State-of-the-art research using the Google Books Ngram Viewer: Improving the method and investigating cultural change

Doctoral thesis for obtaining the

academic degree Doctor of

Natural Sciences (Dr.rer.nat.)

submitted by

Younes, Nadja

at the

Faculty of Sciences

Department of Psychology

Konstanz, 2019

Date of the oral examination: 30.10.2019

1. Reviewer: Prof. Dr. Ulf-Dietrich Reips

2. Reviewer: Prof. Dr. Wolfgang Gaissmaier

Acknowledgements

First and foremost, I would like to deeply thank my supervisor Prof. Dr. Ulf-Dietrich Reips, who did not only give me the opportunity to write a doctoral thesis in Psychology but also the courage to go through a comprehensive aptitude assessment procedure in advance. I am very grateful for Ulf's openness to my research ideas, his constant support, and constructive feedback, as well as all of the additional knowledge, insights, and expertise that I have gained during the process.

In this regard, I would also like to thank my friends and colleagues at the Department of Psychology, particularly, Deborah Canales, Dr. Tim Kuhlmann, and Dr. Vaka Vésteinsdóttir, for creating a superb working environment with lots of inspiring discussions and mutual appreciation.

I furthermore thank Laura Höhner for the fruitful collaboration that led to the third research article of this thesis.

Finally, I would like to express my sincere gratitude to my family – my parents Hassan and Barbara, and my fiancé Torsten – for their never-ending encouragement, unconditional love, and ongoing support.

Konstanz, April 2019 Nadja Younes

Abstract

The present thesis comprises three stand-alone research articles on the Google Books Ngram Viewer – a search engine that allows for a systematic quantification of cultural change over time by charting the relative frequencies of individual words' appearances in a corpus of millions of scanned books. The thesis is concerned with the use of the Google Books Ngram Viewer to examine cross-cultural changes. In addition, each research article emphasizes on an optimization of the research method, previously applied in Google Ngram studies. The thesis thereby aims at introducing state-of-the-art research methods using the Google Books Ngram Viewer.

The first research article made use of the Google Books Ngram Viewer to investigate cultural change in German-speaking countries. By analyzing the development of individualistic and collectivistic German words and by considering additional synonyms, as an improvement on the method, results indicate that Germany and the German-speaking countries around became more individualistic over time. However, a reversal for the time before and during World War II indicates the countries' development towards more collectivistic societies during times of crisis, supporting the argument that cultural changes can be bidirectional.

The second research article presents a comprehensive guideline that aims at increasing the reliability of Google Ngram studies. The guideline builds up on the previously introduced systematic use of synonyms but further suggests the consideration of four additional hands-on procedures, to conduct a composite analysis, and to take the consequences of potential censorship and propaganda into account. Further, in a step-by-step example all of these suggestions were applied by examining religious trends in a cross-cultural setting with a particular focus on the situation during World War II. Results indicate a decline in the importance of religion over time. However, there is again a reversal for the years before and during World War II, particularly for German and Italian words. Based on variances of results, the study

points out that applying and particularly combining the suggested procedures can prevent researchers from deriving wrong and undifferentiated assumptions.

The third research article took advantage of the Google Books Ngram Viewer to investigate changes in the prevalence of positive and negative emotions in China. Based on the application of the procedures suggested in the second research article, and by developing an additional corpus-specific approach that reinforces the reliability of the Chinese corpus, an overall increase in the frequencies of positive Chinese emotion words and a decrease in the frequencies of negative Chinese emotion words is observed. Considering the differences between the perception and the use of positive and negative emotions in collectivistic and individualistic cultures, results support China's theory-compliant development towards a more individualistic society over time.

Taken together, this thesis offers important empirical insights on the cross-cultural development towards increased individualism, even for rather collectivistic countries. Furthermore, each research article provides methodological guidance and valuable advice regarding the use of a sophisticated and robust methodological standard to examine data obtained from the Google Books Ngram Viewer.

Zusammenfassung

Die vorliegende Dissertation besteht aus drei eigenständigen Forschungsartikeln und befasst sich mit dem Google Books Ngram Viewer – einer Suchmaschine, die durch die Darstellung der relativen Häufigkeit einzelner Worte, basierend auf einem Korpus von Millionen Büchern, die systematische Quantifizierung von kulturellem Wandel über die Zeit hinweg erlaubt. Neben der Untersuchung von Annahmen hinsichtlich eines globalen und gesellschaftlichen Wandels beschäftigt sich die vorliegende Arbeit insbesondere mit der Optimierung der bisher in Zusammenhang mit dem Google Books Ngram Viewer angewandten Analysemethoden. Dabei hat die vorliegende Arbeit das Ziel, einen methodischen Standard für Forschungszwecke zu etablieren.

Der erste Forschungsartikel machte sich den Google Books Ngram Viewer zu Nutze, um den gesellschaftlichen Wandel in deutschsprachigen Ländern zu erforschen. Die Ergebnisse der Untersuchung individualistisch und kollektivistisch geprägter deutscher Worte sowie die systematische Berücksichtigung zusätzlicher Synonyme, die eine Verbesserung der soweit vorherrschenden Google Ngram Analysemethodik darstellt, lassen darauf schließen, dass Deutschland und die angrenzenden deutschsprachigen Länder über die Zeit hinweg individualistischer wurden. Ein gegenläufiger Trend zur Zeit des zweiten Weltkriegs deutet allerdings darauf hin, dass in Krisenzeiten wieder vermehrt kollektivistische Werte an Bedeutung gewannen und demnach die Entwicklung gesellschaftlicher Werte bidirektional sein kann.

Der zweite Forschungsartikel präsentiert einen umfassenden Leitfaden, der darauf abzielt, die Reliabilität von Google Ngram Studien zu erhöhen. Der Leitfaden greift die im ersten Forschungsartikel vorgestellte systematische Berücksichtigung von Synonymen auf. Darüber hinaus motiviert er die Verwendung vier weiterer Verfahren und einer Kombinationsanalyse sowie die Beachtung potentieller Konsequenzen von Zensur und Propaganda. Alle Vorschläge des Leitfadens wurden

im Rahmen einer multikulturellen Untersuchung religiöser Trends, unter besonderer Berücksichtigung des zweiten Weltkriegs, beispielhaft angewendet. Die Ergebnisse deuten auf einen kontinuierlichen Rückgang der Bedeutung von Religion im Zeitverlauf hin. Insbesondere für die deutsche und italienische Sprache zeigt sich jedoch abermals ein gegenläufiger Trend zu Zeiten des zweiten Weltkriegs. Auf Basis differierender empirischer Befunde macht der Artikel deutlich, dass durch die Anwendung der im Leitfaden vorgestellten Verfahren Forscher vor dem Treffen falscher und undifferenzierter Annahmen bewahrt werden können.

Der dritte Forschungsartikel machte sich den Google Books Ngram Viewer zu Nutze um Veränderungen in der Verwendung positiver und negativer Emotionen in China zu untersuchen. Basierend auf der Berücksichtigung der im zweiten Forschungsartikel vorgestellten methodischen Verbesserungen sowie der Anwendung eines weiteren Korpus-spezifischen Verfahrens, dass insbesondere die Reliabilität des fehleranfälligeren chinesischen Korpus berücksichtigt, zeigt diese Studie einen Anstieg der Frequenz von positiven chinesischen Emotionswörtern und einen Rückgang der Frequenz von negativen chinesischen Emotionswörtern im Zeitverlauf. In Anbetracht der Unterschiede in der Wahrnehmung und Verwendung positiver und negativer Emotionen in kollektivistisch und individualistisch geprägten Kulturen lassen die Ergebnisse auf einen theoriekonformen Wandel Chinas hin zu einer individualistischeren Gesellschaft schließen.

Zusammenfassend legt die vorliegende Dissertation wichtige empirische Befunde im Hinblick auf eine immer individualistischer ausgerichtete, gesellschaftliche Entwicklung – auch in eher kollektivistisch geprägten Ländern – dar. Darüber hinaus liefert die Arbeit wertvolle Erkenntnisse in Bezug auf die Vorteile der im Rahmen dieser Arbeit dargelegten Analyseverfahren und setzt einen Standard zur verbesserten Untersuchung der durch den Google Books Ngram Viewer bereitgestellten Daten.

Table of Contents

List of Figures

General Introduction

First Research Article: The Changing Psychology of Culture in German-Speaking Countries: A Google Ngram Study

Second Research Article: Guideline for Improving the Reliability of Google Ngram Studies: Evidence from Religious Terms

Third Research Article: The Influence of Social Change and Human Development on Emotions: Evidence from China

List of Tables

First Research Article: The Changing Psychology of Culture in German-Speaking Countries: A Google Ngram Study

Second Research Article: Guideline for Improving the Reliability of Google Ngram Studies: Evidence from Religious Terms

Third Research Article: The Influence of Social Change and Human Development on Emotions: Evidence from China

General Introduction

The advent of computers and the Internet, as part of the Digital Revolution, has spread knowledge globally by transforming even age-old wisdom into digital resources that can be accessed all over the world. In this regard, also research methods, i.e., strategies to generate knowledge, have seen unprecedented opportunities. In particular, cultural product analysis – a research method that infers patterns of culture from tangible public representations of culture such as books, news coverage, or song lyrics (see e.g., Morling & Lamoreaux, 2008 and Lamoreaux & Morling, 2012, for a review) – has been revolutionized. Until recently, analyzing cultural products to unearth cultural change was very costly, particularly in terms of labor. The spread of new technologies such as word coding computer programs and the Internet, however, has allowed researchers to analyze language use in cultural products in more voluminous, accurate, and efficient ways (Twenge et al., 2012a). In this regard, particularly big data tools offer extraordinary opportunities for this field. For instance, if "reading small collections of carefully chosen works enables scholars to make powerful inferences about trends in human thought", imagine the power of analyzing a corpus consisting of millions of books and how much more precise the measurement of the underlying phenomena will be (Michel et al., 2011, p.176).

The Google Books Ngram Viewer – an online search engine that charts word frequencies from a corpus of millions of scanned books – constitutes one of these powerful opportunities. Along with its launch in 2010, the developers argued that one of the tool's main advantages for scientific purposes is the hands-on quantification of cultural change (Michel et al., 2011). In particular, Michel et al. (2011) presented different ways of analysis to subject cultural trends to quantitative investigation and to derive new understanding from the massive aggregation of data. By now, these suggestions have been gratefully embraced by dozens of studies. In this regard, also the present thesis exploits the power of the Google Books Ngram Viewer to examine

cultural change, particularly, by investigating social change and human development based on the concepts of individualism and collectivism (Greenfield, 2009). However, as with any new and evolving methodology, a careful consideration of potential limitations and its consequences on data interpretation is inevitable. By taking the shortfalls of the Google Books Ngram Viewer into account, the present thesis is not only concerned with the use of the Google Books Ngram Viewer to examine cultural change but aims at improving the currently applied method by carefully addressing criticism that has been voiced by the scientific community (see, e.g., Gooding, 2012; Pechenick et al., 2015; Pettit, 2016).

As part of this general introduction, I will first present the Google Books Ngram Viewer in detail. In this regard, I will summarize raised criticism. Second, I will explicate the concepts of individualism and collectivism, particularly with respect to Greenfield's (2009) social change and human development theory that constitutes the main theoretical framework of this thesis. Third, I will point out the theoretical and methodological interconnectedness of the underlying research articles.

The Google Books Ngram Viewer

The Google Books Ngram Viewer charts the relative frequency of a given string of characters as found in a corpus of books, printed between 1500 and 2008. A string of characters – a so-called n-gram – can consist of one word such as "America" (1-gram) or a sequence of several words such as "the United States of America" (5-gram). As of today, 22 different language-corpora are available, assigned to one of two versions – an old and a new version – of the Google Books Ngram Viewer. Whereas the base corpus of the old version consists of more than 5 million books, covering the languages English (e.g., American, British, Fiction, One Million), French, Spanish, German, Chinese, Russian, and Hebrew (Michel et al., 2011), the new version with improved optical character recognition compromises a total of more than 8 million books,

including Italian-language literature (Lin et al., 2012). Figure 1 shows the surface of the Google Books Ngram Viewer as accessible via https://books.google.com/ngrams.

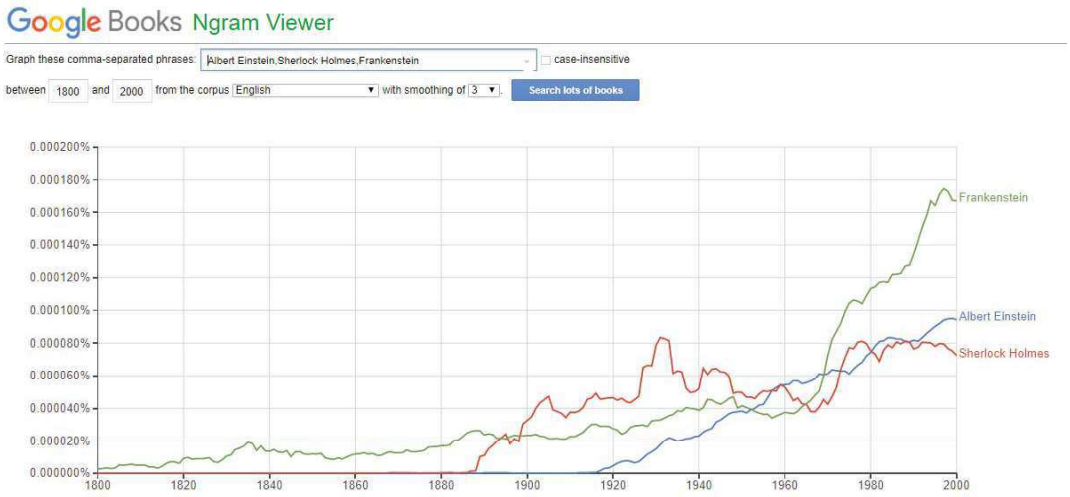

Figure 1. The Google Books Ngram Viewer.

The graphs are generated by using the new English corpus and by dividing the number of the words' (Albert Einstein, Frankenstein, and Sherlock Holmes) yearly appearances by the total number of words in the corpus in the respective year. However, only frequencies of words that appear at least in 40 books across the corpus are considered. In general, users can manually enter search terms (n-grams consisting of up to 5 words) and select the time span (between 1500 and 2008), the underlying corpus (by language, while "(2009)" indicates affiliation to the old version and no date affiliation to the new version of the corpus), whether graphs shall be displayed considering raw data or moving averages (i.e., with or without smoothing, where the selected number indicates the number of years over which an average frequency is computed), and the preferred search type (i.e., case-insensitive). Further forms of advanced searches include wildcard search (by putting a * in place of a word the Google Books Ngram Viewer charts the top ten substitutions), inflection search (by appending _INF to a word its modifications are displayed), part-of-speech-tags (by appending tags such as _VERB, _NOUN, _ADV, etc. the Google Books Ngram

Viewer distinguishes between verbs, nouns, or adjectives etc.), and compositions (by using operators such as "+", "-", "/", "*", ":" searches are combined).[1]

Based on its easy access to big data on cultural products, related to various cultures and centuries, at almost no labor costs, the Google Books Ngram Viewer provides far-reaching opportunities for research purposes. However, there are several shortfalls that require a cautious interpretation of data. Main points of criticism relate to insufficient optical character recognition regarding the digitization of text. For example, semantic scanning errors that often affect words including the letters "f" and "s" might lead to a systematically biased computation of word frequencies. Along these lines also metadata that is incorrectly populated, i.e., wrong publication dates or genres, can lead to similar biases. Additional critique refers to the corpora's large amount of scientific literature, owed to the fact that most books were drawn from university libraries. Because books are not chosen based on transparent criteria such as popularity, single authors might be able to influence the corpus with specific words and phrases. As a consequence, these words may rather represent an elite minority than a board population and therefore bias assumptions on the general development of cultural change over time (see, e.g., Gooding, 2012; Pechenick et al., 2015; Pettit, 2016 for a review on the Google Books Ngram Viewer's limitations for research purposes).

Social Change and Human Development

This thesis uses the Google Books Ngram Viewer to test theoretical predictions on cultural change as mainly derived from the social change and human development theory by Greenfield (2009). The theory links social changes to developmental changes by arguing that shifts in sociodemographic factors alter cultural values and developmental patterns. The theoretical starting points for Greenfield's (2009) theory are the terms *Gemeinschaft* (community) and *Gesellschaft* (society), introduced by the

[1] For further and more detailed information visit https://books.google.com/ngrams/info.

German sociologist Tönnies (1887/1957). Both concepts are characterized by different sociodemographic dimensions. A *Gemeinschaft*, for example, describes a rural, simple, and small scale ecology with low technological standards, limited educational access, and a comparably poor economy that pursues subsistence. A *Gesellschaft*, in contrast, refers to an urban, large scale, and complex ecology with high technological standards, a well-developed educational system, and a rather wealthy and commercial economy. Moreover, whereas relations in a *Gemeinschaft* are characterized by being permanent and mostly lifelong, social relations in a *Gesellschaft* are less enduring and more often fleeting. The classical concepts of "collectivism and individualism summarize social adaptations to the two types of environment" (Greenfield, 2009, p.401).[2] In particular, collectivistic values, behavior, and psychology are adapted to *Gemeinschaft* environments whereas individualistic values, behavior, and psychology are adapted to *Gesellschaft* environments. Although previous research has already recognized that sociodemographic shifts move values from more collectivistic to more individualistic manifestations (see, e.g., Triandis, 1993 and Hofstede, 2001), Greenfield's (2009) theory provides an in-depth understanding of the transformation and the development of cultural changes based on the two distinct environments.

Using the Google Books Ngram Viewer, Greenfield (2013) and Zeng and Greenfield (2015) empirically tested the assumptions of the social change and human development theory for the US, UK, and China. In particular, both studies examined the influence of dominant sociodemographic trends such as the movement from rural environments, informal education, and a subsidence economy to urban environments, formal education, and commerce on values and behavior. By a visual and quantitative examination of frequencies of individualistic (*Gesellschaft*-indexed) and collectivistic (*Gemeinschaft*-indexed) American and British English as well as Chinese words,

[2] Although Greenfield (2009) suggests that "the terms individualism and collectivism do not adequately describe cognitive adaptations to the two types of environment [and that] the ecologies have greater explanatory power" (p.403), within the scope of this thesis, the terms *Gemeinschaft* and collectivism as well as the terms *Gesellschaft* and individualism are used interchangeably.

respectively, both studies noted a shift towards increased individualism over time. To reinforce the relationship between changes in ecology and values, Zeng and Greenfield (2015) placed special emphasis on an additional investigation of the correlations between word frequencies and changes in the sociodemographic factors urbanization, wealth, and formal schooling.

Present Research

The first research article, published in the *International Journal of Psychology*, builds on Greenfield's (2009) theoretical predictions by replicating Greenfield's (2013) study using the German Google Ngram corpus. Further, we systematically considered synonyms as an improvement on the method. With this study, we aimed at providing further evidence for the general universality of Greenfield's (2009) social change and human development theory. In addition, we strove for showing that cultural changes can be bidirectional. In particular, due to the Germans' special situation during World War II, we expected a reversal from individualistic to collectivistic values in this period. To conduct the analysis, we translated Greenfield's (2013) English words to German. We then used the Google Books Ngram Viewer to compare frequencies of individualistic and collectivistic German words between the years 1800 and 2000, with a special focus on World War II. Because the criteria for Greenfield's (2013) explicit choice of words were not clearly indicated and the subjective choice of a specific word may influence the overall results (e.g., by substituting Greenfield's (2013) collectivistic word "worship" with the synonym "ritual" an increase instead of a decrease is observed), we additionally selected the first three synonyms for any translated German word, as listed in a thesaurus. Descriptive as well as quantitative investigations show an overall increase in the frequency of individualistic words, whereas the frequency of collectivistic words decreased. As expected, during the years of war there is a reversal that indicates the development towards a more collectivistic culture at this time. Overall, our findings suggest that based on the Germans'

continuous movement from rural to urban areas, also the society became more individualistic over time.

The second research article, published in *PLOS ONE*, constitutes a guideline on how to conduct Google Ngram studies and increase the reliability of results. The guideline proposes five procedures, namely the use of (I) multiple language corpora, (II) an English fiction corpus, (III) word inflections, (IV) synonyms, and (V) a standardization procedure that accounts for Google Ngram's data influx while treating all investigated words equally. In addition, it provides a clear roadmap on how to use the procedures in combination and a way to deal with potential influences of censorship and propaganda. The procedures are illustrated in a step-by-step example, investigating the development of the importance of religion between the years 1900 and 2000. By comparing the results obtained through the different methods, the study points out that the application and especially the combination of our suggested procedures can increase the reliability of results and prevent researchers from deriving wrong and undifferentiated assumptions. The study thereby builds on the first research article where we already introduced the systematic consideration of synonyms as an improvement on the method. We incorporated this valuable procedure in the guideline and provided suggestions on how to apply it in a cross-cultural setting. Further, by exemplarily studying the development of the importance of religion, the second research article focuses likewise on the development of social change, particularly the decrease of collectivistic values. In line with theoretical predications, we observe a drop in the importance of religion over time. However, in accordance with the findings of the first study, there is a reversal during World War II which supports the argument that cultural changes can be bidirectional.

The third research article studies the development of the expression of emotions in China over time by focusing on the Chinese Google Ngram corpus. In close relation to the first article, we use Greenfield's (2009) social change and human development theory as the underlying conceptual framework and apply – whenever

possible – the procedures suggested in the guideline of the second research article. In this study, we specifically focus on China for the following two reasons. First, as a traditional collectivistic society, China's recent transformative changes in its social ecology have required its people to face great challenges concerning their long-established values and behavior. Thus, we expect to observe a development towards individualism in a more distinct way. In particular, based on the differences in the use and perception of positive and negative emotions in individualistic and collectivistic societies, we predict that the frequency of positive Chinese emotion words increased, whereas the frequency of negative Chinese emotion words decreased between the years 1970 to 2008. Second, because the Chinese language is not based on a Latin alphabet, we show how the guideline of the second research article can be applied and extended for studies that use the more error-prone Chinese corpus. In this respect, we developed and applied an additional corpus-specific procedure that researchers can consider to reinforce the reliability of results. In particular, we investigated whether corpus-specific biases such as scanning errors may lead to the occurrence of similar frequency patterns for a set of emotion words and a further set of optically similar but unrelated expressions. Overall, results do not seem to be influenced by any corpus-specific biases. In addition, based on China's urbanization, the increase in the level of wealth, and the trend towards formal education, the findings suggest that the country's development towards a more individualistic society results in an increase in the use of positive emotions and a decrease in the use of negative emotions.

First Research Article:

The Changing Psychology of Culture in German-Speaking Countries:

A Google Ngram Study

Nadja Younes and Ulf-Dietrich Reips

University of Konstanz

International Journal of Psychology

International Journal of Psychology, 2018
Vol. 53, No. S1, 53–62, DOI: 10.1002/ijop.12428

The changing psychology of culture in German-speaking countries: A Google Ngram study

Nadja Younes[1] and Ulf-Dietrich Reips[1,2]

[1]Graduate School of Decision Sciences, University of Konstanz, Konstanz, Germany
[2]Department of Psychology, University of Konstanz, Konstanz, Germany

This article provides evidence for the long-term affiliation between ecological and cultural changes in German-speaking countries, based on the assumptions derived from *social change and human development theory*. Based on this theory, the increase in urbanisation, as a measure of ecological change, is associated with significant cultural changes of psychology. Whereas urbanisation is linked to greater individualistic values and materialistic attitudes, rural environments are strongly associated with collectivistic values like allegiance, prevalence of religion, and feelings of belonging and benevolence. Due to an increase in the German urbanisation rate over time, our study investigates whether Germany and the German-speaking countries around show the presumed changes in psychology. By using Google Books Ngram Viewer, we find that word frequencies, signifying individualistic (collectivistic) values, are positively (negatively) related to the urbanisation rate of Germany. Our results indicate that predictions about implications of an urbanising population for the psychology of culture hold true, supporting international universality of the *social change and human development theory*. Furthermore, we provide evidence for a predicted reversal for the time during and after World War II, reflecting Nazi propaganda and influence.

Keywords: Social change; Cultural change; Internet science; Google Ngram Viewer.

Tönnies (1887/1957) presented a normative theory that contained ideal types of social organisations. His conception of a social systems' nature relies on the distinction between *Gemeinschaft* (community) and *Gesellschaft* (society). *Gemeinschaft* is characterised by a rural environment, simple face-to-face relationships, low levels of technology, limited education and a life of need rather than of wealth. On the other hand, he defines *Gesellschaft* as an urban environment—a modern society with high levels of technology and wealth. Whereas the bonds of family and religion are weaker, economies are oriented not only towards subsistence but also towards commercial growth. Education becomes important and takes place at schools. Tönnies' (1887/1957) understanding of *Gemeinschaft* and *Gesellschaft* built the cornerstones of Greenfield's (2009) *social change and human development theory*, which suggests that changes in sociodemographic ecologies such as the global increase in the rate of urbanisation impacts on cultural values and influences learning and development.

Although previous research has recognised and discussed the importance of macro-level social changes on human development (cf. Silbereisen & Chen, 2010 for a review), Greenfield (2013) and Zeng and Greenfield (2015) provide novel empirical evidence for their hypothesis of urbanisation as a major driving factor of cultural change. In particular, they show that over time, word frequencies representing *Gesellschaft* (i.e., individualistic) values increase with growing urban population in the United States, United Kingdom and China whereas *Gemeinschaft* (i.e., collectivistic) values decrease. In line with Grossmann and Varnum (2015), who suggest that the development of *Gesellschaft* values is less robust to changes in urbanisation, Zeng and Greenfield (2015) show that besides urbanisation other sociodemographic factors such as shifts in socioeconomic status, for example, measured by personal wealth, are also associated with a cultural change.

Within the last centuries, the rate of urbanisation has increased drastically and globally. Related to the

Correspondence should be addressed to Nadja Younes, Graduate School of Decision Sciences, University of Konstanz, Universitätsstraße 10, Box 141, D-78457 Konstanz, Germany. (E-mail: nadja.younes@uni-konstanz.de).

Nadja Younes did the data collection and analyses. Ulf-Dietrich Reips added to the methodology. Nadja Younes and Ulf-Dietrich Reips designed the study and wrote this article jointly. The authors would like to thank Patricia Greenfield and two anonymous referees for their valuable comments. Further, we would also like to thank Torsten Twardawski for helpful suggestions with respect to the analyses. Nadja Younes is grateful for the support of the Graduate School of Decision Science (University of Konstanz).

industrialisation in Europe and the United States, Bähr (2007) documents massive migration from rural to urban regions leading to a strong increase in the urbanisation rate. Although urbanisation arrived later in some countries, it has become a global trend. In this article, we investigate whether Germany and the German-speaking regions in the countries around moved from *Gemeinschaft* towards *Gesellschaft* during the last two centuries. Moreover, due to "Gleichschaltung" (establishing a system of totalitarian control and coordination, "from the economy and trade associations to the media, culture and education," Strupp, 2013) of personal beliefs, desires and interests, and the sustained propaganda of collectivistic virtues such as responsibility, initiative and courage by National Socialist (i.e., Nazi) ideology (Föllmer, 2010), we predict a reversal at the time of World War II. In particular, given that changes can be bidirectional and shifts from *Gesellschaft* back to *Gemeinschaft* are possible (Inglehart & Baker, 2000; Park, Twenge, & Greenfield, 2014), we propose that although there is an overall trend from *Gemeinschaft* to *Gesellschaft*, during and shortly after World War II the German society moved from *Gesellschaft* towards *Gemeinschaft*.

Cultural values are reflected in terms used in writings. Thus, to provide evidence for the long-term cultural change in values, we use the Google Books Ngram Viewer (https://books.google.com/ngrams). By searching millions of digitalised books and quantitatively analysing changes in word frequencies, this tool has been designed to investigate cultural trends and salient semiotic developments (Michel et al., 2011).

Social change and human development

Classical concepts of individualism and collectivism characterise individualistic cultures as being autonomous from their in-groups and prioritising the achievement of personal goals, whereas people in collectivistic cultures are considered to be more interdependent with their in-groups and more concerned with relationships (cf. Triandis, 2001 for a review). Although previous research has noted the impact of sociodemographic influences with respect to increasing individualism (cf. Hofstede, 2001; Triandis, 1993), Greenfield's (2009) theory of *social change and human development* provides an in-depth explanation for the transformation and the development of cultural values based on different ecological types. In particular, Greenfield (2009) considers *Gemeinschaft* and *Gesellschaft* as prototypical sociocultural ecologies and defines them with contrasting characteristics representing individualism and collectivism, for example, technology and prosperity levels, religious awareness, economic systems, social relations and educational practices.

In line with Tönnies (1887/1957), Greenfield (2009) assigns values such as, for example, obligations and conscientiousness, general welfare, and a lifestyle according to religious duties (cf. Chalfant & Heller, 1991) to *Gemeinschaft*, that is, the focus lies on the community. In contrast, feelings, desires and personal aspects are more important in a *Gesellschaft* (cf. Inglehart & Baker, 2000; Kraus, Piff, Mendoza-Denton, Rheinschmidt, & Keltner, 2012), that is, the focus lies on the individual. Greenfield (2009) points out that in a *Gesellschaft*, individuals of lower social status as well as those living in rural areas, are rather associated to a *Gemeinschaft* than a *Gesellschaft* with its respective values. Therefore, it is assumed that individuals of lower status are more benevolent, and socially dependent than people of higher status (Kraus et al., 2012). Although shifts from *Gesellschaft* to *Gemeinschaft* are possible, on a whole, cultural changes are supposed to move in the direction of *Gesellschaft* and therefore lead to more materialism and individualism, based on an increase in urbanisation, wealth, technological progress and better educational systems (Greenfield, 2009; Park et al., 2014; Uhls & Greenfield, 2011).

METHODOLOGY AND PROCEDURE

Greenfield (2013) uses the level of urbanisation as a proxy for the whole complex of *Gesellschaft* factors and provides evidence that for the United States and United Kingdom, the movement from rural to urban environments is a potential driver of a change in values. As the German urbanisation rate highly correlates with the rates of the United States and United Kingdom (cf. Bähr, 2007), we assume a similar change in values for Germany and the German-speaking regions in neighbouring countries over time. In order to test this assumption, we examine the change in word frequencies in German language books to measure changes in cultural values, using the Google Books Ngram Viewer.

Research methods have seen unprecedented improvements and opportunities from the Internet (cf. Reips, 2008 for a review). In 2004, Google Books constructed a corpus of digitalised texts containing approximately 15 million of the estimated 129 million books published between 1500 and 2008. These books, mainly scanned from large libraries around the world (e.g., Harvard University Library, New York Public Library, Stanford University Library, University of Michigan Library, University of Oxford Library and The Bavarian State Library), contain approximately 500 billion words in seven languages, namely Chinese, English, French, German, Hebrew, Russian and Spanish. For our study of German terms, we can rely on a corpus of around 37 billion words. Whereas in early centuries only a few books per year were available, the main publication time span ranges from 1800 to 2000. The Google Books Ngram Viewer computes word

TABLE 1
Overview of value-carrying words representing *Gemeinschaft* and *Gesellschaft*

	Greenfield (2013)		*German translation*	
Gemeinschaft	obliged	duty	versprechen	Pflicht
	give	benevolence	geben	Güte
	act	deed	handeln	Handlung
	obedience, authority, pray, belong, power, worship, join, conformity		Gehorsam, Amtsgewalt, beten, gehören, Macht, Verehrung, zugesellen, Gleichförmigkeit	
Gesellschaft	choose	decision	auswählen	Entscheidung
	get	acquisition	bekommen	Kauf
	feel	emotion	spüren	Emotion
	individual, unique, child, self, personal, ego, baby, special		individuell, einzigartig, Kind, Selbst, persönlich, Ego, Baby, speziell	

Note: Column I presents *Gemeinschaft*-indexed and *Gesellschaft*-indexed words selected by Greenfield (2013). Column II displays German words, equivalent to Greenfield's (2013) *Gemeinschaft*-indexed and *Gesellschaft*-indexed words.

frequency by dividing the number of respective words' yearly instances by the total number of words in the respective year (Michel et al., 2011).

In order to obtain a qualitatively equal sample of terms as in Greenfield (2013), we also choose the time span 1800–2000. Greenfield (2013) selects six verbs (or adjectives) and six-related noun synonyms. Furthermore, she uses eight additional words as robustness, each indexing either *Gemeinschaft* or *Gesellschaft*. We translate each of Greenfield's (2013) 28 English value-indexed words into German using the first adequate entry of the online dictionary *Pons*.[1] Following Zeng and Greenfield (2015), we additionally asked two native German speakers, a 31-year-old male and a 56-year-old female to judge each translation on its usage contexts and underlying meaning. Without discussion, they both validated that the semantic interpretations of the German words do not differ from the English originals selected by Greenfield (2013).

Table 1 presents an overview of words selected by Greenfield (2013) and their translations to German. However, as Greenfield's selection of *Gemeinschaft*-indexed and *Gesellschaft*-indexed words can be argued to lack a clear criterion,[2] we improve on the method by collecting the first three one-word synonyms for the semantically most adequate grouping listed in *Duden Synonymwörterbuch* (2014) for each translation of Greenfield's (2013) robustness words (obedience, authority, pray, belong, power, worship, join, conformity individual, unique, child, self, personal ego, baby, special). Furthermore, we model and test for a temporal trend in the Google Ngram word frequency data by using various regression models. Finally, we follow Zeng and Greenfield (2015) and establish a link between the change in values and ecological conditions by computing correlations between the annual frequency of German words and the urbanisation rate in the corresponding year.

RESULTS

The first contrasting words and underlying values that were examined by Greenfield (2013) are "obliged" and "duty" (*Gemeinschaft*-indexed, i.e., collectivistic) versus "choose" and "decision" (*Gesellschaft*-indexed, i.e., individualistic). While the first two words show a decrease over time the last two words exhibit an increase. In order to examine the respective value change for German, we select "Pflicht" (duty) and "versprechen" (obliged) as well as "Entscheidung" (decision) and "auswählen" (choose).

Figure 1 shows the difference in word frequency over the time span 1800–2000. In line with the findings of Greenfield (2013), the relative word frequency of "Entscheidung" (decision) almost tripled, whereas "Pflicht" (duty) displays an approximately fivefold decrease over time. Similar results are obtained for the verbs "versprechen" (obliged) and "auswählen" (choose) but we refrain from including their word frequencies in the same figure due to scaling differences.

The difference between getting something for oneself (*Gesellschaft*-indexed) and contributing to common good (*Gemeinschaft*-indexed) is represented by the words "get" and "acquisition" as well as "give" and "benevolence," respectively. Greenfield (2013) observed a decrease for

[1] http://de.pons.com/. Exceptions are "versprechen," "Handlung," "Amtsgewalt," "Gleichförmigkeit" and "Kauf." We use "versprechen" (promise) rather than "verpflichten" (a direct translation of obliged) because "verpflichten" is mostly used in a military context. For "deed" the word "Tat" cannot be selected because it has a second meaning as the past tense of "tun" (do). For this reason, we choose the second translation "Handlung." For "conformity" we only obtain "Konformismus." As this technical term is not used steadily over time, we use the *Duden* dictionary (http://www.duden.de/suchen/englisch/conformity) to recheck and obtain "Gleichförmigkeit." For "acquisition" with respect to purchase we receive "Anschaffung," a rarely used term. Due to a higher frequency we use the word "Kauf."

[2] Greenfield (2013) states that words were chosen according to the following criteria: theory relevance, high frequency and narrow range of semantic interpretations.

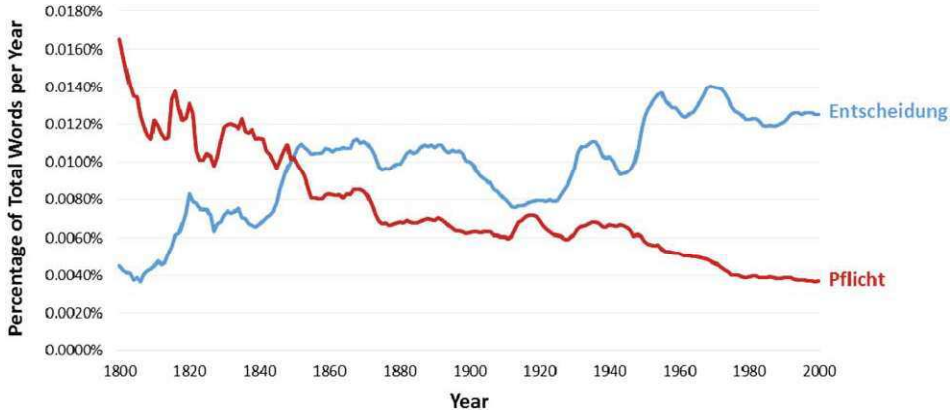

Figure 1. Frequency of the words "Pflicht" (duty) and "Entscheidung" (decision) in the Google corpus of German books from the years 1800 to 2000. The graph was made with the Google Books Ngram Viewer (Michel et al., 2011), with the standard smoothing of 3.

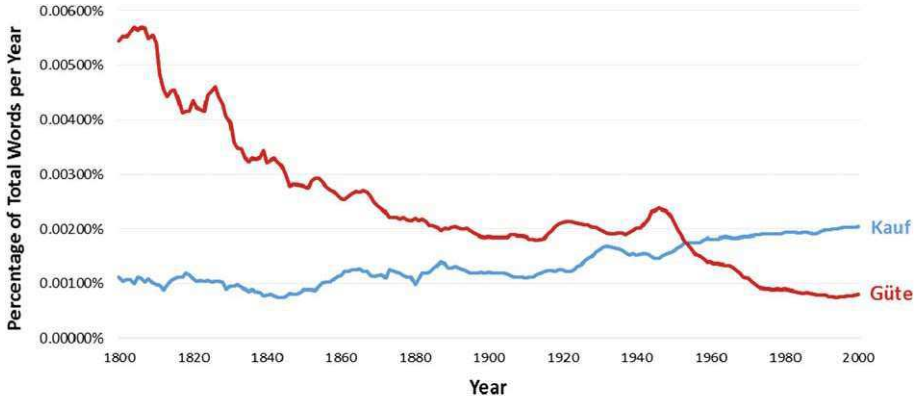

Figure 2. Frequency of the words "Güte" (benevolence) and "Kauf" (acquisition) in the Google corpus of German books from the years 1800 to 2000. The graph was made with the Google Books Ngram Viewer (Michel et al., 2011), with the standard smoothing of 3.

the first and an increase for the latter terms for the United States and United Kingdom. The respective German terms are "Güte" (benevolence) and "geben" (give) as well as "Kauf" (acquisition) and "bekommen" (get).

Figure 2 shows the *Gesellschaft*-indexed word "Kauf" (acquisition) increased over time, whereas the *Gemeinschaft*-indexed "Güte" (benevolence) decreased drastically. However, between 1940 and 1950, an increase in "Güte" (benevolence) is observed (and a corresponding slight decline of "Kauf"). This reversal of cultural change likely occurs due to the end of World War II in 1945 and Nazi ideology. While conformity and equality were highly propagandised during the Nazi Regime, to return to normal life, benevolence among people was of utmost importance after the war. One example is *Trümmerfrauen* (ruin women), who helped clear and reconstruct bombed cities as part of a new community. In line with these findings, Greenfield (2013) found a decline in "get" for the United States. She also assumed that this might represent a decrease in self-interest, based on the entering of the War in 1940. Although "bekommen" (get) does not increase as strong as "Kauf" (acquisition), overall, "bekommen" (get) and "geben" (give) show similar patterns as the respective nouns. Again, due to scaling differences, verbs and nouns are not displayed simultaneously.

Furthermore, Greenfield (2013) used "act" and "deed" (*Gemeinschaft*-indexed) as well as "feel" and "emotion" (*Gesellschaft*-indexed) to contrast that in *Gemeinschaft* individuals appreciate the social world associated with action or behaviour, whereas in a *Gesellschaft* the inner psychological process is focused on. To replicate the findings in German, we examine "Handlung" (deed) and "handeln" (act) as well as "Emotion" (emotion) and "spüren" (feel). In line with Greenfield (2013), who reported a rise in "feel" and a strong decline in "act," our German translations display a consistent pattern. "Handlung" (deed) declined fourfold compared to the approximate 200-fold increase in "Emotion." For "handeln" (act) and "spüren" (feel) we find similar results. Whereas "handeln" (act) declined over time, "spüren" (feel) rose and reached its peak around 1940. In 2000,

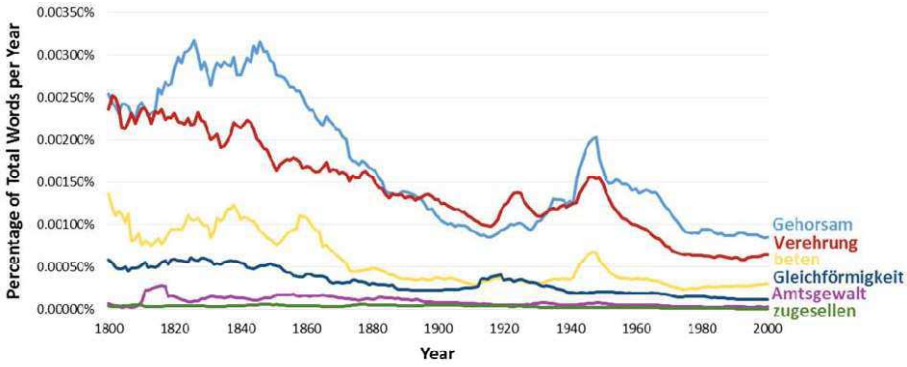

Figure 3. Frequency of words representing *Gemeinschaft*-indexed values from the years 1800 through 2000. The graph was made with the Google Books Ngram Viewer (Michel et al., 2011), with the standard smoothing of 3.

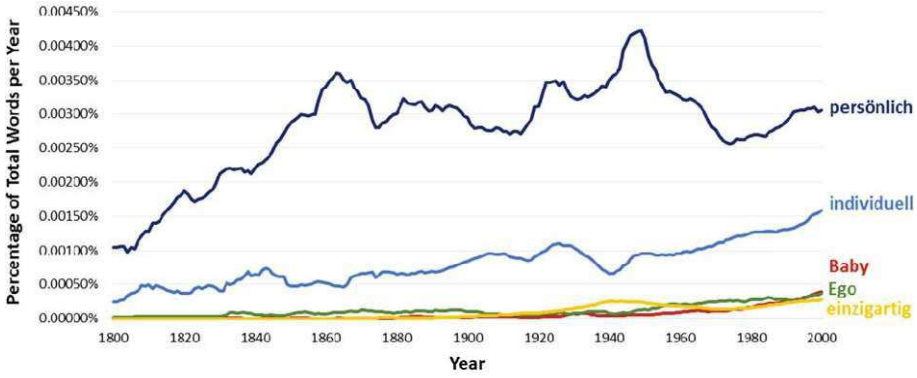

Figure 4. Frequency of words representing *Gesellschaft*-indexed values from the years 1800 through 2000. The graph was made with the Google Books Ngram Viewer (Michel et al., 2011), with the standard smoothing of 3.

"spüren" (feel) exhibited a more than fivefold increase since 1800.

Finally, in order to highlight the "child"-centred socialisation to a "unique," "individual," "self" in a *Gesellschaft*, the quoted words were investigated. To test for robustness, Greenfield (2013) selected "baby," "personal," "special" and "ego" (*Gesellschaft*-indexed). An overall rise was observed, whereas words such as "obedience," "authority," "belong" and "pray" with the synonyms "conformity," "power," "join" and "worship" (*Gemeinschaft*-indexed) were reported to display an overall decrease over time. Thus, due to an increase in rural population, we expect the frequency of words with respect to the individual, the center of a *Gesellschaft*, to rise. Frequencies of words related to the community, religion and belonging are expected to decline. We replicate Greenfield's (2013) results by using the words "individuell" (individual), "einzigartig" (unique), "Kind" (child) and "Selbst" (self) with "persönlich" (personal), "Ego" (ego), "Baby" (baby) and "speziell" (special).

Figure 3 presents the predicted decrease in *Gemeinschaft* values by a decline in the frequency of the words "Gehorsam" (obedience), "Verehrung" (worship), "beten" (pray), "Gleichförmigkeit" (conformity), "Amtsgewalt" (authority) and "zugesellen" (join). A similar picture is given for the synonyms "Macht" (power) and "gehören" (belong).[3] However, as predicted due to Nazi propaganda in World War II and in contrast to the United States and United Kingdom, an increase in word frequencies for the majority of examined terms is observed for the years between 1941 and 1945. On the one hand, Nazi ideology tried to create a feeling of affiliation, on the other hand total obedience and worship (Verehrung) of Hitler was demanded (Welch, 2004). The observable increase in "beten" (pray) might likely be associated with people writing about God and faith to cope with the atrocities and hardships of dictatorship and war.

Figure 4 displays results for words representing *Gesellschaft* values, confirming the results of Greenfield (2013) by presenting a continuous increase for words associated with the development of a unique individual.

To review whether our results are indeed transferable to the *social change and human development* theory by Greenfield (2009), we conduct a set of robustness checks.

[3]"Macht" (power) and "gehören" (belong) as well as "Kind" (child), "speziell" (special) and "Selbst" (self) are not displayed due to scaling differences. However, overall results do not change.

TABLE 2
Overview of value-carrying words representing *Gesellschaft* and *Gemeinschaft*

Panel A: *Gesellschaft* *Greenfield (2013)*	*German translation*		*Synonyms*	
personal	persönlich	eigen	individuell	privat
individual	individuell	eigen	persönlich	privat
baby	Baby	Kind	Neugeborenes	Säugling
child	Kind	Baby	Neugeborenes	Säugling
ego	Ego		—	
self	Selbst		—	
special	speziell	außergewöhnlich	eigenwillig	extravagant
unique	einzigartig	außergewöhnlich	beispiellos	einmalig
Panel B: *Gemeinschaft* *Greenfield (2013)*	*German translation*		*Synonyms*	
obedience	Gehorsam	Ergebenheit	Folgsamkeit	Fügsamkeit
authority	Amtsgewalt		—	
pray	beten	bitten	flehen	
belong	gehören	angehören	zugehören	
power	Macht	Ansehen	Autorität	Einfluss
worship	Verehrung	Anbetung	Glaube	Kult
join	zugesellen		—	
conformity	Gleichförmigkeit		—	

Note: Column I presents *Gesellschaft*-indexed words (Panel A) and *Gemeinschaft*-indexed words (Panel B) selected by Greenfield (2013). Column II displays German words, equivalent to Greenfield's (2013) *Gesellschaft*-indexed words (Panel A) and *Gemeinschaft*-indexed words (Panel B). Columns III–V (Panels A and B) provide synonyms for each German equivalent, obtained via *Duden Synonymwörterbuch* (2014).

TABLE 3
Descriptive statistics of synonyms obtained from *Duden Synonymwörterbuch* (2014)

	Gesellschaft	*Gemeinschaft*	*Total*
Number of Greenfield's (2013) robustness words	8	8	16
Number of *Duden* synonyms	18 (13)	13 (13)	31 (26)
Number of synonyms *with* predicted change in frequency	15 (11)	9 (9)	24 (20)
Number of synonyms *without* predicted change in frequency	3 (2)	4 (4)	7 (6)

Note: The number of unique words is put in parentheses.

First, we collect three synonyms for each translation of Greenfield's (2013) robustness words (obedience, authority, pray, belong, power, worship, join, conformity individual, unique, child, self, personal ego, baby, special). We obtain our synonyms by using the first three one-word synonyms for the semantically most adequate grouping listed in *Duden Synonymwörterbuch* (2014). We avoid choosing words that appear to be colloquial, dialect, humourous or obsolete. Based on this procedure, we exclude "selber" as a form of "persönlich" (personal), as well as "Bébé," "Wickelkind" and "Kleinstkind" as synonyms for "Baby" (baby). The main advantage of this approach is that the selection of words is performed mechanically, using the major German dictionary of synonyms. Hence, the analysis is robust to the choice of translation.

Table 2 provides an overview of Greenfield's (2013) *Gesellschaft*-indexed robustness words (Panel A) and *Gemeinschaft*-indexed robustness words (Panel B), as well as their German translations, and respective German synonyms obtained from *Duden Synonymwörterbuch* (2014).[4]

Table 3 shows descriptive statistics of the change in frequency for all synonyms presented in Table 2. For Greenfield's (2013) *Gesellschaft*-indexed robustness words (personal, individual, baby, child, ego, self, special and unique), we obtain 18 synonyms in total, using the first three synonyms listed in *Duden*

[4]For the *Gesellschaft*-indexed words "Ego" (ego) and "Selbst" (self) and the *Gemeinschaft*-indexed words "Amtsgewalt" (authority), "zugesellen" (join) and "Gleichförmigkeit" (conformity) *Duden Synonymwörterbuch* (2014) lists no synonyms. For "Kind" (child), one of the three proposed synonyms by *Duden Synonymwörterbuch* (2014) is "Wurm," which we excluded because it mainly signifies the animal worm. Further, we skip "ausgefallen" as a synonym for "speziell" because it is often translated with "out of order."

TABLE 4
Results of the regression analysis

	t	*p*	*ß*	*F*	*df*	*p*	*Adj.* R^2
Model I				764.61	199	.000	.7924
Year	27.65	.000	4.32e-06				
Model II				1797.89	199	.000	.8998
Year	−42.40	.000	−.0000125				
Model III				4925.56	199	.000	.9610
Year	−70.18	.000	−.0000168				
Model IV				1841.42	197	.000	.9650
Year	−73.27	.000	−.000017				
Nazi Regime	−3.39	.001	−.0664793				
Year × Nazi Regime	3.40	.001	.0000343				

Note: The dependent variable for Model I (and II) is the mean word frequency of *Gesellschaft*-indexed (and *Gemeinschaft*-indexed) words. The dependent variable for Models III and IV is the mean word frequency of *Gemeinschaft* minus *Gesellschaft*. $N = 201$. Nazi Regime is a dummy equal to one if the year is between 1935 and 1950 and zero otherwise.

Synonymwörterbuch (2014) wherever possible. However, as her *Gesellschaft*-indexed robustness words are semantically very similar, only 13 of 18 synonyms are unique words. Out of these words, only "Kind" (child) and "eigen" (personal/individual) do not show an overall increase. Nevertheless, substituting "eigen" with one of its inflections, that is, "eigene," "eigener," or "eigenes," the predicted increase over time is observed. For Greenfield's (2013) *Gemeinschaft*-indexed robustness words (obedience, authority, pray, belong, power, worship, join, conformity), we also obtain 13 different synonyms. "Autorität" (power) and "Kult" (worship) show an overall increase with peaks in 1950. Since that time "Autorität" (power) displays a steep decline, whereas "Kult" (worship) has been rising again since 1990. On the other hand, "Fügsamkeit" (obedience) and "angehören" (belong) reach their peaks between 1860 and 1880. From that time on, both words show a steady decline. Overall, we observe the predicted theory-based increase for *Gesellschaft*-indexed and decrease for *Gemeinschaft*-indexed synonyms over time.

Second, in addition to our descriptive analysis, we go further by modelling and testing for a temporal trend in the Google Ngram word frequency data, that is, by identifying changes in values over time. In particular, we access annual raw data for the years 1800–2000 of all translated words and *Duden Synonymwörterbuch* (2014) synonyms.[5] With this data, we run various regression models displayed in Table 4.

The regression result of Model I shows that the mean frequency of *Gesellschaft*-indexed words (dependent variable) significantly increases with time (independent variable), whereas the mean frequency of *Gemeinschaft*-indexed words decreases (Model II). Model III shows that the difference between the mean frequency of *Gemeinschaft*-indexed words and the mean frequency of *Gesellschaft*-indexed words decreases significantly with respect to the year indicating a change in values off from *Gemeinschaft* towards *Gesellschaft* over time. In Model IV, we additionally include a dummy variable that is equal to one if the year indicates the time between 1935 and 1950 (lagged time of Nazi Regime)[6] and zero otherwise, and an interaction term of the year and this dummy variable. The interaction term evidences that there is a significant countermovement in values during the Nazi Regime. This manifests the reliance of the method, that is, changes in word frequency seem to indicate changes in social and cultural values, as well as predicted changes for the German-speaking population.

Third, in order to establish the link between the change in values and changes in ecological conditions, we follow Zeng and Greenfield (2015) and compute correlations between the annual frequency of German words covered by Table 1 and the size of the urban population in the corresponding year. For this purpose, we obtain data on the urban population of Germany between 1961 and 2000 from the World Bank.[7]

Table 5 (column 1) shows the correlation coefficients for German *Gesellschaft*-indexed and

[5]We add one inflection of "eigen" namely "eigene." Furthermore, by choosing the time span 1800–2000, we keep our sample well comparable to Greenfield (2013). However, results remained unchanged for the time span 1800–2008.

[6]We investigate a lagged time horizon to account for the fact that there might be a delay in value development and publication during the Nazi regime. Results remain robust for the exact time period of the Nazi regime, 1933–1945.

[7]The data (http://data.worldbank.org) roughly covers 40 years of our sample period. This should be long enough to identify a structural relationship between urbanisation and social values. Furthermore, because Grossmann and Varnum (2015) argue that urbanisation might be less robustly associated with a change in individualism we calculate (lagged) correlations between wealth increase as a proxy of socioeconomic status and German word frequencies representing *Gesellschaft* and *Gemeinschaft* values. Results remained stable and can be requested from the authors.

TABLE 5
Correlations between urban population and German word frequencies representing *Gesellschaft* and *Gemeinschaft* values

Panel A: Gesellschaft	*Urban population*	*Urban population time lag one year*	*Urban population time lag two years*
decision—Entscheidung	−.1327	−.1947	−.2858*
choose—auswählen	.7688***	.7754***	.7785***
get—bekommen	.7782***	.7629***	.7448***
acquisition—Kauf	.9370***	.9419***	.9323***
feel—spüren	.0370	.0781	.1335
emotion—Emotion	.8741***	.8664***	.8616***
individual—individuell	.9116***	.9225***	.9328***
unique—einzigartig	.6954***	.6954***	.6982***
child—Kind	.6775***	.6572***	.6508***
self—Selbst	.7281***	.7228***	.7180***
personal—persönlich	−.0668	−.0366	−.0025
ego—Ego	.8273***	.8455***	.8671***
baby—Baby	.8425***	.8396***	.8360***
special—speziell	.4023**	.3778**	.3525**
Panel B: Gemeinschaft	*Urban population*	*Urban population time lag one year*	*Urban population time lag two years*
duty—Pflicht	−.8428***	−.8558***	−.8715***
benevolence—Güte	−.8641***	−.8686***	−.8700***
obliged—versprechen	.5258	.5647***	.5802***
give—geben	−.8542***	−.8658***	−.8785***
deed—Handlung	−.3571**	−.3675**	−.3719**
act—handeln	−.6667***	−.6693***	−.6820***
authority—Amtsgewalt	−.7841***	−.7996***	−.8092***
obedience—Gehorsam	−.8033***	−.8199***	−.8362***
power—Macht	−.8272***	−.8359***	−.8481***
belong—gehören	−.8753***	−.8859***	−.8951***
pray—beten	−.5127***	−.4942***	−.4619***
join—zugesellen	−.8815***	−.8903***	−.8964***
worship—Verehrung	−.8131***	−.8073***	−.7980***
conformity—Gleichförmigkeit	−.8815***	−.8782***	−.8749***

Note: Data of urban population are from the World Bank. Column I shows the correlation coefficients for German *Gesellschaft*-indexed and *Gemeinschaft*-indexed words with urban population in the respective year. Columns II respectively III calculate correlations for the urban population in year x, using word frequency of year $x+1$ and year $x+2$, respectively.
*$p < .10$, **$p < .05$, ***$p < .01$.

Gemeinschaft-indexed words with urban population in the respective year. All significant correlations exhibit the predicted direction. In particular, words representing *Gesellschaft* values, that is, individualistic values—"auswählen" (choose), "bekommen" (get), "Kauf" (acquisition), "Emotion" (emotion), "individuell" (individual), "einzigartig" (unique), "Kind" (child), "Selbst" (self), "Ego" (ego), "Baby" (baby), "speziell" (special)—are significantly positively correlated with size of the urban population ($r = .402$ to $r = .937$, $p < .05$). In addition, words representing *Gemeinschaft* values, that is, collectivistic values—"Pflicht" (duty), "Güte" (benevolence), "geben" (give), "Handlung" (deed), "handeln" (act), "Amtsgewalt" (authority), "Gehorsam" (obedience), "Macht" (power), "gehören" (belong), "beten" (pray)—are significantly negatively correlated with size of the urban population ($r = -.368$ to $r = -.890$, $p < .05$).

We additionally control for possible time lags between writing and publishing in columns 2 and 3. In particular, we recalculate the correlations between size of the urban population in the period 1961−1999 (respectively 1961−1998) and word frequencies in the period 1962−2000 (respectively 1963−2000) and test for significance. Almost all correlations show unchanged patterns. The frequency of *Gesellschaft*-indexed words correlates significantly positive with size of the urban population ($r = .353$ to $r = .933$, $p < .05$) whereas *Gemeinschaft*-indexed words correlate significantly negative with size of the urban population ($r = -.368$ to $r = -.896$, $p < .05$).

Although correlations exhibit predicted relations and suggest that urbanisation might drive social values as measured via word frequencies, we cannot argue that there is a clear causal relationship. The relationship could be reverse, bidirectional, or there might exist latent variables that influence both, a change in social

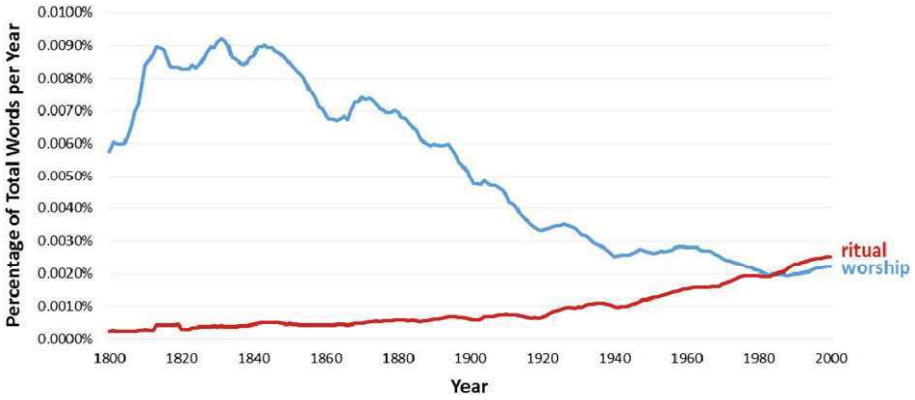

Figure 5. Frequency of the words "ritual" and "worship" in the Google corpus of American English books from the years 1800 to 2000. The graph was made with the Google Books Ngram Viewer (Michel et al., 2011), with the standard smoothing of 3.

structure as well as in word frequency. However, the large number of tested words attributes at least a predictive power to urbanisation with respect to cultural changes.

DISCUSSION

The theory of *social change and human development* predicts a global shift from *Gemeinschaft* to *Gesellschaft* based on sociodemographic changes. By using the Google Ngram Viewer, Greenfield (2013) and Zeng and Greenfield (2015) confirm this hypothesis for the United States, United Kingdom and China by respective shifts in word frequency in American English, British English and Chinese books. We replicate these findings for the German language culture and improve on the methodology by testing for the first three one-word synonyms listed in the most widely referenced synonym dictionary *Duden Synonymwörterbuch* (2014). Investigating changes in the German culture is especially interesting because of the Germans' role in World War II. In particular, we use the time period during and shortly after the war as a specific historical event with expected reversal of the global shift in cultural changes. We assume that the reported increase in the frequency of *Gemeinschaft*-indexed words might originate from Nazi-controlled publishing and the sustained propaganda of collectivistic values (Föllmer, 2010). Moreover, our results show that the overall increase in individualistic values is highly correlated with an increase in urbanisation. However, no causal relationship can be determined and urbanisation can only be viewed as a potential driver for this cultural development. Furthermore, in our study as well as in the study of Zeng and Greenfield (2015) a small fraction of words does not display the predicted directional trend. The persistence of or even increase in *Gemeinschaft* values may indicate a persisting cultural heritage that in some circumstances might contradict the process of social change (Hamamura, 2012).

Several studies have used Google Ngram to highlight cultural changes (Greenfield, 2013; Hamamura & Xu, 2015; Kesebir & Kesebir, 2012; Michel et al., 2011; Oishi, Graham, Kesebir, & Galinha, 2013; Twenge, Campbell, & Gentile, 2012a, 2012b; Xu & Hamamura, 2014; Yu et al., 2016; Zeng & Greenfield, 2015). With respect to this methodology, Twenge et al. (2012a) argue that studying language use in books allows to capture cultural changes from an individual level to group level. In fact, they suggest that books as cultural constructs mirror the viewpoint of authors, representing cultural changes of an influential portion of population, but might also highlight shifts in preferences, that is, what is in demand. We are aware that this reasoning simultaneously highlights the limitations of the methodology. In their work, Michel et al. (2011) admit that the corpus of literature is limited by the books obtained and preserved by the library partners of Google Books, raising criticism that the selection of books might not be enough to represent the majority of books and therefore culture in general. Finke and McClure (2015) address these concerns by investigating whether cultural and religious trends become visible through Ngram traces. Although the authors suggest to make use of a mix of methods in order to improve understanding, their results show that the Ngram Viewer is able to chart significant historical events in cultural and religious context. An issue with the original study by Greenfield (2013) remains in the criteria for the selection of terms and synonyms. The problem continues with the selection of translations.

Figure 5 highlights this problem. Replacing Greenfield's (2013) "worship" with the synonym "ritual," the picture changes completely and the predicted decrease turns into an increase. The same reversal is found for the German translation. By substituting "Verehrung" (worship) with the synonym "Kult" the story changes and an increase instead of a decrease is displayed. Thus, by including a predefined number of and a clear procedure

for selecting synonyms, we hope to increase the robustness of our findings.

Overall our study provides deeper insight into the cultural development of German values over time and highlights similarities in cultural changes among different countries. The impact of the World War II makes this investigation especially valuable and offers a good starting point for further research.

Manuscript received September 2016
Revised manuscript accepted March 2017
First published online May 2017

REFERENCES

Bähr, J. (2007). *Entwicklung von Urbanisierung*. Berlin, Germany: Berlin Institut für Bevölkerung und Entwicklung.. Retrieved from http://www.berlin-institut.org/online-handbuchdemografie/bevoelkerungsdynamik/auswirkungen/entwicklung-von-urbanisierung.html

Chalfant, H. P., & Heller, P. L. (1991). Rural/urban versus regional differences in religiosity. *Review of Religious Research*, *33*, 76–86.

Duden Synonymwörterbuch. (2014). Dudenredaktion. Synonymwörterbuch, D. D. *Ein Wörterbuch sinnverwandter Wörter*, 6.

Finke, R., & McClure, J. M. (2015). Reviewing millions of books: Charting cultural and religious trends with Google's Ngram Viewer. Association of Religion Data Archive's Working Paper Series. Retrieved from http://thearda.com/rrh/papers/book.asp

Föllmer, M. (2010). Was Nazism collectivistic? Redefining the individual in Berlin, 1930–1945. *The Journal of Modern History*, *82*(1), 61–100.

Greenfield, P. M. (2009). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology*, *45*(2), 401–418.

Greenfield, P. M. (2013). The changing psychology of culture from 1800 through 2000. *Psychological Science*, *24*(9), 1722–1731.

Grossmann, I., & Varnum, M. E. (2015). Social structure, infectious diseases, disasters, secularism, and cultural change in America. *Psychological Science*, *26*(3), 311–324.

Hamamura, T. (2012). Are cultures becoming individualistic? A cross-temporal comparison of individualism–collectivism in the United States and Japan. *Personality and Social Psychology Review*, *16*(1), 3–24.

Hamamura, T., & Xu, Y. (2015). Changes in Chinese culture as examined through changes in personal pronoun usage. *Journal of Cross-Cultural Psychology*, *46*(7), 930–941.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, CA: Sage.

Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American Sociological Review*, *65*, 19–51.

Kesebir, P., & Kesebir, S. (2012). The cultural salience of moral character and virtue declined in twentieth century America. *The Journal of Positive Psychology*, *7*(6), 471–480.

Kraus, M. W., Piff, P. K., Mendoza-Denton, R., Rheinschmidt, M. L., & Keltner, D. (2012). Social class, solipsism, and contextualism: How the rich are different from the poor. *Psychological Review*, *119*(3), 546–572.

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., … Pinker, S. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*(6014), 176–182.

Oishi, S., Graham, J., Kesebir, S., & Galinha, I. C. (2013). Concepts of happiness across time and cultures. *Personality and Social Psychology Bulletin*, *39*(5), 559–577.

Park, H., Twenge, J. M., & Greenfield, P. M. (2014). The Great Recession Implications for adolescent values and behavior. *Social Psychological and Personality Science*, *5*(3), 310–318.

Reips, U.-D. (2008). How Internet-mediated research changes science. In A. Barak (Ed.), *Psychological aspects of cyberspace: Theory, research, applications* (pp. 268–294). Cambridge, U.K.: Cambridge University Press.

Silbereisen, R. K., & Chen, X. (2010). *Social change and human development: Concepts and results*. London, U.K.: Sage.

Strupp, C. (2013, January 30). Only a phase: How diplomats misjudged Hitler's rise [Spiegel online]. Retrieved from http://www.spiegel.de/international/germany/marking-eighty-years-since-hitler-took-power-in-germany-a-880565.html

Tönnies, F. (1957). In C. P. Loomis (Ed. & Trans.), *Community and society: Gemeinschaft und Gesellschaft*. East Lansing, MI: Mich (Original work published 1887).

Triandis, H. C. (1993). Collectivism and individualism as cultural syndromes. *Cross-Cultural Research*, *27*(3–4), 155–180.

Triandis, H. C. (2001). Individualism-collectivism and personality. *Journal of Personality*, *69*(6), 907–924.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012a). Increases in individualistic words and phrases in American books, 1960–2008. *PLoS One*, *7*(7), e40181.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012b). Changes in pronoun use in American books and the rise of individualism, 1960–2008. *Journal of Cross-Cultural Psychology*, *44*(3), 406–415.

Uhls, Y. T., & Greenfield, P. M. (2011). The rise of fame: A historical content analysis. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *5*(1), 1. Retrieved from https://journals.muni.cz/cyberpsychology/article/view/4243.

Welch, D. (2004). Nazi propaganda and the Volksgemeinschaft: Constructing a people's community. *Journal of Contemporary History*, *39*(2), 213–238.

Xu, Y., & Hamamura, T. (2014). Folk beliefs of cultural changes in China. *Frontiers in Psychology*, *5*, 1066.

Yu, F., Peng, T., Peng, K., Tang, S., Chen, C. S., Qian, X., … Chai, F. (2016). Cultural value shifting in pronoun use. *Journal of Cross-Cultural Psychology*, *47*(2), 310–316.

Zeng, R., & Greenfield, P. M. (2015). Cultural evolution over the last 40 years in China: Using the Google Ngram Viewer to study implications of social and political change for cultural values. *International Journal of Psychology*, *50*(1), 47–55.

Second Research Article:

Guideline for Improving the Reliability of Google Ngram Studies:

Evidence from Religious Terms

Nadja Younes and Ulf-Dietrich Reips

University of Konstanz

PLOS ONE

RESEARCH ARTICLE

Guideline for improving the reliability of Google Ngram studies: Evidence from religious terms

Nadja Younes*, Ulf-Dietrich Reips

Department of Psychology, University of Konstanz, Konstanz, Germany

* nadja.younes@uni-konstanz.de

OPEN ACCESS

Citation: Younes N, Reips U-D (2019) Guideline for improving the reliability of Google Ngram studies: Evidence from religious terms. PLoS ONE 14(3): e0213554. https://doi.org/10.1371/journal.pone.0213554

Editor: Christopher M. Danforth, University of Vermont, UNITED STATES

Received: November 30, 2018

Accepted: February 24, 2019

Published: March 22, 2019

Data Availability Statement: The data used in this study is from https://books.google.com/ngrams and publicly available.

Funding: Support for this research was provided to U.-D. Reips by the Committee on Research at the University of Konstanz and the Excellence Clusters "Politics of Inequality" (EXC 2035/1) and "Collective Behaviour" (EXC 2117/1).

Competing interests: The authors have declared that no competing interests exist.

Abstract

The Google Books Ngram Viewer (Google Ngram) is a search engine that charts word frequencies from a large corpus of books and thereby allows for the examination of cultural change as it is reflected in books. While the tool's massive corpus of data (about 8 million books or 6% of all books ever published) has been used in various scientific studies, concerns about the accuracy of results have simultaneously emerged. This paper reviews the literature and serves as a guideline for improving Google Ngram studies by suggesting five methodological procedures suited to increase the reliability of results. In particular, we recommend the use of (I) different language corpora, (II) cross-checks on different corpora from the same language, (III) word inflections, (IV) synonyms, and (V) a standardization procedure that accounts for both the influx of data and unequal weights of word frequencies. Further, we outline how to combine these procedures and address the risk of potential biases arising from censorship and propaganda. As an example of the proposed procedures, we examine the cross-cultural expression of religion via religious terms for the years 1900 to 2000. Special emphasis is placed on the situation during World War II. In line with the strand of literature that emphasizes the decline of collectivistic values, our results suggest an overall decrease of religion's importance. However, religion re-gains importance during times of crisis such as World War II. By comparing the results obtained through the different methods, we illustrate that applying and particularly combining our suggested procedures increase the reliability of results and prevents authors from deriving wrong assumptions.

Introduction

Since its launch in 2010, the possibilities and limitations of using the Google Books Ngram Viewer (Google Ngram) for research purposes have been controversially discussed. Although the large number of Google Ngram studies indicates scientific recognition, several papers rightly address methodological issues (see, e.g., [1,2,3]). Yet, no set of applicable solutions is given. Our present article aims at mitigating concerns about the trustworthiness of Google

Ngram studies by presenting a guideline on different hands-on procedures that can increase reliability.

Google Ngram is a search engine that charts word frequencies from a large corpus of books that were printed between 1500 and 2008. The tool generates charts by dividing the number of a word's yearly appearances by the total number of words in the corpus in that year. Thereby a book's content is split into case-sensitive text blocks–so called *n-grams*. The word "book", for example, is a 1-gram, whereas "holy book" is a 2-gram and so on, up to a maximum of five words. Because early decades contain significantly fewer books, the overall corpus of Google Ngram becomes sufficiently large for scientific use by the year 1800 [4]. When the tool was released in 2010, the total corpus consisted of more than 5 million books, covering the languages English, French, Spanish, German, Chinese, Russian, and Hebrew. These books were drawn from over 40 different university libraries [4]. Since the corpus' latest update in 2012, users can access 22 different sub-corpora, encompassing 8 million books in total. The new version is characterized by improved optical character recognition (OCR) as well as better underlying library and publisher metadata [5]. By now, two corpora for each of the above-mentioned languages exist–an old version and an updated version. One additional corpus contains Italian books. Four further corpora split English into British and American English. Finally, there are two fiction corpora, which include predominately English fiction books and one corpus, called "English one Million", which includes a balanced text-collection of 6000 English language books, published between 1500 and 2008, and chosen from any one year.

When Google Ngram was released, Michel et al. [4] argued that one of the tool's main opportunities for scientific purposes is the hands-on quantification of cultural development that can be measured through changes in word frequencies. Related to the tool's free and easy access, Google Ngram has further contributed to an ongoing debate–the ease of replicability. By now, several dozen studies have embraced Google Ngram as an opportunity to gain insight into the development of cultural changes (see Table A in S1 Appendix for an overview of psychological Google Ngram research, published between 2010 and 2018). A vast strand of research hereby empirically tests theoretical predictions on cultural changes with a particular focus on individualism and collectivism. Twenge et al. [6,7], for example, report an increase in the frequency of individualistic words and phrases in American books. Kesebir and Kesebir [8] document a decline in the frequency of moral terms, while Twenge et al. [9] highlight an increase in the frequency of swear words. In addition, Greenfield's [10] findings suggest a relationship between increased individualism and ecological changes in the US and UK, as derived from word frequency changes in the American and British English corpora. The usage of Google Ngram has, however, not been limited to American and British English books. By tracing concepts of folk believes or personal pronouns, several studies document rising individualism also for rather collectivistic societies like China [11,12,13]. Further, such developments were also examined for German-speaking countries [14] and (Soviet) Russia [15,16]. Besides these types of studies, Google Ngram research shed light on topics such as gender differences [17,18,19], the expression of emotions [20,21,22], personality [23,24], and cognition [25,26,27].

In spite of Google Ngram's rising acceptance among researchers, critics justifiably do not grow tired in highlighting potential problems that may challenge the reliability of existing results. Main points of criticism relate to insufficient OCR, particularly with respect to semantic scanning errors (which can affect words such as *fail* and *sail* due to similarities in the letters "ſ" and "s"), and messy metadata that may lead to the display of word frequencies in wrong or unrelated time intervals. Another often heard critique refers to the overall corpora's large proportion of scientific literature and the possibility for single authors to heavily influence the data set with specific words and phrases (see, e.g., [1,2,3]).

In this guideline, we propose how to address these concerns by introducing several methodological procedures such as cross-validations via the examination of different language corpora, the use of word inflections and synonyms, as well as the use of a newly-developed standardization procedure that all aim at increasing the reliability of Google Ngram studies. Further, in a step-by-step example we apply all of these procedures by investigating cross-cultural religious trends over the last century (1900–2000) for cultures whose language is covered by Google Ngram and based on a Latin alphabet. To be able to derive assumptions from language- to country-level, we focus primarily on American and British English, German, and Italian, but consider multinational languages like Spanish and French to verify the general validity of our findings.

Assessing the development of religion via Google Ngram

While the world population almost quintupled from approximately 1.5 billion in 1900 to 6.9 billion in 2010, the number of people reporting to have no religion increased by more than 265 times (from 3 million in 1900 to 797 million in 2010) over the same period [28]. In other words, nowadays, roughly one in every ten persons on the globe reports to have no religion compared to one in every 500, a hundred years ago. This development stands in line with the vast amount of literature that emphasizes the decline of collectivistic values such as religious belonging, in the wake of an increasing industrialization (e.g., [10,29]; see also [30] for a review on collectivism and individualism). Inglehart and Baker's [31] profound study on *cultural change and the persistence of traditional values* confirms these shifts, but also points out that cultural changes can be bidirectional. Using Google Ngram, Younes and Reips [14] report on such a bidirectional change by showing an overall decline for collectivistic German words but a reversal during the time of the Nazi Regime and World War II (WWII), indicating the Germans' movement towards a more collectivistic society during that time. Accordingly, Pargament et al. [32] suggest that especially in times of crises such as wars, people cling to religious beliefs. Along these lines, several studies also mention the importance of religion as a coping mechanism during traumatic and stressful life events (e.g., [33]; see also [34] for a review).

We believe that the cross-cultural development of religion constitutes a suitable research topic to present how Google Ngram can be used for scientific purposes and what can be done to improve the investigation method of previous Google Ngram studies. Because we expect a theory-based cross-cultural decline in the importance of religion, we hypothesize that the relative frequency of religious terms in books covered by Google Ngram will decrease over time. We further hypothesize that–despite an overall decrease in the frequency of religious terms–there will be a reversal during crises that affect a large proportion of a country's population. In particular, based on the German population's complete involvement in the severe and lengthy crisis at the time of the Nazi Regime, i.e., in the years before and during WWII, we expect to observe a positive trend for religious German terms during the years 1933 and 1945. Throughout the study, we refer to this time span as WWII.

Method and results

This section describes how to conduct a Google Ngram study. We discuss several methodological improvements exemplarily in a cross-cultural setting by analyzing the development of frequencies for 20 religious terms in the American and British English, German, and Italian Google Ngram corpora. We use the latest version of each corpus and restrict our investigation period to the last century because the number of books before 1900 is relatively small and recently published books, i.e., those after the year 2000, may still need to be included in the data set [20]. For reasons of comparability to previous studies, we focus on 1-grams.

Collection of words and verification procedure

We obtain a set of 23 religious English terms from Ritter and Preston [35] who surveyed the literature of words that were used to prime religious concepts. Their list contains common nouns that represent general religious concepts. To facilitate the collection of religious German and Italian words and to establish a word identification process that follows clear criteria, we translated the religious English terms into German and Italian using PONS online dictionary (https://en.pons.com/translate). We excluded the word "faith" because the concept is already represented by the word "belief". Further, we excluded the words "holy day" and "scripture" because the first expression consists of two words and the translation of the latter leads to terminologies consisting of two words (e.g., "Heilige Schrift" in German and "Sacre Scritture" in Italian). As suggested by Zeng and Greenfield [13] and Younes and Reips [14], we recommend asking several native speakers to check independently from each other whether the respective translations are in line with the original terms. In our setting, we asked two native speakers per language, who are also proficient in English. All agreed that the presented terms were translated in an adequate manner. Table 1 shows the final list of words and their corresponding translations.

Baseline analysis

After extracting word frequencies from Google Ngram, prior studies have investigated cultural changes by examining the correlation coefficients between word frequencies and years. Based on the idea that the natural frequency of a word is relevant for assessing cultural change, it is reasonable to sum up the frequencies of single words per year (and language), and to run an aggregated correlational analysis (see, e.g., [6,7,9,18]). In this respect, the more frequent a word, the larger its proportional influence. By analyzing the summed frequencies of our

Table 1. List of religious English terms with their German and Italian translations.

Original	German	Italian
altar	Altar	altare
angel	Engel	angelo
belief	Glaube	fede
clergy	Geistlichkeit	clero
creed	Überzeugung	credo
doctrine	Doktrin	dottrina
God	Gott	Dio
heaven	Himmel	paradiso
miracle	Wunder	miracolo
pilgrimage	Pilgerfahrt	pellegrinaggio
prayer	Gebet	preghiera
prophet	Prophet	profeta
religion	Religion	religione
revelation	Offenbarung	rivelazione
ritual	Ritual	rituale
saint	Heiliger	santo
sermon	Predigt	predica
shrine	Schrein	santuario
soul	Seele	anima
spirit	Geist	spirito

https://doi.org/10.1371/journal.pone.0213554.t001

religious words (by year and language), we observe a negative correlation between the years and the religious terms' frequencies for American English ($r = -0.89$, $p<0.001$), British English ($r = -0.96$, $p<0.001$), German ($r = -0.47$, $p<0.001$), and Italian ($r = -0.64$, $p<0.001$). In contrast, investigating the development of religious terms in times of crisis such as WWII, German terms show a significant positive trend ($r = 0.79$, $p<0.01$). The findings suggest that religion became less important over time but showed more prevalence during a time of existential crisis (see Figures A-H in S1 Appendix for visual inspection).

Procedure I: Multiple languages

As highlighted by our current analysis, one big advantage of Google Ngram is the possibility to compare cultural changes in a cross-cultural setting. Although transferring assumptions from language- to country-level, particularly for multinational languages, can be a difficult procedure, several studies suggest comparing differences and similarities of several languages to derive assumptions on the general validity of a certain theory or concept [36,37,38]. Following this reasoning, we further examined religious words' frequencies during the years 1900 to 2000 using the Spanish and French corpora. Hence, we first translated the original religious English terms into Spanish and French using PONS. Again, two French- and two Spanish-speaking natives confirmed our translations. Table B in S1 Appendix displays the original English terms with their respective translations to Spanish and French. In line with our initial findings, religious words in Spanish ($r = -0.47$, $p<0.001$) and French ($r = -0.83$, $p<0.001$) show a significant negative trend over time (see Figures I-L in S1 Appendix for visual inspection).

Procedure II: English fiction corpus

So far, the majority of Google Ngram studies focuses on the English language, i.e., by investigating the American and British English corpora. Besides the traditional English corpora, Google Ngram also offers access to an English fiction corpus. The corpus is limited to fiction books in the Google database and does not distinguish between American and British English. Although it might seem counterintuitive, the fiction corpus may provide an additional opportunity to test the reliability of derived assumptions on cultural changes from the American and British English corpora. Following the argument that "a fiction writer may aim to capture realistic modern dialogue" ([7], p. 407), cultural changes should be also reflected in the fiction corpus—at least to some extent. Because the traditional corpora contain a large amount of scientific text, particularly throughout the 1900s [2], Virues-Ortega and Pear [27] argue that fiction books are less influenced by scientific trends and may therefore provide more general results. Accordingly, several studies used the fiction corpus to test the robustness of their findings. For instance, Twenge et al. [7] confirm their findings for a decrease in first person plural pronouns and an increase in first person singular pronouns by using the fiction corpus. Acerbi et al. [20] document an overall decrease in the use of mood words, which is also reflected in the fiction corpus. By analyzing manifestations of individualism, Twenge et al. [9] confirm the significant increase in the use of swear words also for the fiction corpus. Dependent on the particular analysis, contemplating fictional literature may provide additional insights. Brysbaert et al. [39], for example, report that word frequencies extracted from the fiction corpus predict word processing better than word frequencies obtained from the regular corpora. Because the fiction corpus is only available in English, we investigated the trend pattern for the original religious English terms. By using the fiction corpus, the significant negative overall trend ($r = -0.96$, $p<0.001$) can be confirmed (see Figures M and N in S1 Appendix for visual inspection).

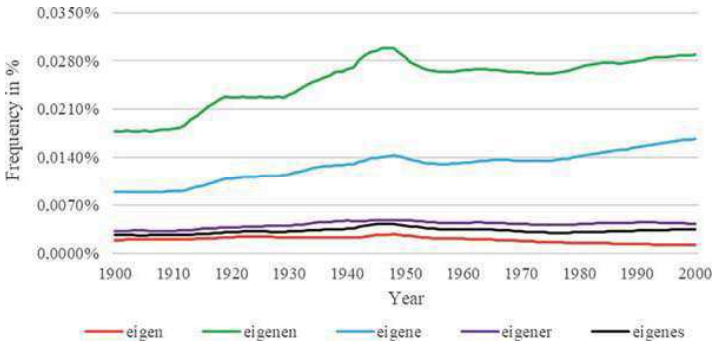

Fig 1. Higher frequency inflections for the German word "eigen".

https://doi.org/10.1371/journal.pone.0213554.g001

Procedure III: Word inflections

Using Google Ngram to analyze long-term relationships between ecological and cultural changes in German-speaking countries, Younes and Reips [14] report that in spite of the theory-based prediction of an increase, the relative frequency of the individualistic word "eigen" (personal/individual) dropped over time. As highlighted by Fig 1, all of the word's higher frequency inflections (i.e., "eigenen", "eigene", "eigener", or "eigenes") display, however, the expected rise over the course of time.

Because "the most robust historical trends are associated with frequent n-grams" ([4], Supplementary Online Material, p.12), one potential reason for the counter-predictive trend displayed for the word "eigen" may be the word's comparatively low overall frequency. Considering the use of the tag "_INF" may help to address that issue in a systematic way. Applying the tag, Google Ngram provides graphs for yet available inflections of a certain word. To obtain inflections, Google Ngram uses Wiktionary entries (www.wiktionary.org) and supplement them with automatically generated inflection tables. "Because Wiktionary is an evolving resource, results for a particular [Google Ngram] query may change over time" ([40], p. 4). In this respect, the tool allows for a direct comparison of inflection frequencies. Thus, the feature is not only helpful for identifying words of highest frequency within a similar group, but also highlights inconsistencies of lower frequency words such as in the case of "eigen".

In this analysis, we checked the consistency between and the frequencies of the original and translated terms and their respective inflections. We exemplify the procedure on the word "saint". Searching for the word's inflections (by using "saint_INF") in the American and British English corpora yields three modifications ("saints", "sainted", and "sainting"). As indicated by Fig 2(A), for the American English corpus, the term "saints" shows a similar development as

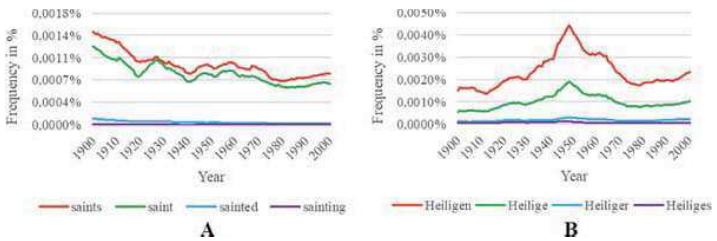

Fig 2. Frequencies of inflections for the word "saint". Frequencies of given inflections for the word "saint" using the American English corpus (A) and the three most frequent inflections for the German translation "Heiliger" (B).

https://doi.org/10.1371/journal.pone.0213554.g002

Table 2. Overview of original words and their higher frequency inflections.

American English			British English		
Original	High	Ratio	Original	High	Ratio
angel	angels	1.08	angel	angels	1.12
saint	saints	1.18			
German			Italian		
Original	High	Ratio	Original	High	Ratio
Engel	Engels	2.28	angelo	angeli	1.52
Glaube	Glauben	1.91			
Prophet	Propheten	2.54			
Heiliger	Heiligen	13.10			

The columns "Original" present original words, whereas the columns "High" present the original words' higher frequency inflections. The columns "Ratio" display the average yearly ratios between the frequencies of higher frequency inflections and the frequencies of their original counterparts.

https://doi.org/10.1371/journal.pone.0213554.t002

"saint" but exhibits a higher relative frequency. The picture for British English is similar. However, the terms "saint" and "saints" have an almost identical frequency. In contrast, searching for inflections of the respective German translation "Heiliger" yields 13 inflections. Fig 2(B) displays the three most frequent inflections ("Heiligen", "Heilige", and "Heiliges") and thereby illustrates that the term "Heiliger" shows a similar pattern but an approximately three times lower relative frequency than "Heilige" and a 13 times lower relative frequency compared to the most frequent word "Heiligen".

Finally, no higher frequency inflections are given for the Italian word "santo". While all of our originally selected terms display similar patterns as their semantically similar inflections, the original English terms as well as our translations were not always those with the highest frequency compared to their inflections. Table 2 provides an overview of those terms for which we found higher frequency inflections as well as the ratios between the frequencies of higher frequency inflections and the frequencies of their original counterparts.

Re-conducting our initial analysis by using inflections that display a higher frequency than the originally selected religious terms confirms previous findings. In particular, we find a significant negative correlation between years and frequencies of religious terms for American English ($r = -0.89$, $p<0.001$), British English ($r = -0.96$, $p<0.001$), German ($r = -0.49$, $p<0.001$), and Italian ($r = -0.65$, $p<0.001$). German terms continue to exhibit a positive trend during WWII ($r = 0.79$, $p<0.01$).

Despite the benefits of choosing words with more robust historical trends, one disadvantage of the tag "_INF" is the mechanical adjustment of the words' endings. The inflected terms' original meanings can sometimes be significantly distorted. One example relates to the Italian word "predica" (sermon), which is far from being represented by the higher frequency modifications "pred**etto**" (aforementioned) and "predica**to**" (predicate). Thus, a manual reconsideration of the words' meanings, particularly for non-native speakers, is inevitable.

Procedure IV: Synonyms

To verify that investigated words reflect true underlying concepts rather than idiosyncrasies, re-checking initial findings with several synonyms is a strong robustness check. Pettit [3], for example, argues that especially 1-grams may be risky to analyze due to potential changes in the words' meanings over time. One example that illustrates Pettit's [3] concerns is the word "nice". Nowadays, the word is used to express that something is pleasant. However, in the fourteenth century the term was primarily related to negative quality such as "foolish" or "silly"

(see, e.g., https://blog.oxforddictionaries.com/2012/10/01/change-in-word-meanings/). Using synonyms reduces the impact of a particular word and therefore decreases the probability of incorrect assumptions due to, e.g., varying meanings, idiosyncrasies, messy metadata, and OCR errors. Younes and Reips [14] introduced the use of synonyms for the German corpus by recommending the collection of the first three one-word synonyms out of the semantically most adequate grouping (i.e., only synonyms of "date" as a type of fruit, not as "going out"), listed in the *Duden Synonymwörterbuch*, the standard reference for the German language. In the current study, we followed this procedure for selecting German synonyms. In addition, we collected the first three one-word English synonyms for the most adequate grouping using the standard reference for English that is *Roget's Thesaurus*. For selecting Italian synonyms, we used and recommend *Zanichelli's Sinonimi e Contrari*.

Table C in S1 Appendix presents an overview of collected synonyms. We obtained at least one synonym in 95% of all cases. A dictionary-based approach may however not always be the optimal choice to obtain synonyms. Although it avoids an arbitrary selection of words, not for every word there are synonyms and some of the obtained synonyms are often used in a non-related context. Hence, dependent on the investigated language and particularly the topic, it might be helpful to consult several native speakers to suggest (additional) synonyms or rate the obtained synonyms according to their suitability. In accordance with this notion, we let our synonyms rate by two native speakers per language on a scale of 0 (no synonym) to 10 (perfect synonym). The average rating for all synonyms over all languages was relatively high, m = 7.5 (6.8 for English, 7.1 for German, and 8.5 for Italian).

Google Ngram provides frequencies for all synonyms except for the German word "Brandopferstätte" (altar). Analyzing correlations for our set of synonyms, we can confirm the results obtained for the sample of original words. In particular, we find significant negative correlations between years and American English ($r = -0.96$, $p<0.001$), British English ($r = -0.93$, $p<0.001$), German ($r = -0.92$, $p<0.001$), and Italian ($r = -0.81$, $p<0.001$) word frequencies. With respect to WWII, German terms continue to exhibit a positive trend, which is however not significant ($r = 0.12$, $p>0.1$).

Procedure V: Standardization of word frequencies

In spite of Michel et al.'s [4] argument that "the most robust historical trends are associated with frequent n-grams" ([4], Supplementary Online Material, p.12), high frequency words significantly drive average results if their frequency is relatively larger than the frequencies of other words that are summed up. In order to account for the influence of individual words, several studies transform word frequencies into z-scores prior to summing them (see, e.g., [6,9,15]). Although this procedure mitigates the disproportionately large influence of single words by giving each word an equal weight, it does not account for the varying size of the corpora, i.e., the influx of data over time. To address this concern, Acerbi et al. [20] and Bentley et al. [41] suggest normalizing word frequencies by expressing a word's frequency relative to the frequency of a very common word.

In this section, we demonstrate that the combination of z-scoring and normalizing raw word frequencies by common words is a beneficial procedure that can mitigate biased estimations. By themselves, z-scoring and normalizing by a common word either account for unequal weights or data-related trends, respectively. To address both concerns simultaneously, we propose a new procedure that accounts for data-related trends and gives each word an equal weight. In particular, we recommend subtracting the summed z-scored frequencies of various very common words obtained from Lin et al. [5], from the summed z-scored frequencies of the original terms. This procedure is beneficial for several reasons. First, we can mitigate

any disproportionately large influence of single words by giving each of the original terms an equal weight. Second, we do not only focus on one common word, whose pattern might deviate from the pattern of other very common words, but on various common words. Third, by also z-scoring the set of common words, we further ensure to treat all common words equally.

In the following, we highlight the differences in results between several standardization procedures. In particular, we compare correlation coefficients between time and religious word frequencies considering (I) summed raw frequencies, (II) summed z-scores of raw frequencies, (III) summed raw frequencies that were previously normalized by the raw frequency of a very common word, and (IV) summed raw frequencies that were previously normalized by the summed raw frequencies of various very common words. In a further step, we present the results for (V) summed z-scores of raw frequencies that were previously normalized by the raw frequency of a very common word and (VI) summed z-scores of raw frequencies which were previously normalized by the summed raw frequencies of various very common words. Finally, we present the correlation coefficients for summed z-scores of raw frequencies that were normalized by subtracting the summed z-scores of raw frequencies of various very common words (VII).

We calculate z-scores as

$$z_t = \frac{w_t - \mu}{\sigma}. \quad (1)$$

w_t represents the frequency of a religious word w in year t, μ the mean of w_t, and σ the standard deviation of w_t.

To account for the influx of data, we first identified the most common word in each corpus by manually comparing the frequencies for each of the most common words as indicated by Lin et al. [5]. Lin et al.'s [5] list contains the 16 most common words in each language, equally split into the word classes *adjectives*, *prepositions*, *adverbs*, *conjunctions*, *articles*, *nouns*, *pronouns*, and *verbs*. The most common word in the American and British English corpora is "the", in the German corpus it is "der", and in the Italian corpus it is "di".

Fig 3 shows the extent of variation across time for the most frequent word in each corpus. Theoretically, the relative frequency of these neutral words should remain rather constant over time. However, even these words are subject to a trend. For example, the word "the" decreased by approximately 1% (which is ca. 20% of the overall frequency) over the last century, whereas the word "di" exhibited a 0.3% increase (which is ca. 10% of the overall frequency) over the same time. Hence, Fig 3 reinforces the concern that the influx of data should be taken into account by applying an appropriate standardization procedure to avoid biased estimations.

To obtain a reference set of various common words, we consider Lin et al.'s [5] total list of common words and compare it to Hughes et al.'s [42] list of neutral words, i.e., words with little or no specific meaning. This procedure ensures that results are not driven by any specific underlying context of the common words. In particular, we select all words that appear on both lists but refrain from including pronouns, based on their categorization as content-full words by previous literature (see, e.g., [7,11,18,37,38]). Because Hughes et al.'s [42] list comprises only English terms, we translated Lin et al.'s [5] most common German and Italian words to English using PONS. We then searched for the translated words on Hughes et al.'s [42] word list. Applying this procedure, we obtain the words "other", "such", "of", "in", "not", "when", "and", "or", "the", "a", "is", and "was" for American and British English. For German, we identified "anderen", "ersten", "in", "von", "auch", "so", "und", "daß", "der", "die", "ist", and "werden". Finally, our list of common Italian words comprises "stesso", "di", "in", "non", "piú", "che", "ed", "la", "il", and "é" (see Figures O-R in S1 Appendix for visual inspection).

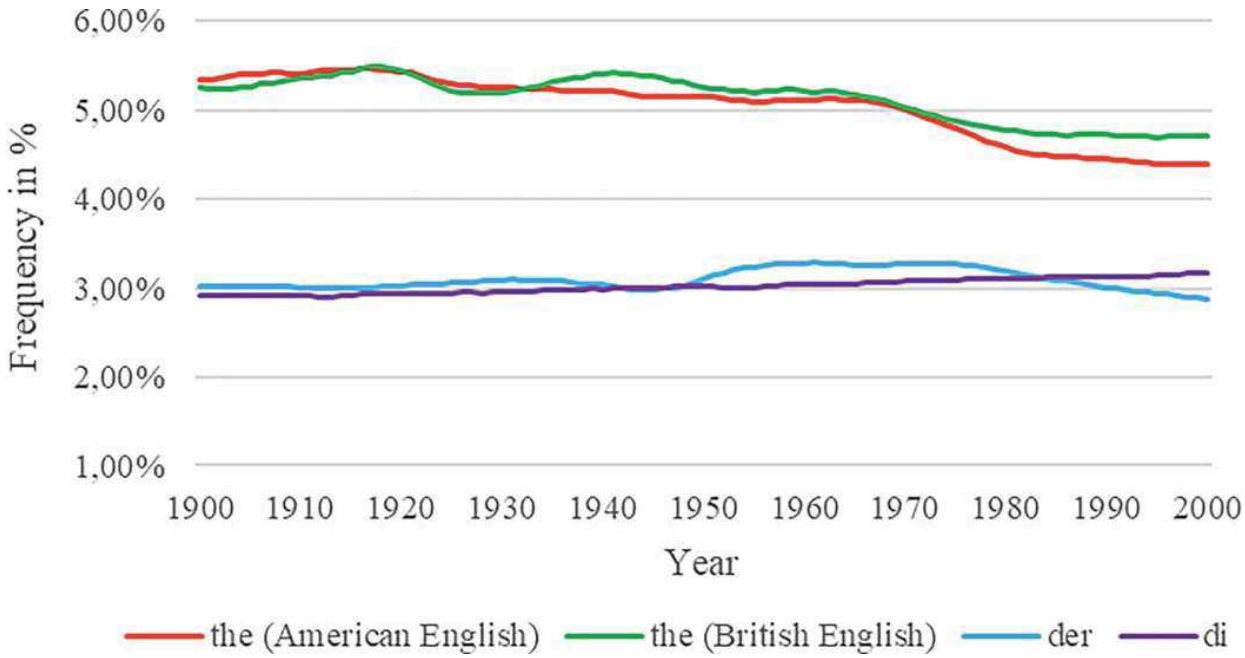

Fig 3. Frequencies for most frequent words. Frequencies for the most frequent words in the American English, British English, German, and Italian Google Ngram corpora.

https://doi.org/10.1371/journal.pone.0213554.g003

Table 3 presents the correlation coefficients for the different standardization procedures and highlights how sensitive results can be, depending on an author's choice of methodology.

Whereas results for the American and British English corpora remain relatively stable, Model VII shows that the significance for the German corpus vanishes by accounting simultaneously for unequal weights and data-related time trends. For the Italian corpus results are very inconsistent. As indicated by Model I, considering raw frequencies, we find a significant negative trend over time. If we z-score raw frequencies (Model II), we obtain a positive trend that, however, turns negative if we normalize frequencies beforehand by a very common word (Model III). Finally, by addressing both, unequal weights and the influx of data (Model VII)

Table 3. Correlation coefficients for different standardization procedures.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Original Words	**Σ Raw**	**Σ z-scores**	$\sum \frac{\text{Raw}}{\text{Most common}}$	$\sum \frac{\text{Raw}}{\Sigma \text{ Most common}}$	**Σ z–scores of** $\left[\frac{\text{Raw}}{\text{Most common}}\right]$	**Σ z–scores of** $\left[\frac{\text{Raw}}{\Sigma \text{ Most common}}\right]$	**Σ z–scores – Σ z–scores of most common**
American English	r = -0.89, p<0.001	r = -0.88, p<0.001	r = -0.71, p<0.001	r = -0.77, p<0.001	r = -0.70, p<0.001	r = -0.76, p<0.001	r = -0.49, p<0.001
British English	r = -0.96, p<0.001	r = -0.94, p<0.001	r = -0.95, p<0.001	r = -0.96, p<0.001	r = -0.91, p<0.001	r = -0.92, p<0.001	r = -0.81, p<0.001
German WWII	r = -0.47, p<0.001 r = 0.79, p<0.01	r = -0.29, p<0.01 r = 0.87, p<0.001	r = -0.51, p<0.001 r = 0.81, p<0.001	r = -0.46, p<0.001 r = 0.79, p<0.01	r = -0.32, p<0.001 r = 0.89, p<0.001	r = -0.26, p<0.01 r = 0.87, p<0.001	r = -0.13, p>0.1 r = 0.85, p<0.001
Italian	r = -0.64, p<0.001	r = 0.24, p<0.05	r = -0.76, p<0.05	r = -0.58, p<0.001	r = -0.01, p>0.1	r = 0.35, p<0.001	r = 0.51, p<0.001

https://doi.org/10.1371/journal.pone.0213554.t003

Italian terms exhibit a positive trend over time. As the main religion in the US, UK, Germany, and Italy has been Christianity, we added a further set of 14 Christian-specific nouns from Ritter and Preston [35] in an additional robustness check (i.e., the words baptism, Bible, Christmas, church, commandments, communion, cross, gospel, Jesus, messiah, preacher, Sabbath, salvation, and worship). Replicating the results in Table 3, Model VII, for a total set of 34 religious terms per language, we obtain similar results as for the original set of religious words for American English ($r = -0.51$, $p<0.001$), British English ($r = -0.84$, $p<0.001$), German ($r = -0.20$, $p<0.05$), WWII ($r = 0.80$, $p<0.01$), and Italian ($r = 0.56$, $p<0.001$).

Composite analysis

In this section, we combine our suggested procedures by using higher frequency words, synonyms, and the standardization procedure that accounts for unequal weights and the influx of data. In particular, we re-calculated previous analyses including the sets of additional synonyms and using the tag _INF to identify the most frequent words for our sets of original terms and synonyms (see Table D in S1 Appendix for an overview of corresponding higher frequency synonyms). Table 4 presents correlation coefficients for different standardization procedures using higher frequency words for original terms and additional synonyms.

The results reinforce the significant negative trend for German terms over time suggesting that prior results were affected by the set of original words. Indeed, re-running Model VII by only considering synonyms also displays a significant negative trend ($r = -0.20$, $p<0.05$). Thus, particularly the consideration of additional synonyms provides a more robust picture. Further, for Italian terms, a positive trend (as suggested by using the standardization procedure that accounts for an influx of data and unequal weights) cannot be confirmed. Instead, conducting the composite analysis, we do not find any significant trend (Model VII). In fact, we even obtain a significant negative trend over time by re-running Model VII but only considering synonyms ($r = -0.36$, $p<0.001$). As a final robustness check, we re-run Model VII, by dropping synonyms that received an average rating of less than the mean rating of 7.5. Results do not change for American English ($r = -0.64$, $p<0.001$), British English ($r = -0.89$, $p<0.001$), German ($r = -0.23$, $p<0.05$), WWII ($r = 0.89$, $p<0.001$), and Italian ($r = -0.09$, $p>0.1$).

Overall, this analysis strongly emphasizes that only the combination of different methodological procedures mitigates biased estimations and therefore prevents researchers from deriving wrong and undifferentiated assumptions.

Table 4. Correlation coefficients for different standardization procedures using higher frequency words and synonyms.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Original Terms + Synonyms	**Σ Raw**	**Σ z-scores**	$\sum \frac{\text{Raw}}{\text{Most common}}$	$\sum \frac{\text{Raw}}{\Sigma \text{ Most common}}$	**Σ z–scores of** $\left[\frac{\text{Raw}}{\text{Most common}}\right]$	**Σ z–scores of** $\left[\frac{\text{Raw}}{\Sigma \text{ Most common}}\right]$	**Σ z–scores – Σ z–scores of most common**
American English	$r = -0.92$, $p<0.001$	$r = -0.83$, $p<0.001$	$r = -0.77$, $p<0.001$	$r = -0.83$, $p<0.001$	$r = -0.60$, $p<0.001$	$r = -0.68$, $p<0.001$	$r = -0.72$, $p<0.001$
British English	$r = -0.95$, $p<0.001$	$r = -0.91$, $p<0.001$	$r = -0.94$, $p<0.001$	$r = -0.95$, $p<0.001$	$r = -0.87$, $p<0.001$	$r = -0.89$, $p<0.001$	$r = -0.89$, $p<0.001$
German WWII	$r = -0.73$, $p<0.001$ $r = 0.66$, $p<0.05$	$r = -0.33$, $p<0.001$ $r = 0.86$, $p<0.001$	$r = -0.76$, $p<0.001$ $r = 0.72$, $p<0.01$	$r = -0.74$, $p<0.001$ $r = 0.65$, $p<0.05$	$r = -0.36$, $p<0.001$ $r = 0.88$, $p<0.001$	$r = -0.29$, $p<0.01$ $r = 0.86$, $p<0.001$	$r = -0.26$, $p<0.01$ $r = 0.85$, $p<0.001$
Italian	$r = -0.76$, $p<0.001$	$r = -0.28$, $p<0.01$	$r = -0.84$, $p<0.001$	$r = -0.72$, $p<0.001$	$r = -0.49$, $p<0.001$	$r = -0.19$ $p<0.1$	$r = -0.14$, $p>0.1$

https://doi.org/10.1371/journal.pone.0213554.t004

Censorship and propaganda

Finally, we discuss the importance of taking any potential censorship and propaganda into account. Michel et al. [4] address these issues by stressing, for example, the decreasing popularity of the Jewish artist Marc Chagall in the German corpus during the Nazi regime, the absence of the 1976 and 1989 Tiananmen Square incidents in the Chinese corpus, and the reflection of the Hollywood Ten's blacklisting in the American corpus. In this section, we aim to increase awareness that censorship and propaganda may not only affect certain people or historical events but whole subject-areas. In particular, we graphically illustrate the impact of the Soviet regime's (1922–1991) enforcement of secularization and religious persecution on the Russian corpus by plotting the overall frequencies for the original religious terms translated to Russian. For this reason, we translated our original English terms to Russian using PONS. Two native speakers re-checked and then confirmed the translations.

Fig 4 shows the results. The graphs include all of the original translated terms except for "Бог" (God), "душа" (soul), and "дух" (spirit). We did not include these words due to scaling differences. However, they display a similar curvature.

The graphs highlight a drop around the 1920s, followed by rather flat movement until the 1990s, where the frequencies start to rise again. In this respect, they likely mirror the Soviet regime's objective to eliminate religion. Because there is evidence that Russians acted out religion secretly rather than becoming atheists (see, e.g., [31,43]), there is a gap between what was prevalent in people's minds and what was artificially created. Therefore, the assumptions which we would derive from visual inspection and also by calculating a correlation coefficient, would mirror the development of the importance of religion for the Soviet Union's population only undifferentiated and incomplete. Hence, a country's historical context must always be taken into account. In case of any potential concerns about distorted time periods, we recommend following Hamamura and Xu [11] and investigate sub-periods (e.g., decades) to iteratively discover cultural trends.

Discussion

Google Ngram allows for hands-on quantification of cultural change using millions of books. In spite of the tool's unique opportunities for research purposes, several studies justifiably emphasize the existence of potential problems. With this paper, we contribute to the ongoing debate of weighting the tool's advantages and disadvantages against each other. To the best of

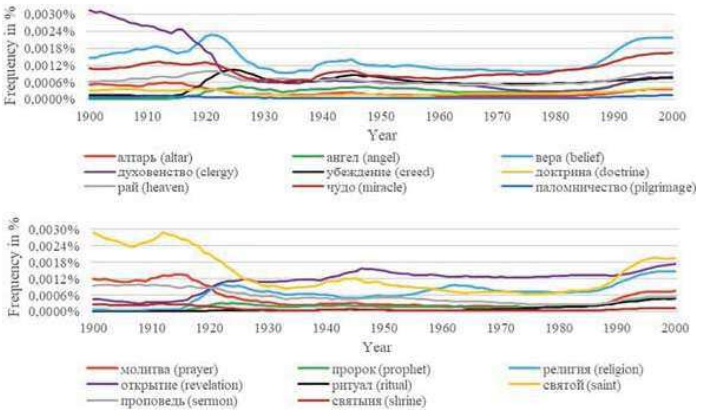

Fig 4. Frequencies of religious terms translated to Russian.

https://doi.org/10.1371/journal.pone.0213554.g004

our knowledge, this is the first summary of the tool's limitations that comes with a set of methodological procedures that are suited to improve the reliability of results. In this guideline we present five methodological procedures, a roadmap on how to use them in combination, and a way to check for influences of censorship and propaganda. In particular, to test the universality of a certain theory or hypothesis and to examine derived results cross-culturally, we propose (I) the use of different language corpora. Further, (II) using the fiction corpus, which is not heavily impacted by scientific literature, allows researchers to reinforce results obtained for the American and British English corpora. We additionally recommend to (III) compare frequencies of words and their inflections using the feature "_INF". We further advise researchers to (IV) make use of synonyms to diminish the probability of wrong assumptions caused by the prevalence of, e.g., OCR errors, flaws related to messy metadata, idiosyncrasies, or the disproportionately large influence of single authors' contributions. Finally, we also strongly recommend to (V) apply a standardization procedure that accounts for both, the influx of data and unequal weights.

In addition to the individual benefits of each procedure, we strongly advise researchers to use them in combination. As shown by a composite analysis, the most robust results are obtained by combining several procedures. So far, most prior studies did not challenge their results by taking different procedures into account. Based on the literature review displayed in Table A in S1 Appendix, only approximately 3% of previous studies considered both, the use of synonyms and any form of standardization procedure. None of the studies took the particular use of higher frequency inflections into account. As with any new and evolving methodology, many such procedures are only discovered as beneficial over time. Our guideline demonstrates how sensitive results can be to specific choices an author makes and how to resolve such dependencies. For example, by focusing on raw frequencies or by applying a standardization procedure that only accounts for the influx of data over time, we obtain a significant negative trend for Italian religious terms (see Table 3, Model I and III). However, accounting for unequal weights (see Table 3, Model II) leads to the impression that in Italy, religion became more important over time. Furthermore, taking unequal weights and time trends simultaneously into account, the significant positive trend remains only for our set of original terms (see Table 3, Model VII) but turns negative for our set of synonyms. Finally, combining words of comparatively high frequency and the most advanced standardization procedure with an additional set of synonyms, no significant trend is observable any longer (see Table 4, Model VII). Consequently, this example illustrates that only analyzing Google Ngram data in pre-determined appropriate ways can prevent authors from deriving "first-glance" assumptions that may turn out as not valid when a more robust analysis procedure is applied.

Finally, by the example of forced secularization during the Soviet regime, we draw attention to the danger of biased estimations that can arise from censorship and propaganda. However, we would like to point out that what we believe is censorship might be also attributable to the lack of metadata. In particular, Koplenig [44] shows that with the lack of proper metadata, it cannot be ruled out that trends arise due to changes in the composition of the underlying data. Thus, some of the trends we observe for the Russian corpus may not necessarily result from censorship as part of forced secularization, but some change in the data.

By applying all of the above-mentioned procedures to religious 1-grams, this study further contributes to the body of research that investigates the development of collectivism and individualism (see, e.g., [30] for a review) as well as to the literature that discusses religion as a coping strategy during crises (see, e.g., [34] for a review). In particular, we study the cross-cultural development of religious trends for the years 1900 to 2000, with a particular focus on the development in times of crisis such as WWII. Except for the Italian corpus, our analyses reveal a

relative overall decrease for religious terms. With respect to WWII, German terms display an upward trend, indicating the importance of religion during a severe crisis along with the Germans' movement towards a more collectivistic society during the war [14]. The trend towards an increased expression of religion during a severe time of crisis is, however, not only observable in the German corpus. Additional analyses further reveal a significant positive and robust trend for Italian religious terms. The positive trend for Italian terms likely relates to the alliance between Nazi-Germany and Italy during WWII. In contrast, for the American and British English corpora we do not find such a reversal but a constant and significant negative time trend over the whole observation period. Although both countries also participated in the war, we believe that in contrast to Italy, results strongly relate to the larger distance of the respective countries' population to the main war zone (see Table E in S1 Appendix for an overview of detailed results).

Our study is not without limitations. First, in spite of our suggestions on how to mitigate the influence of the potentially large number of scientific texts in the Google Ngram's corpora, we did not discuss concerns related to the corpora's general level of diversity. Although the developers have addressed this issue by providing a corpus with a better-balanced text collection, i.e., "English One Million", we do not recommend the use of this corpus because it has never been updated and is therefore considered to be per se more error prone. Nevertheless, re-calculating previous analyses by using the corpus "English One Million", we obtain similar results. Second, we used the new updated corpora to exploit the advantages of improved OCR and better underlying library and publisher metadata. However, as suggested by Twenge et al. [9], the new corpora might entail errors that the old corpora did not capture. We therefore re-examined all analyses using the old corpora. Overall, previous findings are confirmed. However, there was no Italian corpus before the update. Finally, we agree that the key assumption of Google Ngram studies, i.e., that print culture represents culture as a whole, is certainly an undifferentiated view that may not always hold.

Overall, Google Ngram has allowed scholars to shed further light on various topics such as gender differences [17,18,19], emotions [20,21,22,45], personality [23,24], cognition [25,26,27], hypnosis and psychotherapy [46], moral values [47], education [48], nature [49], astrology and phrenology [50], and the development of individualism and collectivism (e.g., [6,7,10,51,52]). In this respect, despite limitations, we believe that Google Ngram is a beneficial tool for research purposes and that the procedures presented in this guideline can reinforce the reliability of derived results.

Supporting information

S1 Appendix.
(PDF)

Acknowledgments

We thank participants at the Annual Meeting of the Society for Computers in Psychology (SCiP) in 2018 for valuable comments. We further thank Deborah Canales, Tim Kuhlmann, Torsten Twardawski, Vaka Vésteinsdóttir, and Christopher Wolfe for constructive suggestions that improved the quality of this article.

Author Contributions

Conceptualization: Nadja Younes, Ulf-Dietrich Reips.

Data curation: Nadja Younes.

Investigation: Nadja Younes.

Methodology: Nadja Younes, Ulf-Dietrich Reips.

Writing – original draft: Nadja Younes, Ulf-Dietrich Reips.

Writing – review & editing: Nadja Younes, Ulf-Dietrich Reips.

References

1. Gooding P. (2012). Mass digitization and the garbage dump: The conflicting needs of quantitative and qualitative methods. *Literary and Linguistic Computing*, 28(3), 425–431.
2. Pechenick E. A., Danforth C. M., & Dodds P. S. (2015). Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PloS One*, 10(10), e0137041. https://doi.org/10.1371/journal.pone.0137041 PMID: 26445406
3. Pettit M. (2016). Historical time in the age of big data: Cultural psychology, historical change, and the Google Books Ngram Viewer. *History of Psychology*, 19(2), 141. https://doi.org/10.1037/hop0000023 PMID: 27100927
4. Michel J. B., Shen Y. K., Aiden A. P., Veres A., Gray M. K., Pickett J. P., et al. (2011). Quantitative analysis of culture using millions of digitized books [With Supporting Online Material]. *Science*, 331(6014), 176–182. https://doi.org/10.1126/science.1199644 PMID: 21163965
5. Lin Y., Michel J. B., Aiden E. L., Orwant J., Brockman W., & Petrov S. (2012, July). Syntactic annotations for the Google Books Ngram corpus. In *Proceedings of the ACL 2012 System Demonstrations* (pp. 169–174). Association for Computational Linguistics.
6. Twenge J. M., Campbell W. K., & Gentile B. (2012a). Increases in individualistic words and phrases in American books, 1960–2008. *PloS One*, 7(7), e40181.
7. Twenge J. M., Campbell W. K., & Gentile B. (2013). Changes in pronoun use in American books and the rise of individualism, 1960–2008. *Journal of Cross-Cultural Psychology*, 44(3), 406–415.
8. Kesebir P., & Kesebir S. (2012). The cultural salience of moral character and virtue declined in twentieth century America. *The Journal of Positive Psychology*, 7(6), 471–480.
9. Twenge J. M., VanLandingham H., & Campbell W. K. (2017). The seven words you can never say on television: Increases in the use of swear words in American books, 1950–2008. *SAGE Open*, 7(3), 2158244017723689.
10. Greenfield P. M. (2013). The changing psychology of culture from 1800 through 2000. *Psychological Science*, 24(9), 1722–1731. https://doi.org/10.1177/0956797613479387 PMID: 23925305
11. Hamamura T., & Xu Y. (2015). Changes in Chinese culture as examined through changes in personal pronoun usage. *Journal of Cross-Cultural Psychology*, 46(7), 930–941.
12. Xu Y., & Hamamura T. (2014). Folk beliefs of cultural changes in China. *Frontiers in Psychology*, 5, 1066. https://doi.org/10.3389/fpsyg.2014.01066 PMID: 25309491
13. Zeng R., & Greenfield P. M. (2015). Cultural evolution over the last 40 years in China: Using the Google Ngram Viewer to study implications of social and political change for cultural values. *International Journal of Psychology*, 50(1), 47–55. https://doi.org/10.1002/ijop.12125 PMID: 25611928
14. Younes N., & Reips U. -D. (2018). The changing psychology of culture in German-speaking countries: A Google Ngram study. *International Journal of Psychology*, 53, 53–62. https://doi.org/10.1002/ijop.12428 PMID: 28474338
15. Skrebyte A., Garnett P., & Kendal J. R. (2016). Temporal relationships between individualism–collectivism and the economy in Soviet Russia: A word frequency analysis using the Google Ngram corpus. *Journal of Cross-Cultural Psychology*, 47(9), 1217–1235.
16. Velichkovsky B. B., Solovyev V. D., Bochkarev V. V., & Ishkineeva F. F. (2017). Transition to market economy promotes individualistic values: Analysing changes in frequencies of Russian words from 1980 to 2008. *International Journal of Psychology*.
17. Del Giudice M. (2012). The twentieth century reversal of pink-blue gender coding: A scientific urban legend? *Archives of Sexual Behavior*, 41(6), 1321–1323. https://doi.org/10.1007/s10508-012-0002-z PMID: 22821170
18. Twenge J. M., Campbell W. K., & Gentile B. (2012b). Male and female pronoun use in US books reflects women's status, 1900–2008. *Sex Roles*, 67(9–10), 488–493.

19. Ye S., Cai S., Chen C., Wan Q., & Qian X. (2018). How have males and females been described over the past two centuries? An analysis of Big-Five personality-related adjectives in the Google English Books. *Journal of Research in Personality*, 76, 6–16.
20. Acerbi A., Lampos V., Garnett P., & Bentley R. A. (2013). The expression of emotions in 20th century books. *PloS One*, 8(3), e59030. https://doi.org/10.1371/journal.pone.0059030 PMID: 23527080
21. Mohammad S. M. (2012). From once upon a time to happily ever after: Tracking emotions in mail and books. *Decision Support Systems*, 53(4), 730–741.
22. Morin O., & Acerbi A. (2017). Birth of the cool: a two-centuries decline in emotional expression in Anglophone fiction. *Cognition and Emotion*, 31(8), 1663–1675. https://doi.org/10.1080/02699931.2016.1260528 PMID: 27910735
23. Roivainen E. (2013). Frequency of the use of English personality adjectives: Implications for personality theory. *Journal of Research in Personality*, 47(4), 417–420.
24. Roivainen E. (2015). Personality adjectives in twitter tweets and in the Google books corpus. An analysis of the facet structure of the openness factor of personality. *Current Psychology*, 34(4), 621–625.
25. Ellis D. A., Wiseman R., & Jenkins R. (2015). Mental representations of weekdays. *PloS One*, 10(8), e0134555. https://doi.org/10.1371/journal.pone.0134555 PMID: 26288194
26. Hills T. T., & Adelman J. S. (2015). Recent evolution of learnability in American English from 1800 to 2000. *Cognition*, 143, 87–92. https://doi.org/10.1016/j.cognition.2015.06.009 PMID: 26117487
27. Virues-Ortega J., & Pear J. J. (2015). A history of "behavior" and "mind": Use of behavioral and cognitive terms in the 20th century. *The Psychological Record*, 65(1), 23–30.
28. FAZ. (n.d.). Verteilung der Weltbevölkerung nach Religionen in den Jahren 1900 und 2010. In Statista–Das Statistik Portal. Retrieved on 06.03.2019 from https://de.statista.com/statistik/daten/studie/256878/umfrage/verteilung-der-weltbevoelkerung-nach-religionen/.
29. Greenfield P. M. (2009). Linking social change and developmental change: shifting pathways of human development. *Developmental Psychology*, 45(2), 401. https://doi.org/10.1037/a0014726 PMID: 19271827
30. Triandis H. C. (2018). *Individualism and Collectivism*. Routledge.
31. Inglehart R., & Baker W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American Sociological Review*, 19–51.
32. Pargament K. I., Ishler K., Dubow E. F., Stanik P., Rouiller R., Crowe P., et al. (1994). Methods of religious coping with the Gulf War: Cross-sectional and longitudinal analyses. *Journal for the Scientific Study of Religion*, 347–361.
33. Ano G. G., & Vasconcelles E. B. (2005). Religious coping and psychological adjustment to stress: A meta-analysis. *Journal of Clinical Psychology*, 61(4), 461–480. https://doi.org/10.1002/jclp.20049 PMID: 15503316
34. Pargament K. I. (2001). *The psychology of religion and coping: Theory, research, practice*. Guilford Press.
35. Ritter R. S., & Preston J. L. (2013). Representations of religious words: Insights for religious priming research. *Journal for the Scientific Study of Religion*, 52(3), 494–507.
36. Oishi S., Graham J., Kesebir S., & Galinha I. C. (2013). Concepts of happiness across time and cultures. *Personality and Social Psychology Bulletin*, 39(5), 559–577. https://doi.org/10.1177/0146167213480042 PMID: 23599280
37. Uz I. (2014). Individualism and first person pronoun use in written texts across languages. *Journal of Cross-Cultural Psychology*, 45(10), 1671–1678.
38. Yu F., Peng T., Peng K., Tang S., Chen C. S., Qian X., et al. (2016). Cultural value shifting in pronoun use. *Journal of Cross-Cultural Psychology*, 47(2), 310–316.
39. Brysbaert M., Keuleers E., & New B. (2011). Assessing the usefulness of Google Books' word frequencies for psycholinguistic research on word processing. *Frontiers in Psychology*, 2, 27. https://doi.org/10.3389/fpsyg.2011.00027 PMID: 21713191
40. Mann, J., Zhang, D., Yang, L., Das, D., & Petrov, S. (2014). Enhanced search with wildcards and morphological inflections in the Google Books Ngram Viewer. In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations (pp. 115–120).
41. Bentley R. A., Acerbi A., Ormerod P., & Lampos V. (2014). Books average previous decade of economic misery. *PloS One*, 9(1), e83147. https://doi.org/10.1371/journal.pone.0083147 PMID: 24416159
42. Hughes J. M., Foti N. J., Krakauer D. C., & Rockmore D. N. (2012). Quantitative patterns of stylistic influence in the evolution of literature. *Proceedings of the National Academy of Sciences*, 109(20), 7682–7686.

43. Froese P. (2004). Forced secularization in Soviet Russia: Why an atheistic monopoly failed. *Journal for the Scientific Study of Religion*, 43(1), 35–50.

44. Koplenig A. (2017). The impact of lacking metadata for the measurement of cultural and linguistic change using the Google Ngram data sets—Reconstructing the composition of the German corpus in times of WWII. *Digital Scholarship in the Humanities*, 32(1), 169–188.

45. Scheff T. (2015). Toward defining basic emotions. *Qualitative Inquiry*, 21(2), 111–121.

46. Rossi E., Mortimer J., & Rossi K. (2013). Therapeutic hypnosis, psychotherapy, and the digital humanities: The narratives and culturomics of hypnosis, 1800–2008. *American Journal of Clinical Hypnosis*, 55(4), 343–359. https://doi.org/10.1080/00029157.2012.696078 PMID: 23724569

47. Mooijman M., Meindl P., Oyserman D., Monterosso J., Dehghani M., Doris J. M., et al. (2018). Resisting temptation for the good of the group: Binding moral values and the moralization of self-control. *Journal of Personality and Social Psychology*, 115(3), 585–599. https://doi.org/10.1037/pspp0000149 PMID: 28604018

48. Roivainen E. (2014). Changes in word usage frequency may hamper intergenerational comparisons of vocabulary skills: An Ngram analysis of wordsum, WAIS, and WISC test items. *Journal of Psychoeducational Assessment*, 32(1), 83–87.

49. Kesebir S., & Kesebir P. (2017). A growing disconnection from nature is evident in cultural products. *Perspectives on Psychological Science*, 12(2), 258–269. https://doi.org/10.1177/1745691616662473 PMID: 28346112

50. Genovese J. E. (2015). Interest in astrology and phrenology over two centuries: A Google Ngram study. *Psychological Reports*, 117(3), 940–943. https://doi.org/10.2466/17.PR0.117c27z8 PMID: 26595286

51. Li Y., Tan X., Huang Z., & Liu L. (2017). Relationship between collectivism and corruption in American and Chinese books: A historical perspective. *International Journal of Psychology*.

52. Grossmann I., & Varnum M. E. (2015). Social structure, infectious diseases, disasters, secularism, and cultural change in America. *Psychological Science*, 26(3), 311–324. https://doi.org/10.1177/0956797614563765 PMID: 25656275

Third Research Article:

The Influence of Social Change and Human Development on Emotions:

Evidence from China

Nadja Younes, Laura Höhner, and Ulf-Dietrich Reips

University of Konstanz

Abstract

This article examines the prevalence of positive and negative emotions in China based on the country's recent transformative changes in its social ecology. Conceptual framework of this study is the *social change and human development theory* that argues that changes in sociodemographic factors such as rapid economic growth or urbanization contribute to the development towards a more individualistic culture. There are fundamental differences in the perception and expression of emotions between individualistic and collectivistic cultures. In particular, avoiding negative and promoting positive emotions is expected to resonate with the cultural ideal of individualistic societies, whereas the avoidance of negative and promotion of positive emotions is not prioritized in collectivistic cultures. Based on China's recent development towards a more individualistic society the current study investigates whether there is a related shift in the prevalence of positive and negative emotions. By using the Google Books Ngram Viewer, we show that frequencies of positive (negative) emotion words exhibit an upward (a downward) trend between the years 1970 and 2008. Furthermore, a positive (negative) correlation between frequencies of positive (negative) emotion words and China's rising level of wealth, urbanization, and formal education suggests that values and behavior shift along with changes in ecology.

Introduction

During the last decades, collectivistic China has undergone massive transformations, driven by numerous socio-political changes such as the Cultural Revolution from 1966 to 1976 and the economic reforms in 1978. Particularly since the latter, China's economy has been substantially growing, leading to rising levels of wealth, formal education, and a continuous movement from rural towards urban environments (Zeng & Greenfield, 2015). As part of her theory on social change and human development, Greenfield (2009) suggests that such ecological shifts encourage cultural development, particularly, the development from collectivism towards individualism. The theory relies on Tönnies (1887/1957) conceptualization of a *Gemeinschaft* (community) and a *Gesellschaft* (society). A *Gemeinschaft* describes relatively impoverished, rural, and low-tech environments with rather low levels of formal education, whereas a *Gesellschaft* relates to relatively wealthy, urbanized, and high-tech environments with well-developed educational systems. Greenfield (2009) links individualistic values, behavior, and psychology to a *Gesellschaft* and collectivistic values, behavior, and psychology to the features of a *Gemeinschaft*. In line with China's recent development towards a *Gesellschaft's* direction, several studies provide evidence for the culture's development towards a more individualistic society (see, e.g., Moore, 2005; Yan, 2009; Sun & Wang, 2010; Cai et al., 2012; Xu & Hamamura, 2014; Hamamura & Xu 2015; Zeng & Greenfield, 2015; Yu et al.; 2016), particularly, by pointing out that "the Chinese are increasingly prioritizing individualist factors in assessments of their own happiness and life satisfaction" (Steele & Lynch, 2013, p. 450).

In this paper, we took advantage of China's unique cultural context and massive transformative changes during the last decades. In particular, we investigated whether there is a related shift in the prevalence of positive and negative emotions based on fundamental differences in the perception and expression of emotions

between individualistic and collectivistic cultures. Oishi and colleagues (2013), for example, provide an overview of variations in the concept of happiness across cultures, with collectivistic cultures being more likely to have luck-based concepts, whereas individualistic societies consider happiness as actively pursuable. Happiness, life-satisfaction, and individual well-being are conscious goals for individualists. In particular, individualistic cultures pursue the maximization of pleasure and strive for a minimization of displeasure because both aspects significantly contribute to the individual's well-being (Schimmack et al., 2002). In accordance with this notion, Kuppens et al. (2008) report that life satisfaction relates more closely to the experience of positive emotions than to the absence of negative emotions, with negative emotions being more negatively associated with life satisfaction in individualistic than in collectivistic countries. In fact, negative emotions like sorrow, hostility, and anger are even perceived as dangerous or noxious (Potter, 1988). Accordingly, the desire to avoid negative emotions and promote positive emotions, one's autonomy, as well as a positive view of the self, relates to the cultural ideal of individualistic societies, which emphasizes the pursuit of standing out and feeling good.

In contrast, in rather collectivistic cultures, paying attention to cultural norms is of greater importance than maximizing pleasure. The achievement of group harmony, e.g., by avoiding interpersonal disharmony, is considered one of the most important goals (Ho, 1979; Kuppens et al., 2008). In this respect, "the individual's actual position on an issue is subordinated to his or her desire to protect the group's integrity by side stepping open disagreement" (Wilson 1974, p.196). The pursuit of group harmony thus often causes the bearing of negative emotions. Therefore, the avoidance of negative emotions is infrequent and not prioritized because its impact on life satisfaction is considered to be of minor importance (Kuppens et al., 2008). Indeed, Potter (1988) points out that in collectivistic cultures, negative emotions are openly expressed because they are not considered to be intrinsically dangerous. The capacity

to work and suffer for others thereby reflects the foundation for social as well as intimate relationships, whereas in individualistic cultures positive emotions such as, e.g., *love* serve as the symbolic basis for social relationships.

By linking China's development towards a more individualistic society to the differences in the prevalence of positive and negative emotions in collectivistic and individualistic cultures, we hypothesize that the prevalence of positive emotions increased while the prevalence of negative emotions decreased over time. In addition, to study the relationship between changes in ecological factors and emotions, we examine to what extent shifts in the prevalence of negative and positive emotions are directly related to China's rising level of wealth, formal education, and urbanization.

Method and Procedure

Language as an Assessment of Culture

In this paper, we investigate changes in the prevalence of positive and negative emotional expressions in the Chinese society by examining frequencies of positive and negative emotion words in Chinese books. The rationale behind this analysis is the assumption that cultural patterns can be inferred from cultural products such as writings (Hamamura & Xu, 2015). We conducted this analysis by using the Google Books Ngram Viewer – a publicly and freely accessibly search engine that charts word frequencies by dividing the total number of a word's yearly appearances by the total number of words in the Google Books Ngram Viewer's corpus in that year. The latest version of the overall corpus consists of about 8 million digitized books, published between 1500 and 2008. The books were provided by more than 40 libraries and split into 8 languages (English, German, Russian, French, Italian, Hebrew, Spanish and Chinese), which are stored in individual corpora. Accessing the latest version of the Google Books Ngram Viewer, the current study is based on a corpus of 302,652 Chinese books (Michel et al., 2011; Lin et al., 2012).

To conduct our analysis, we followed Zeng and Greenfield (2015) and focused on the period between 1970, the midst of the Chinese Cultural Revolution, and 2008. We tested for quantitative changes in emotion words by examining Pearson correlation coefficients between the emotion words' frequencies and years or sociodemographic factors, respectively.[1] Based on the idea that the natural frequency of a word is relevant for assessing its cultural change, we first focused on each word's raw frequency, independently. In a second step, we summed up the frequencies of single words per year (and emotion type, i.e., positive or negative), and ran an aggregated correlational analysis (see, e.g., Twenge et al., 2012 a,b, 2013). Accordingly, the more frequent an emotion word, the larger its proportional influence. Finally, we followed Younes and Reips (2019) by applying a standardization procedure that accounts for the proportional influence of single words as well as data-related trends.[2] In particular, we subtracted the summed z-scored frequencies of several common words, obtained from Lin et al. (2012), from the summed z-scored frequencies of our original emotion words.

To obtain a reference set of suitable common words, we considered Lin et al.'s (2012) list of very common Chinese words and compared them to Hughes et al.'s (2012) list of neutral words. We chose all words that appear on both lists. This procedure verifies that common words are free of any underlying meaning that might potentially bias the results. In accordance, we did not include pronouns based on their categorization as content-full words by previous literature (see, e.g., Twenge et al., 2012b, 2013; Uz, 2014; Hamamura & Xu, 2015; Yu et al., 2016). Because Hughes et al.'s (2012) list comprises only English terms, we asked two Chinese native speakers, who are also fluent in English, to translate Lin et al.'s (2012) list of Chinese common words to English. We then searched for all of the suggested translations on Hughes et al.'s (2012) word list. Applying this procedure, the final sample of suitable Chinese

[1] For all analyses we followed Hamamura and Xu (2015) and used a smoothing of 1.
[2] See Younes and Reips (2019) for a detailed explanation of the procedure.

common words consists of “在” (at, in), “不” (not, no), “也” (also, too, as well), “和” (and, together with), “与” (and, together with, against) “这” (this, these), “各” (all, each, every) “是” (is, are, am) “有” (there is, there are). Figure 1 visualizes the development of the common words’ frequencies over time and reinforces the concern that data influx may potentially bias the results and should be taken into account.

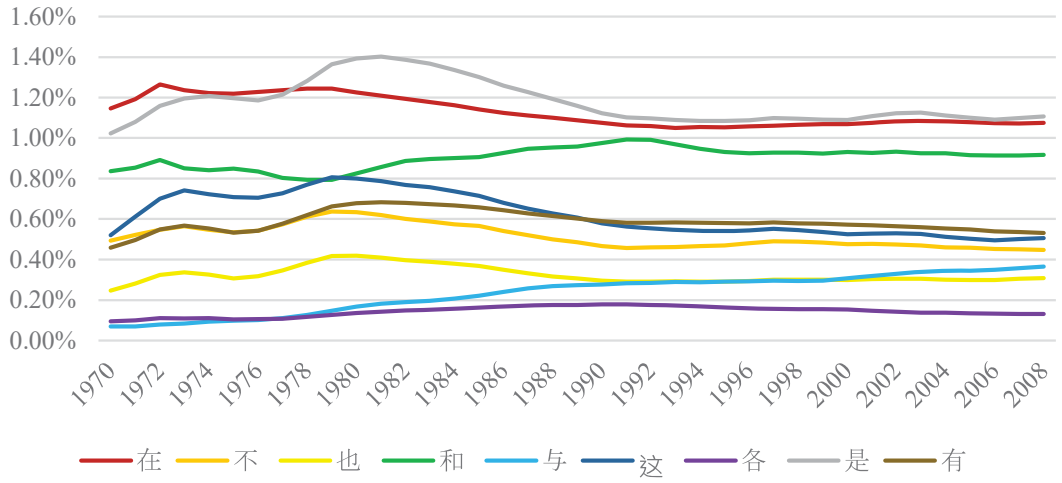

Figure 1. Pattern of Chinese common words from 1970 to 2008.

Emotion Words

We started with a set of 24 English emotion words, obtained from Fontaine et al. (2007). Because we only focus on positive and negative emotion words, we excluded the neutral emotions “interest” and “surprise” as well as one expression that consists of two words (“being hurt”). Hence, our final sample comprises 21 emotion words, with seven words describing positive and 14 words describing negative emotions. Again, two native speakers translated the list of English emotion words to Chinese. Independently from each other, both agreed on the same Chinese expressions. Table 1 provides an overview of the Chinese emotion words and their English translations.

Table 1. Overview of emotion words.

Positive Emotion Words	Negative Emotion Words	
开心 (joy)	害怕 (fear)	生气 (irritation)
幸福 (happiness)	悲伤 (sadness)	内疚 (guilt)
满意 (contentment)	羞愧 (shame)	仇恨 (hate)
乐趣 (pleasure)	绝望 (despair)	妒忌 (jealousy)
爱 (love)	愤怒 (anger)	鄙视 (contempt)
骄傲 (pride)	恐惧 (anxiety)	恶心 (disgust)
倾慕 (affection)	失望 (disappointment)	精神压力 (stress)

One common criticism related to research using the Google Books Ngram Viewer is that results may be prone to biases arising from poor optical character recognition (OCR), particularly with respect to semantic scanning errors, due to similarities in letters or signs. In their supporting online material, Michel et al. (2011) argue that in order to ensure high data quality, volumes with poor OCR quality were filtered out. However, due to comparatively fewer books in the Chinese corpus, no OCR filter was applied. Although previous research suggests that the Google Books Ngram Viewer's Chinese corpus "can be used validly in inferring cross-temporal trends in the Chinese culture" (Hamamura & Xu, 2015, p. 932), we aimed at verifying the reliability of our results by considering several sets of synonyms (Younes & Reips, 2018, 2019) as well as by taking the development of an additional set of control words, i.e., words with similar optical characters but different meanings, into account.

Synonyms

To verify that results are not driven by idiosyncrasies or varying meanings, we followed Younes and Reips (2018, 2019) and collected up to three synonyms (whenever possible) for each of the above stated Chinese emotion words using the thesaurus "*1700 Groups of Frequently Used Chinese Synonyms*". To ensure that the synonyms are adequate substitutes, we further followed Younes and Reips' (2019)

suggestion and asked several Chinese native speakers to rate the synonyms' suitability on a scale of 0 (no synonym) to 200 (perfect synonym). In particular, 12 Chinese native speakers rated our synonyms through an online survey.[3] All native speakers were students (83% female, average age 21 years) at a Chinese University and reported to have at least a high school diploma. Table 2 provides an overview of the dictionary-based synonyms and their respective average ratings.

Table 2. Overview of synonyms (dictionary-based) and their respective average ratings.

Positive Emotion Words	Synonyms	Average Rating
开心 (joy)	高兴 (glad, happy)	152.5
	愉快 (happy, joyful, cheerful)	151.5
幸福 (happiness)	福 (luck, happiness, good fortune, blessing)	83.8
满意 (contentment)	满足 (satisfied, contented)	127.9
	如意 (be satisfied, as one wishes)	100.9
乐趣 (pleasure)	兴趣 (interest)	112.2
爱 (love)	热爱 (ardently love, have deep love (or affection) for)	107.0
骄傲 (pride)	自豪 (have a proper sense of pride or dignity, be proud of sth.)	129.3
Negative Emotion Words	**Synonyms**	**Average Rating**
害怕 (fear)	恐惧 (fear, dread)	162.3
	畏惧(fear, dread)	159.3
	怕 (fear, dread, be afraid of)	153.8
悲伤 (sadness)	悲痛 (grieved, sorrowful)	142.7
羞愧 (shame)	惭愧 (be ashamed)	157.0
愤怒 (anger)	发怒 (get angry, flare up, lose one's temper)	151.1
	恼怒 (angry, resentful, enrage)	143.6
恐惧 (anxiety)	害怕 (be afraid, be scared, be in fear of)	169.3
失望 (disappointment)	扫兴 (have one's spirit dampened, feel	156.7

[3] We considered only complete forms.

	disappointed)	
	灰心 (lose heart, be discouraged)	131.1
生气 (irritation)	发火 (get angry, flare up, lose one's temper)	90.8
仇恨 (hate)	仇 (enemy, foe)	136.8
	愤恨 (indignantly resent; furiously detest)	123.8

Our final list of synonyms comprises 21 words, covering 14 out of the 21 original emotion words. The mean rating over all words, 135.4, indicates that on average, chosen synonyms mirror the original emotion words appropriately. However, since we did not obtain any synonym for 7 of the original emotion words using a dictionary-based approach, we also asked the Chinese native speakers to propose one adequate synonym for each original emotion word. To verify the appropriateness of the suggested synonyms, we only considered words that were mentioned by at least two persons and different from the synonyms that we already obtained through the dictionary-based approach.[4] Table 3 provides the final overview of proposed synonyms for positive and negative emotion words.

Table 3. Overview of synonyms (suggested by native speakers) for positive and negative emotion words with their respective translation and total number of mentions.

Positive Emotion Words	Synonyms	# of mentions
开心 (joy)	快乐 (happy, joyful, cheerful, delightful)	5
幸福 (happiness)	美满 (happy, very satisfactory)	4
满意 (contentment)	欣慰 (gratified, satisfied)	2
	合意 (suit, be to one's liking or taste)	2
乐趣 (pleasure)	趣味 (interest, delight, taste)	3
爱 (love)	喜欢 (like, love, be fond of, be keen on)	5
	喜爱 (love, have a liking for, be fond of, be keen on)	2

[4] To avoid the case that synonyms obtained through the dictionary-based approach were only repeated, we asked the 12 Chinese native speakers first to propose own synonyms and then to rate given synonyms.

	恋 (love, love affair)	2
骄傲 (pride)	傲慢 (haughty, arrogant, impudent)	2
	高傲 (supercilious, conceited, arrogant, haughty)	2
	自满 (self-satisfied, smug, complacent)	2
倾慕 (affection)	仰慕 (admire, hold in high esteem)	5
	爱慕 (admire, adore)	4

Negative Emotion Words	Synonyms	# of mentions
悲伤 (sadness)	难过 (feel sorry, feel bad, be distressed)	6
	伤心 (sad, grieved, sorrowful, broken-hearted)	2
羞愧 (shame)	愧疚 (feel ashamed and remorseful, be conscience-stricken)	2
绝望 (despair)	无望 (hopeless, impossible)	9
愤怒 (anger)	生气 (get angry, be furious)	8
	气愤 (indignant, enraged, furious)	2
失望 (disappointment)	失落 (lose)	3
	绝望 (giving up all hope, hopelessness, despair)	2
生气 (irritation)	愤怒 (indignation, fury, anger, wrath)	7
	气愤 (indignant, enraged, furious)	3
内疚 (guilt)	愧疚 (feel ashamed and remorseful, be conscience-stricken)	7
仇恨 (hate)	怨恨 (have a grudge against, resent, hate)	3
	仇视 (look upon with hostility, be hostile to, regard as an enemy)	2
	仇怨 (grudge, grievance, hatred)	2
妒忌 (jealousy)	嫉妒 (be jealous of, envy)	12
鄙视 (contempt)	蔑视 (despise, scorn, look down on, show contempt for)	4
	轻视 (despise, belittle, scorn, look down on)	3

	看不起 (look down upon, belittle, scorn, despise)	2
恶心 (disgust)	厌恶 (detest, abhor, abominate, loathe)	5
	呕吐 (vomit, throw up)	2
精神压力 (stress)	思想包袱 (sth. weighing on one's mind, mental burden)	2

Note. Translations are obtained through the *New Age Chinese-English Dictionary*.

Control Words

As a second robustness check, we developed an approach that can be used to test whether results are likely driven by systematic scanning errors within the Chinese corpus. In particular, we considered frequencies of words that are represented by similar optical characters as our original Chinese emotion words but have a different meaning. Consequently, the pattern of these frequencies should be different from the pattern of the original emotion words. To facilitate this robustness check, we used the *New Age Chinese-English Dictionary* that lists Chinese expressions according to their first character. In a first step, we located the original emotion word. In a second step, we identified the first entry above (upper entry) and the first entry below (lower entry) the original emotion word that has both, an identical number of total characters and the same first character as the original Chinese emotion word.[5] As indicated in Table 4, applying this procedure, we obtained control words for four positive and five negative emotion words.

Table 4. Overview of control words.

Positive Emotion Words	Upper Entry	Lower Entry
开心 (joy)	开销 (pay expenses)	开衅 (outbreak of hostilities)
满意 (contentment)	满眼 (fill one's eyes)	满语 (Manchu)
爱 (love)	瑗 (fine quality jade)	薆 (cover, shelter)

[5] In case the Google Books Ngram Viewer did not display any frequencies for a selected control word, we took the next available word listed in the dictionary. Words were considered as control words when their meaning was different from the meaning of the corresponding emotion. Since the word for "love" only consists of one character, we considered only the first character with additional radicals. Thereby, we only considered lower entries because otherwise the whole symbol would have been changed.

倾慕 (affection)	倾角 (dip, angle of dip)	倾弃 (dumping)
Negative Emotion Words	Upper Entry	Lower Entry
害怕 (fear)	害处 (damage)	害臊 (feel ashamed)
绝望 (despair)	绝嗣 (without offspring)	绝响 (lost music, anything lost to the modern world)
生气 (irritation)	生漆 (raw lacquer)	生前 (during one's lifetime)
内疚 (guilt)	内径 (internal diameter)	内举 (recommend one's relative)
鄙视 (contempt)	鄙人 (humble servant)	鄙俗 (vulgar, philistine)

Results

The Prevalence of Emotions over Time

Focusing on the development of individual raw frequencies of positive emotion words over time, i.e., computing the correlation between raw frequencies and years, we observe a significant upward trend during the relevant years 1970 to 2008 for the Chinese emotion words "joy" ($r=0.77$, $p<0.001$), "pleasure" ($r=0.75$, $p<0.001$), "affection" ($r=0.58$ $p<0.001$), and "contentment" ($r=0.53$, $p<0.001$). "Love" ($r=0.14$, $p>0.1$) exhibits a positive but non-significant trend over time. The only positive emotion words that show a negative time trend are "happiness" ($r=-0.56$, $p<0.001$) and "pride" ($r=-0.75$, $p<0.001$). Figure 2 visualizes the pattern of positive emotion words.

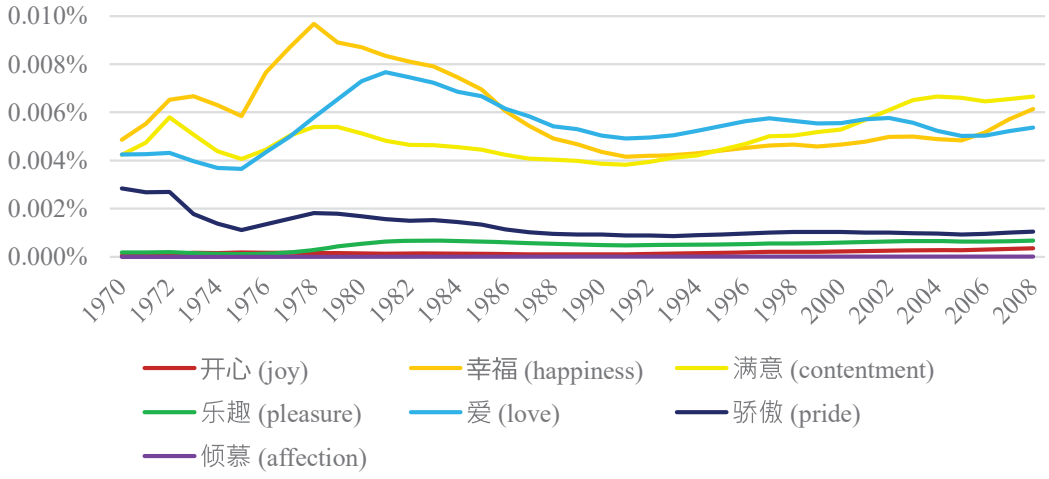

Figure 2. Pattern of positive emotion words in the Chinese corpus from 1970 to 2008.

The frequencies of positive emotion words peaked in the middle of the 1970s. The graph further indicates that the 1980s were a period in which the frequencies of positive emotion words declined substantially. However, since the early 1990s, we observe an increase in frequencies for the majority of positive emotion words. As indicated by Figure 3, the pattern of negative Chinese emotion words is more differentiated.

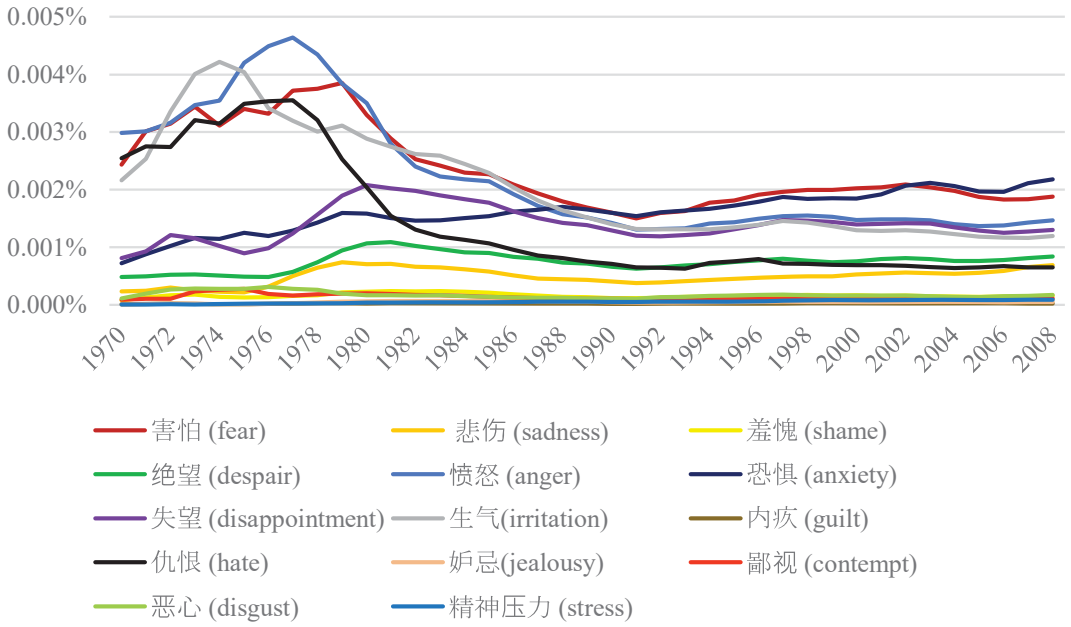

Figure 3. Pattern of negative emotion words in the Chinese corpus from 1970 to 2008.

The visual inspection postulates a general positive trend until the end of the 1970s. In the subsequent years, the trend becomes rather constant and for several words strongly negative.

The correlation analysis suggests that the individual raw frequencies of "stress" ($r=0.98$, $p<0.001$), "anxiety" ($r=0.94$, $p<0.001$), "guilt" ($r=0.48$, $p<0.01$), "jealousy" ($r=0.44$, $p<0.01$), "sadness" ($r=0.43$, $p<0.01$), and "despair" ($r=0.33$, $p<0.05$) rose significantly over time. In contrast, the words for "irritation" ($r=-0.86$, $p<0.001$), "hate" ($r=-0.83$, $p<0.001$), "anger" ($r=-0.81$, $p<0.001$), "fear" ($r=-0.75$, $p<0.001$), "contempt" ($r=-0.61$, $p<0.001$), and "disgust" ($r=-0.53$, $p<0.001$) show the predicted

negative trend over time. Finally, “disappointment” ($r=0.05$, $p>0.1$) and “shame” ($r=-0.12$, $p>0.1$) display no significant pattern.

Raw frequencies of words from the Google Books Ngram Viewer can be misleading. To thus obtain a more reliable picture, i.e., one corrected for a number of issues with the Google Books Ngram Viewer, we followed Younes and Reips (2019) and conducted a set of analyses that can improve the reliability of results of Google Ngram studies. First, we calculated correlation coefficients by summing up the individual raw frequencies of the original positive and negative emotion words, respectively. Second, we repeated both analyses several times by adding three different sets of synonyms (procedure 4, Younes & Reips, 2019). In particular, in a first approach, we added the set of synonyms obtained through the dictionary-based approach. In a second approach, we added the set of synonyms suggested by the Chinese native speakers. In a third approach, we added a set consisting of the synonyms obtained through the dictionary-based approach as well as the synonyms that were suggested by the Chinese native speakers. However, to solely take the most adequate synonyms into account, we dropped synonyms obtained through the dictionary-based approach if their rating was lower than the sample mean of 135.4. Third, we re-conducted each of the above-mentioned analyses by accounting for data influx, i.e., variations in the corpus over time, and unequal weights, i.e., the overproportioned influence of single words. In particular, in this advanced method, we subtracted the summed z-scored frequencies of various common words obtained from Lin et al. (2012), from the summed z-scored frequencies of the emotion words (procedure 5, Younes & Reips, 2019). Finally, as an additional robustness check, we computed the correlation coefficients between years and yearly differences between the frequencies of original emotion words and the frequencies of control words, both, using raw frequencies and the advanced method. Table 5 presents the results.

Table 5. Overview of results.

	I Σ Raw frequencies	II Σ z−scores – Σ z−scores of common Chinese words
Positive emotion words	r=-0.18	r= 0.60***
Negative emotion words	r=-0.77***	r= 0.09
Positive emotion words including synonyms (dictionary)	r=-0.09	r= 0.29†
Negative emotion words including synonyms (dictionary)	r=-0.73***	r=-0.45**
Positive emotion words including synonyms (native speakers)	r= 0.29†	r= 0.70***
Negative emotion words including synonyms (native speakers)	r=-0.80***	r=-0.30†
Positive emotion words including synonyms (dictionary + native speakers)	r=-0.16	r= 0.58***
Negative emotion words including synonyms (dictionary + native speakers)	r=-0.77***	r=-0.41**
Difference between control & emotion words	r=-0.27†	r= 0.48**

Note. †p<.1, **p < .01, ***p < .001

After accounting for unequal weights and data influx over time, i.e., using the advanced method (Table 5, Column II), the original set of positive emotion words displays a significant positive trend, whereas negative emotion words do not show any significant pattern. The picture for negative emotion words, however, changes after including the different sets of synonyms. While there is a significant positive trend for positive emotion words, the trend for negative emotion words turned significantly negative after including any set of additional synonyms. Finally, the pattern of the control words is significantly different from the pattern of the original emotion words, indicating that the results of this study are unlikely driven by a general bias within the Chinese corpus.

Emotions and Ecology

Based on the assumption that the shift towards a more individualistic society relates to changes in the Chinese ecology, we followed Zeng and Greenfield (2015) and

calculated correlation coefficients between frequencies of positive and negative emotion words (with and without synonyms) and the sociodemographic factors wealth (per capital gross national income), formal education (tertiary school enrollment), and urbanization (urban population). The data on sociodemographic factors is provided by the World Bank.[6] Table 6 presents an overview of the results for the different sociodemographic factors.

Table 6. Overview of results considering different sociodemographic factors.

	I Σ Raw frequencies	II Σ z−scores – Σ z−scores of common Chinese words
Panel A: Wealth		
Positive emotion words	r= 0.00	r= 0.82***
Negative emotion words	r=-0.50**	r= 0.35*
Positive emotion words including synonyms (dictionary)	r=-0.08	r= 0.51***
Negative emotion words including synonyms (dictionary)	r=-0.50**	r=-0.27†
Positive emotion words including synonyms (native speakers)	r= 0.36*	r= 0.72***
Negative emotion words including synonyms (native speakers)	r=-0.54***	r=-0.09
Positive emotion words including synonyms (dictionary + native speakers)	r= 0.04	r= 0.63***
Negative emotion words including synonyms (dictionary + native speakers)	r=-0.52***	r=-0.24
Panel B: Formal Education		
Positive emotion words	r= 0.10	r= 0.82***
Negative emotion words	r=-0.49**	r= 0.33†
Positive emotion words including synonyms (dictionary)	r= 0.03	r= 0.58***
Negative emotion words including synonyms (dictionary)	r=-0.49**	r=-0.27
Positive emotion words including synonyms (native speakers)	r= 0.45**	r= 0.77***
Negative emotion words including synonyms (native speakers)	r=-0.53**	r=-0.08
Positive emotion words including	r= 0.12	r= 0.69***

[6] See https://data.worldbank.org/.

synonyms (dictionary + native speakers)		
Negative emotion words including synonyms (dictionary + native speakers)	r=-0.52**	r=-0.23
Panel C: Urbanization		
Positive emotion words	r=-0.16	r= 0.73***
Negative emotion words	r=-0.72***	r= 0.17
Positive emotion words including synonyms (dictionary)	r=-0.14	r= 0.36*
Negative emotion words including synonyms (dictionary)	r=-0.70***	r=-0.44**
Positive emotion words including synonyms (native speakers)	r= 0.29†	r= 0.72***
Negative emotion words including synonyms (native speakers)	r=-0.75***	r=-0.27†
Positive emotion words including synonyms (dictionary + native speakers)	r=-0.13	r= 0.60***
Negative emotion words including synonyms (dictionary + native speakers)	r=-0.73***	r=-0.40**

Note. †p<.1, *p < .05, **p < .01, ***p < .001

Based on the fact that urbanization, wealth, and formal education have been constantly rising, also the results displayed in Table 6 are mainly in line with our hypotheses. In particular, using the advanced method (Column II), we find that the original set of frequencies of positive emotion words exhibits a significant positive correlation with all sociodemographic factors. Additionally, for the combined sets of frequencies of positive emotion words and all types of synonyms, we observe a significant positive correlation with all sociodemographic factors. In contrast, frequencies of the original set of negative emotions words also display a positive correlation. By including any set of synonyms, we obtain, however, the expected negative correlation that remains significant for urbanization.

Discussion

By using the Google Books Ngram Viewer to investigate changes in the frequencies of Chinese negative and positive emotion words between the years 1970 and 2008, this article analyzes the impact of social changes on the prevalence of emotions in China. Based on the theory of social change and human development (Greenfield,

2009), we hypothesize that related to China's development towards a *Gesellschaft* over the last few decades, changes in the prevalence of emotions are also observable. In particular, based on differences in the perception and expression of emotions between individualistic and collectivistic cultures (Potter, 1988; Schimmack et al., 2002; Kuppens et al., 2008), the frequencies of positive emotion words are expected to increase, whereas the frequencies of negative emotion words are expected to decrease.

On an aggregated level and after considering data influx, unequal weights, and synonyms, frequencies of positive emotion words exhibit the expected increase over time. In addition, frequencies of negative emotion words display the expected decrease over time. However, by considering the sets of the original emotion words without adding any synonyms, we observe a significant positive trend for positive emotion words and a positive but insignificant trend for negative emotion words. Although we believe that the mixed result obtained for the set of negative emotion words relates to a methodological bias that can be mitigated by the consideration of additional synonyms, it might be also possible that this result mirrors the impact of the role of cultural heritage on social change (Huntington, 1996). Ingelhart and Baker (2000), for example, provide profound evidence for the assumption that "the broad cultural heritage of a society […] leaves an imprint on values that endures despite modernization" (p.19). Hence, the unexpected increase for some negative emotions could likely be the outcome of a reaction against the new values in China. Further, Greenfield and Zeng (2015) explain the counter-movement of some of their examined collectivistic values with China's current stage of urbanization. A part of the population still lives a simple life with little wealth and low technology, whereas the other part of the population is urban, wealthier, and enjoys high technical standards. In line with the assumption that especially the latter part of the population is more closely linked to the adaption of individualistic values, Zeng and Greenfield (2015) indicate that there are co-existing value systems in current China, which are likely responsible for the partial increase of "*Gemeinschaft*" characteristics.

Finally, by directly linking positive and negative emotion words to sociodemographic changes, our results indicate that urbanization, increases in the level of wealth, and formal education facilitate changes in the prevalence of emotions. In particular, after considering the influx of data, unequal weights, and additional sets of synonyms, we observe a significant positive correlation between all sociodemographic factors and the frequencies of positive emotion words. In contrast, frequencies of negative emotion words show a significant negative correlation with urbanization, and a negative but mostly non-significant correlation with formal education and wealth. However, as part of the most advanced investigation procedure (considering influx of data, unequal weights, and synonyms obtained through a dictionary and suggested by Chinese native speakers), we excluded synonyms that received a below-average rating by the Chinese native speakers. By also including these synonyms, the correlations between frequencies of negative emotion words and formal education as well as wealth were weakly negative. Overall, the results suggest that values and behavior shift along with ecological changes.

Like any other study, this paper is not without limitations. Although the consideration of the development towards individualism as the sole explanation of emotional development does certainly not capture all facets of such a complex and diverse field, this paper still sheds further light on the culture-sensitivity of values and behavior. A further point of criticism that we already noted relates to the Google Books Ngram Viewer's underlying methodology, particularly to the lack of an appropriate OCR filter for the Chinese corpus (Michel et al., 2011, supporting online material). We aimed at addressing this issue not only with the use of different sets of synonyms but also by comparing the frequency pattern of the original emotion words to the frequency pattern of control words, i.e., words with similar Chinese characters but a different meaning. Since control words did not display a similar development as our positive and negative emotion words, potential flaws in Google Books Ngram Viewer's OCR are likely not a major driver of our results. An additional concern that

particularly applies for the Chinese corpus is the fact that Google's library partners are mainly based in Western countries. Hence, a large part of the underlying Chinese literature may be books that drew attention to Western audiences and may not properly represent China's actual cultural development (Xu & Hamamura, 2014; Hamamura & Xu, 2015; Zeng & Greenfield, 2015). In this regard, we want to point out that the assumption that print culture represents culture as a whole may certainly not always hold. However, we believe that this study allows at least for an approximate assessment of the culture-sensitive development of emotions in China.

General Discussion

The three research articles of this thesis are concerned with the use of the Google Books Ngram Viewer. By examining the development of social change over the course of time, each study generates important empirical knowledge on the development towards individualism in different cultures. In addition, the thesis provides beneficial guidance concerning the methodology that can be used to increase the reliability of results. In what follows, I will explicitly point out the research articles' scientific contribution, provide implications for future research, and draw a final conclusion.

Contribution to Knowledge

The main contribution of this thesis is two-fold. First, each research article contributes to the vast amount of literature that focuses on the general development of individualism and collectivism over time (see, e.g., Triandis, 2018, for a review). In this respect, the thesis particularly adds to the novel strand of research that uses the Google Books Ngram Viewer to derive assumptions on cultural change (e.g., Kesebir & Kesebir, 2012; Twenge et al. 2012a,b; Oishi et al., 2013; Greenfield, 2013; Hamamura & Xu, 2015; Xu & Hamamura, 2014; Zeng & Greenfield, 2015; Yu et al. 2016). With regard to the replication of Greenfield's (2013) findings for the German corpus, the first research article provides further evidence for the global universality of the recently developed social change and human development theory that suggests that sociodemographic shifts move values from collectivistic to individualistic manifestations (Greenfield, 2009). In this regard, the first and the second research article especially add to the better understanding of cultural change by revealing a bidirectional change for individualistic and collectivistic values. In particular, the first research article shows a general increase in individualistic values and a general decrease in collectivistic values over time but highlights a reversal during the years of the Nazi Regime and World War II. In accordance, the second research article

demonstrates a similar reversal for collectivistic values despite a general negative trend. By putting its focus on religion, the second research article further contributes to the strand of research that discusses religion as a coping strategy (see, e.g., Pargament, 2001, for a review). Finally, by highlighting the development towards individualism for China, the third research article confirms previous findings on the culture's development towards a more individualistic society also for rather collectivistic countries (Moore, 2005; Yan, 2009; Sun & Wang, 2010; Cai et al., 2012; Xu & Hamamura, 2014; Hamamura & Xu 2015; Zeng & Greenfield, 2015; Yu et al.; 2016). Thereby, the third research article simultaneously adds to the strand of (Google Ngram) research that focuses on emotions (e.g., Mohammed, 2012; Acerbi et al., 2013; Scheff, 2015; Morin & Acerbi, 2017).

The second share of the thesis' main contribution is related to each of the research articles methodological improvements and refinements. Because the Google Books Ngram Viewer offers impressive opportunities for research purposes, it has been increasingly used in recent academic research. Thus, it is important to continuously improve on the method to mitigate shortfalls that may potentially bias derived assumptions (see, e.g., Gooding, 2012; Pechenick et al., 2015; Pettit, 2016). In this respect, each of the three presented research articles improves on the previously applied method. In particular, to reduce the probability of wrong assumptions caused by the prevalence of, e.g., optical character recognition errors, flaws related to messy metadata, idiosyncrasies, and the disproportionately large influence of single words, the first research article introduces the systematic use of synonyms for the German corpus. In addition, the second article constitutes a guideline of state-of-the-art research methods that should be considered when using the Google Books Ngram Viewer. The guideline focuses on the synonym approach and four further methodological procedures that all aim at increasing the reliability of derived results. In particular, to test the universality of a certain theory or to compare results for various cultural environments, the use of different language corpora may be beneficial.

Further, by relying on a corpus that is not heavily impacted by scientific literature, the English fiction corpus allows to reinforce the findings derived from the American and British English corpora. In addition, by using the feature "_INF", frequencies of similar words can be compared. Thereby, inconsistencies among semantically similar words can be detected and taken into account. Finally, the application of a newly developed standardization procedure accounts for both, unequal frequency weights, and the influx of data over time. By presenting a roadmap on how to apply several procedures in combination and by highlighting the empirical and theoretical consequences of their neglect, the second research article points out how sensitive results can be to the methodological choices a researcher can make. Finally, by focusing on Chinese, the third research article applies the previously introduced procedures to a language that is not based on a Latin alphabet. In addition, it adds a newly developed corpus-specific procedure that suggests how to evaluate whether, e.g., errors in optical character recognition, may bias the results. In particular, by comparing frequency patterns for a set of content-related words and a further set of optically similar but completely unrelated expressions, researchers can reinforce the reliability of results when using the Chinese corpus.

Implications for Future Research

This doctoral thesis offers valuable insights associated with the improvement on the method and the examination of cultural change using the Google Books Ngram Viewer. Both aspects offer interesting directions for future research. In particular, by reinforcing universality of Greenfield's (2009) social change and human development theory, this thesis suggests that an investigation of cultural change can be conducted by using any of the Google Books Ngram Viewer's available corpora. In this regard, considering rather understudied corpora such as the Hebrew corpus may provide further beneficial insights and additional knowledge on, for example, the Israeli population's development of values over time. In addition, both, the first and the

second research article focus on World War II to investigate the consequences of a momentous historical event on a population's cultural development. However, also other historically significant events might be interesting to study. The consideration of, e.g., additional wars, changes of governing, i.e., from democracy to a regime and vice versa, or foundations of unions or political parties may have also led to bidirectional value shifts. Further, although this thesis is mainly concerned with the development of individualism and collectivism over time, future research may use the Google Books Ngram Viewer to gain insight into various fields. Previous research, for example, also focused on topics such as gender differences (Twenge et al., 2012b; Del Giudice, 2012; Ye et al., 2018), personality (Roivainen 2013, 2015), or cognition (Hills & Adelman, 2015; Ellis et al., 2015; and Virues-Ortega & Pear, 2015).

In terms of methodology, research using the Google Books Ngram Viewer has primarily focused on visual inspection, correlation analysis, and linear regression models. Accordingly, also this thesis takes these kind of methods into account. While they are suitable to reveal cultural changes, they do not establish a clear cause-effect relationship. In this regard, future research may put its focus more on the establishment of causality. Further, one of the most challenging problems with these types of analysis is the number of words that has to be considered to obtain real effects. For this reason, it would be beneficial to develop a standard, possibly by using some kind of bootstrap method to estimate the minimum amount of words needed to obtain reliable effects. The same holds true for the number of years to be included in the analysis. Another point that might be interesting to examine is the quantification of the quality improvement that comes by using the Google Books Ngram Viewer in general. To examine cultural change, previous studies also considered external corpora such as the Gutenberg online library that is, for example, not affected by the problem of messy metadata but consists of significantly fewer books (see, e.g., Morin & Acerbi, 2017). Future research may strive for conducting some kind of horse race by using several corpora and even go beyond the scope of books in order to determine what type of

cultural product analysis is the most appropriate one to capture cultural change. Finally, this thesis presents procedures on how to conduct Google Ngram studies using the English, German, French, Italian, Spanish, and even Chinese corpora. As already mentioned, we did not study the Hebrew corpus. However, just as for Chinese, the visual similarity between letters of the Hebrew alphabet may also cause biases arising from errors in, e.g., optical character recognition. Particularly the letters Dalet, Hey, and Final Kaf might be difficult to distinguish for a computer algorithm. The same holds true for Russian. Within the scope of the second research article, we did not deeply focus on this corpus but only used it to discuss potential biases arising from censorship and propaganda. Future research should take the specific characteristics of the Hebrew and Cyrillic alphabet into account and strive for developing additional corpora-specific investigation procedures that are able to address potential biases.

Conclusion

Technological progress has spread knowledge globally. In this regard, also research methods, i.e., strategies to generate knowledge, have seen unprecedented opportunities. One of these opportunities has been the launch of the Google Books Ngram Viewer in 2010. The online tool allows to chart a word's annual relative appearance in a corpus of millions of scanned books and thereby makes it possible to quantify cultural change. The present thesis uses the Google Books Ngram Viewer to investigate the development of cultural change over time. In line with previous research, each of the thesis' three research articles highlights a constant development towards individualism also for rather collectivistic societies like China. However, there is also evidence for an expected reversal during severe times of crisis such as wars, suggesting that cultural changes can be bidirectional. Further, by focusing on China, the thesis adds that cultural change also influences the expression of emotions. In addition to the newly established cultural insights, each of the three studies improves on the method by addressing several of the Google Books Ngram Viewer's

shortfalls, i.e., scanning errors, flaws related to messy metadata, idiosyncrasies, and the disproportionately large influence of scientific literature or single words that may potentially bias derived assumptions. The methodological recommendations cumulate in a state-of-the-art guideline on how to conduct Google Ngram research and are extended by a corpus-specific procedure for Chinese. To summarize, this thesis does not only add to research on cultural development but suggests that the Google Books Ngram Viewer offers unprecedented research opportunities. Further, by equipping researchers with a highly relevant and extensive set of methodological tools, this thesis allows to generate reliable knowledge and unfold the Google Books Ngram Viewer's full potential.

Author Contributions

First Research Article

Nadja Younes did the data collection and analyses. Ulf-Dietrich Reips added to the methodology. Nadja Younes and Ulf-Dietrich Reips designed the study and wrote this article jointly.

Second Research Article

Nadja Younes did the data curation and investigation. Ulf-Dietrich Reips added to the methodology. Nadja Younes and Ulf-Dietrich Reips designed the study and wrote this article jointly.

Third Research Article

Nadja Younes did the analyses. Laura Höhner did the data collection. Laura Höhner und Ulf-Dietrich Reips added to the methodology. Nadja Younes, Laura Höhner, and Ulf-Dietrich Reips designed the study and wrote this article jointly.

References

Acerbi, A., Lampos, V., Garnett, P., & Bentley, R. A. (2013). The expression of emotions in 20th century books. *PloS One*, *8*(3), e59030.

Ano, G. G., & Vasconcelles, E. B. (2005). Religious coping and psychological adjustment to stress: A meta-analysis. *Journal of Clinical Psychology*, *61*(4), 461-480.

Bähr, J. (2007). Entwicklung von Urbanisierung. *Berlin Institut für Bevölkerung und Entwicklung, Berlin*. Retrieved from http://www.berlin-institut.org/online-handbuchdemografie/bevoelkerungsdynamik/auswirkungen/entwicklung-von-urbanisierung.html.

Bentley, R. A., Acerbi, A., Ormerod, P., & Lampos, V. (2014). Books average previous decade of economic misery. *PloS One*, *9*(1), e83147.

Brysbaert, M., Keuleers, E., & New, B. (2011). Assessing the usefulness of Google Books' word frequencies for psycholinguistic research on word processing. *Frontiers in Psychology*, *2*, 27.

Cai, H., Kwan, V. S. Y., & Sedikides, C. (2012). A sociocultural approach to narcissism: The case of modern China. *European Journal of Personality*, *26*(5), 529–535.

Chalfant, H. P., & Heller, P. L. (1991). Rural/urban versus regional differences in religiosity. *Review of Religious Research, 33*(1), 76–86.

Del Giudice, M. (2012). The twentieth century reversal of pink-blue gender coding: A scientific urban legend? *Archives of Sexual Behavior*, *41*(6), 1321-1323.

Duden Synonymwörterbuch (2014). Dudenredaktion. Synonymwörterbuch, D. D. *Ein Wörterbuch sinnverwandter Wörter*, 6.

Ellis, D. A., Wiseman, R., & Jenkins, R. (2015). Mental representations of weekdays. *PloS One*, *10*(8), e0134555.

FAZ. (n.d.). Verteilung der Weltbevölkerung nach Religionen in den Jahren 1900 und 2010. *In Statista – Das Statistik Portal*. Retrieved from https://de.statista.com/statistik/daten/studie/256878/umfrage/verteilung-der-weltbevoelkerung-nach-religionen/.

Finke, R., & McClure, J. M. (2015). Reviewing millions of books: Charting cultural and religious trends with Google's Ngram Viewer. *Association of Religion Data Archive's Working Paper Series*. Retrieved from http://thearda.com/rrh/papers/book.asp

Föllmer, M. (2010). Was Nazism collectivistic? Redefining the individual in Berlin, 1930-1945. *The Journal of Modern History, 82*(1), 61–100.

Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050-1057.

Froese, P. (2004). Forced secularization in Soviet Russia: Why an atheistic monopoly failed. *Journal for the Scientific Study of Religion*, *43*(1), 35-50.

Genovese, J. E. (2015). Interest in astrology and phrenology over two centuries: A Google Ngram study. *Psychological Reports*, *117*(3), 940-943.

Greenfield, P. M. (2009). Linking social change and developmental change: shifting pathways of human development. *Developmental Psychology*, *45*(2), 401-418.

Greenfield, P. M. (2013). The changing psychology of culture from 1800 through 2000. *Psychological Science*, *24*(9), 1722-1731.

Grossmann, I., & Varnum, M. E. (2015). Social structure, infectious diseases, disasters, secularism, and cultural change in America. *Psychological Science*, *26*(3), 311-324.

Gooding, P. (2012). Mass digitization and the garbage dump: The conflicting needs of quantitative and qualitative methods. *Li terary and Linguistic Computing*, *28*(3), 425-431.

Hamamura, T. (2012). Are cultures becoming individualistic? A cross-temporal comparison of individualism–collectivism in the United States and Japan. *Personality and Social Psychology Review*, *16*(1), 3-24.

Hamamura, T., & Xu, Y. (2015). Changes in Chinese culture as examined through changes in personal pronoun usage. *Journal of Cross-Cultural Psychology*, *46*(7), 930-941.

Hills, T. T., & Adelman, J. S. (2015). Recent evolution of learnability in American English from 1800 to 2000. *Cognition*, *143*, 87-92.

Ho, D. Y. F. (1979). 'Psychological implications of collectivism: with special reference to the Chinese case and Maoist dialectics'. In Eckensberger, L., Lonner, J. and Poortinga, Y. H. (Eds), *Cross-Cultural Contributions to Psychology*. Lisse, Netherlands: Swets & Zeitlinger, 143–150.

Hofstede, G. (2001). Culture's consequences: *Comparing values, behaviors, institutions and organizations across nations*. Sage.

Hughes, J. M., Foti, N. J., Krakauer, D. C., & Rockmore, D. N. (2012). Quantitative patterns of stylistic influence in the evolution of literature. *Proceedings of the National Academy of Sciences*, *109*(20), 7682-7686.

Huntington, S. P. (1996). *The clash of civilizations and the remaking of world order*. London, England:Simon & Schuster.

Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American Sociological Review*, *65*(1), 19-51.

Kesebir, P., & Kesebir, S. (2012). The cultural salience of moral character and virtue declined in twentieth century America. *The Journal of Positive Psychology*, *7*(6), 471-480.

Kesebir, S., & Kesebir, P. (2017). A growing disconnection from nature is evident in cultural products. *Perspectives on Psychological Science*, *12*(2), 258-269.

Koplenig, A. (2017). The impact of lacking metadata for the measurement of cultural and linguistic change using the Google Ngram data sets—Reconstructing the composition of the German corpus in times of WWII. *Digital Scholarship in the Humanities*, *32*(1), 169-188.

Kraus, M. W., Piff, P. K., Mendoza-Denton, R., Rheinschmidt, M. L., & Keltner, D. (2012). Social class, solipsism, and contextualism: How the rich are different from the poor. *Psychological Review, 119*(3), 546–572.

Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of Personality and Social Psychology*, *95*(1), 66.

Lamoreaux, M., & Morling, B. (2012). Outside the head and outside individualism-collectivism: Further meta-analyses of cultural products. *Journal of Cross-Cultural Psychology*, *43*(2) 299-327.

Li, Y., Tan, X., Huang, Z., & Liu, L. (2019). Relationship between collectivism and corruption in American and Chinese books: A historical perspective. *International Journal of Psychology*, *54*(2), 180-187.

Lin, Y., Michel, J. B., Aiden, E. L., Orwant, J., Brockman, W., & Petrov, S. (2012, July). Syntactic annotations for the Google Books Ngram corpus. In *Proceedings of the ACL 2012 system demonstrations* (pp. 169-174). Association for Computational Linguistics.

Mann, J., Zhang, D., Yang, L., Das, D., & Petrov, S. (2014). Enhanced search with wildcards and morphological inflections in the Google Books Ngram Viewer. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 115-120).

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., ... & Pinker, S. (2011). Quantitative analysis of culture using millions of digitized books [With supporting online material]. *Science*, *331*(6014), 176-182.

Mohammad, S. M. (2012). From once upon a time to happily ever after: Tracking emotions in mail and books. *Decision Support Systems*, *53*(4), 730-741.

Mooijman, M., Meindl, P., Oyserman, D., Monterosso, J., Dehghani, M., Doris, J. M., & Graham, J. (2018). Resisting temptation for the good of the group: Binding moral values and the moralization of self-control. *Journal of Personality and Social Psychology, 115*(3), 585-599.

Moore, R. L. (2005). Generation Ku: Individualism and China's millennial youth. *Ethnology*, *44*(4), 357-376.

Morin, O., & Acerbi, A. (2017). Birth of the cool: a two-centuries decline in emotional expression in Anglophone fiction. *Cognition and Emotion*, *31*(8), 1663-1675.

Morling, B., & Lamoreaux, M. (2008). Measuring culture outside the head: A meta-analysis of individualism-collectivism in cultural products. *Personality and Social Psychology Review*, *12*(3), 199-221.

Oishi, S., Graham, J., Kesebir, S., & Galinha, I. C. (2013). Concepts of happiness across time and cultures. *Personality and Social Psychology Bulletin*, *39*(5), 559-577.

Pargament, K. I., Ishler, K., Dubow, E. F., Stanik, P., Rouiller, R., Crowe, P., ... & Royster, B. J. (1994). Methods of religious coping with the Gulf War: Cross-sectional and longitudinal analyses. *Journal for the Scientific Study of Religion*, *33*(4) 347-361.

Pargament, K. I. (2001). *The psychology of religion and coping: Theory, research, practice*. Guilford Press.

Park, H., Twenge, J. M., & Greenfield, P. M. (2014). The Great Recession Implications for adolescent values and behavior. *Social Psychological and Personality Science*, *5*(3), 310-318.

Pechenick, E. A., Danforth, C. M., & Dodds, P. S. (2015). Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PloS One*, *10*(10), e0137041.

Pettit, M. (2016). Historical time in the age of big data: Cultural psychology, historical change, and the Google Books Ngram Viewer. *History of Psychology*, *19*(2), 141.

Potter, S. H. (1988). The cultural construction of emotion in rural Chinese social life. *Ethos*, 16(2), 181-208.

Reips, U.-D. (2008). How Internet-mediated research changes science. In A. Barak (Ed.), *Psychological aspects of cyberspace: Theory, research, applications* (pp. 268-294). Cambridge: Cambridge University Press.

Ritter, R. S., & Preston, J. L. (2013). Representations of religious words: Insights for religious priming research. *Journal for the Scientific Study of Religion*, *52*(3), 494-507.

Roivainen, E. (2013). Frequency of the use of English personality adjectives: Implications for personality theory. *Journal of Research in Personality*, *47*(4), 417-420.

Roivainen, E. (2014). Changes in word usage frequency may hamper intergenerational comparisons of vocabulary skills: An Ngram analysis of wordsum, WAIS, and WISC test items. *Journal of Psychoeducational Assessment*, *32*(1), 83-87.

Roivainen, E. (2015). Personality adjectives in twitter tweets and in the google books corpus. An analysis of the facet structure of the openness factor of personality. *Current Psychology*, *34*(4), 621-625.

Rossi, E., Mortimer, J., & Rossi, K. (2013). Therapeutic hypnosis, psychotherapy, and the digital humanities: The narratives and culturomics of hypnosis, 1800–2008. *American Journal of Clinical Hypnosis*, *55*(4), 343-359.

Scheff, T. (2015). Toward defining basic emotions. *Qualitative Inquiry*, *21*(2), 111-121.

Schimmack, U., Oishi, S., & Diener, E. (2002). Cultural influences on the relation between pleasant emotions and unpleasant emotions: Asian dialectic philosophies or individualism-collectivism? *Cognition & Emotion*, *16*(6), 705-719.

Silbereisen, R. K., & Chen, X. (2010). *Social change and human development: Concepts and results*. London, U.K.: Sage.

Skrebyte, A., Garnett, P., & Kendal, J. R. (2016). Temporal relationships between individualism–collectivism and the economy in Soviet Russia: A word frequency analysis using the Google Ngram corpus. *Journal of Cross-Cultural Psychology*, *47*(9), 1217-1235.

Steele, L. G., & Lynch, S. M. (2013). The pursuit of happiness in China: Individualism, collectivism, and subjective well-being during China's economic and social transformation. *Social Indicators Research*, *114*(2), 441-451.

Strupp, C. (30 January 2013). Only a phase: How diplomats misjudged Hitler's rise. Spiegel Online. Retrieved from http://www.spiegel.de/international/germany/marking-eighty-years-since-hitler-took-power-in-germany-a-880565.html.

Sun, Z., & Wang, X. (2010). Value differences between generations in China:Astudy in Shanghai. *Journal of Youth Studies*, *13*(1), 65–81.

Tönnies, F. (1957). *Community and Society: Gemeinschaft und Gesellschaft*, translated and edited by Charles P. Loomis; East Lansing, Mich. (Original work published 1887).

Triandis, H. C. (1993). Collectivism and individualism as cultural syndromes. *Cross-Cultural Research*, 27(3-4), 155-180.

Triandis, H. C. (2001). Individualism-collectivism and personality. *Journal of Personality*, 69(6), 907-924.

Triandis, H. C. (2018). *Individualism and Collectivism*. Routledge.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012a). Increases in individualistic words and phrases in American books, 1960–2008. *PloS One*, *7*(7), e40181.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2012b). Male and female pronoun use in US books reflects women's status, 1900–2008. *Sex Roles*, *67*(9-10), 488-493.

Twenge, J. M., Campbell, W. K., & Gentile, B. (2013). Changes in pronoun use in American books and the rise of individualism, 1960-2008. *Journal of Cross-Cultural Psychology*, *44*(3), 406-415.

Twenge, J. M., VanLandingham, H., & Campbell, W. K. (2017). The seven words you can never say on television: Increases in the use of swear words in American books, 1950-2008. *SAGE Open*, *7*(3), 2158244017723689.

Uhls, Y. T., & Greenfield, P. M. (2011). The rise of fame: A historical content analysis. *Cyberpsychology*: *Journal of Psychosocial Research on Cyberspace*, 5(1), 1.

Uz, I. (2014). Individualism and first person pronoun use in written texts across languages. *Journal of Cross-Cultural Psychology*, *45*(10), 1671-1678.

Velichkovsky, B. B., Solovyev, V. D., Bochkarev, V. V., & Ishkineeva, F. F. (2019). Transition to market economy promotes individualistic values: Analysing changes in frequencies of Russian words from 1980 to 2008. *International Journal of Psychology*, *54*(1), 23-32.

Virues-Ortega, J., & Pear, J. J. (2015). A history of "behavior" and "mind": Use of behavioral and cognitive terms in the 20th century. *The Psychological Record*, *65*(1), 23-30.

Welch, D. (2004). Nazi propaganda and the Volksgemeinschaft: Constructing a people's community. *Journal of Contemporary History, 39*(2), 213-238.

Wilson, R. W. (1974). *The moral state: A study of the political socialization of Chinese and American children*. New York: Free Press.

Xu, Y., & Hamamura, T. (2014). Folk beliefs of cultural changes in China. *Frontiers in Psychology*, *5*, 1066.

Yan, Y. (2009). *The individualization of Chinese society*. New York: Oxford.

Ye, S., Cai, S., Chen, C., Wan, Q., & Qian, X. (2018). How have males and females been described over the past two centuries? An analysis of Big-Five personality-related adjectives in the Google English Books. *Journal of Research in Personality*, *76*, 6-16.

Younes, N., & Reips, U. -D. (2018). The changing psychology of culture in German-speaking countries: A Google Ngram study. *International Journal of Psychology*, *53*, 53-62.

Younes, N., & Reips, U. -D. (2019). Guideline for improving the reliability of Google Ngram studies: Evidence from religious terms. *PloS One*, *14*(3): e0213554. https://doi.org/10.1371/journal.pone. 0213554.

Yu, F., Peng, T., Peng, K., Tang, S., Chen, C. S., Qian, X., ... & Chai, F. (2016). Cultural value shifting in pronoun use. *Journal of Cross-Cultural Psychology*, *47*(2), 310-316.

Zeng, R., & Greenfield, P. M. (2015). Cultural evolution over the last 40 years in China: Using the Google Ngram Viewer to study implications of social and political change for cultural values. *International Journal of Psychology*, *50*(1), 47-55.

吴景荣. (2000). *Xin shi dai Han-Ying da ci dian: New Age Chinese-English Dictionary.* Jingrong, W. (Ed.). Bei jing: Shang wu yin shu guan.

寄洲, & 永芬. (2005). *1700 对近义词语用法对比: 1700 Groups of Frequently Used Chinese Synonyms*. 北京语言文化大学出版社.

S1 Appendix

Due to large scaling differences, the Figures A to N of S1 Appendix do not include the words “God”, “Gott”, “Dio”, “Dios”, and “Dieu”, respectively. However, they show a similar curvature.

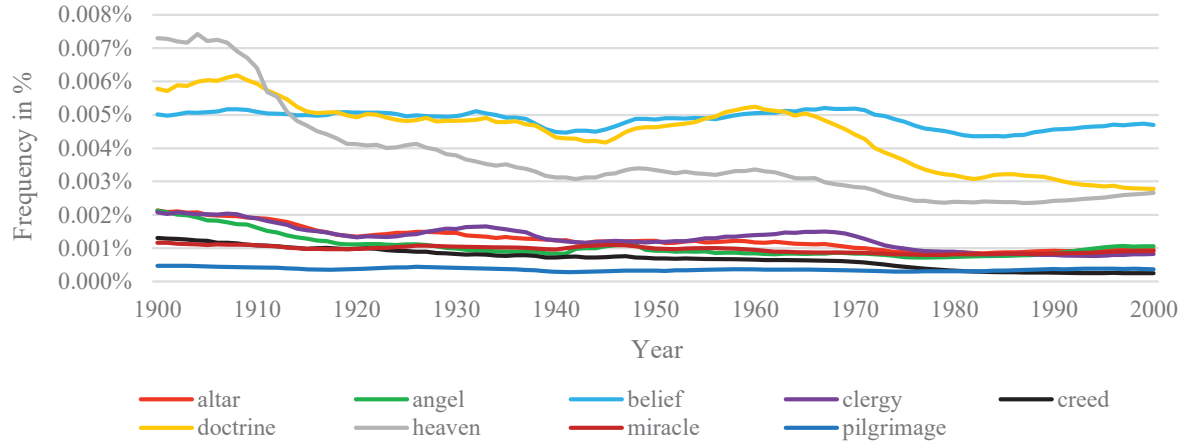

Fig A. Raw frequencies for religious terms using the American English Google Ngram corpus.

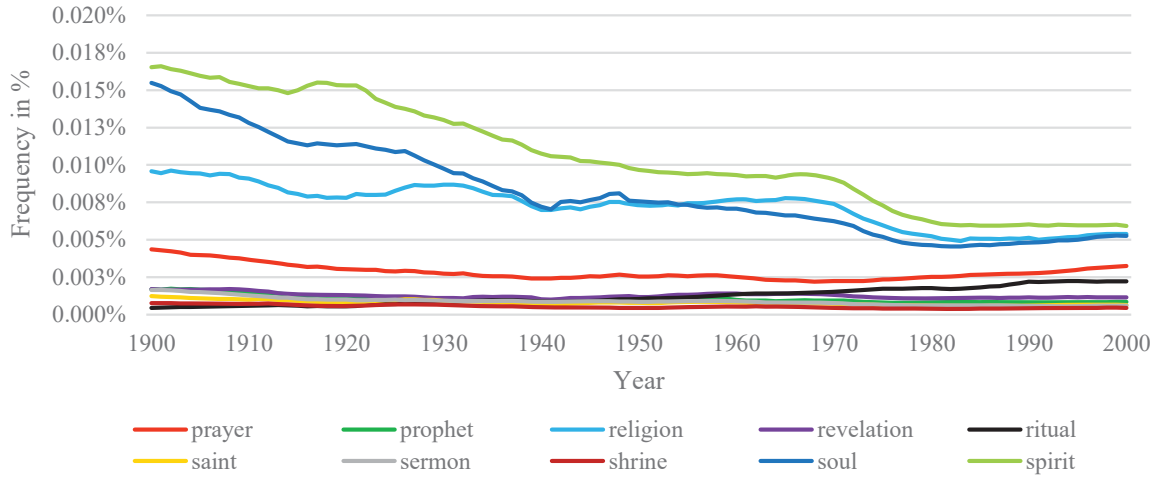

Fig B. Raw frequencies for religious terms using the American English Google Ngram corpus.

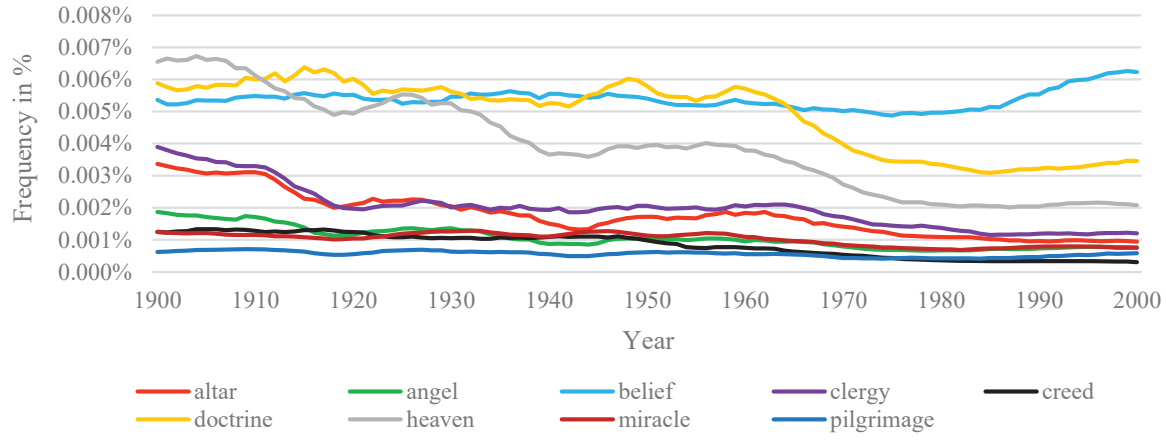

Fig C. Raw frequencies for religious terms using the British English Google Ngram corpus.

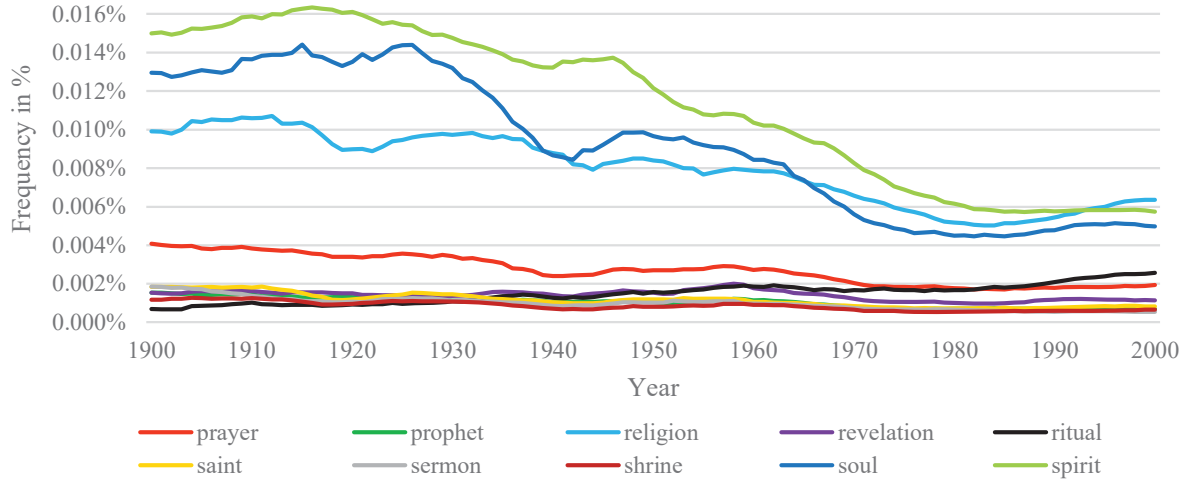

Fig D. Raw frequencies for religious terms using the British English Google Ngram corpus.

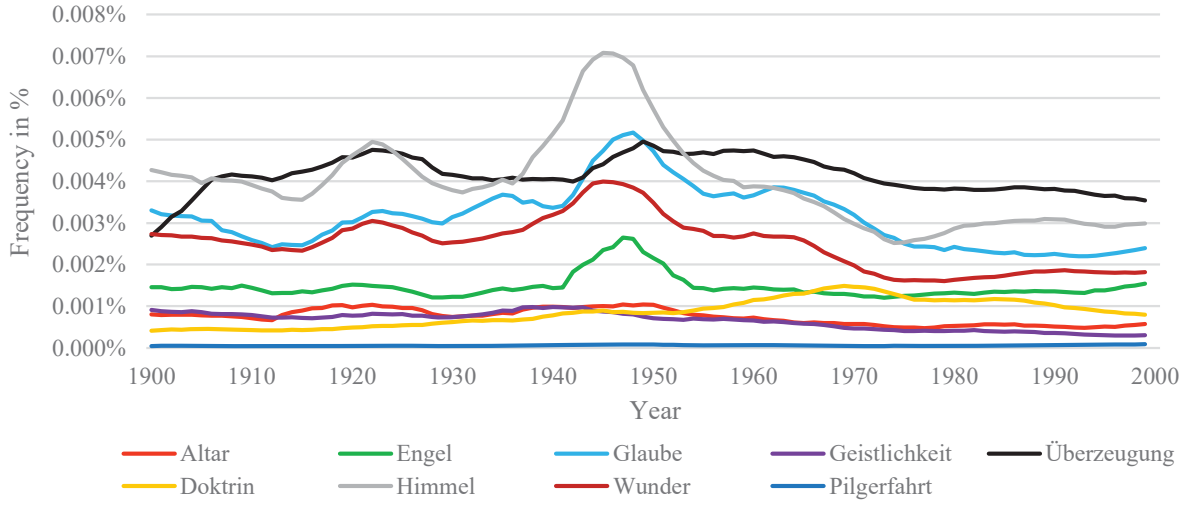

Fig E. Raw frequencies for religious terms using the German Google Ngram corpus.

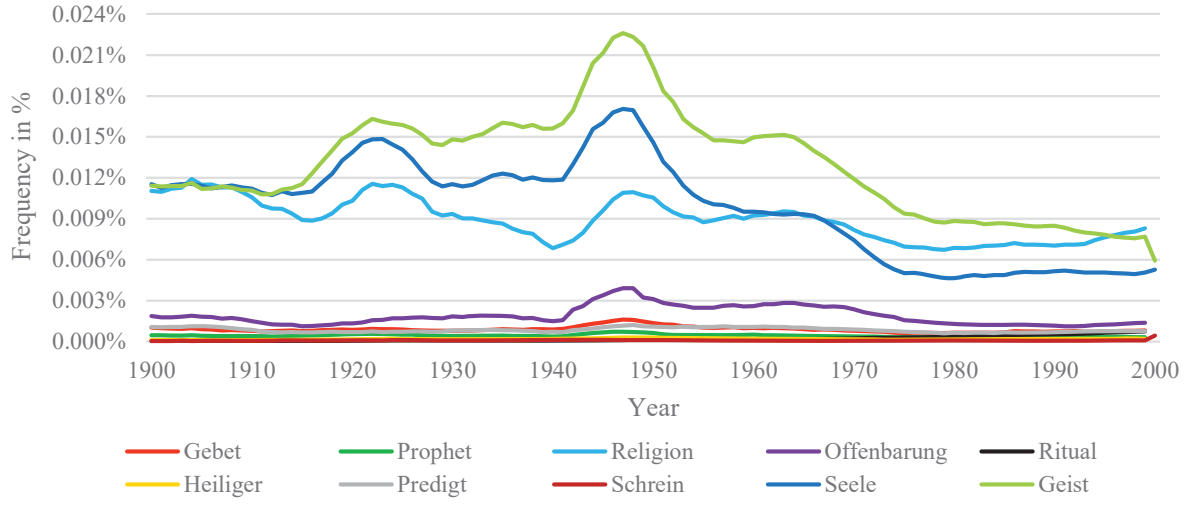

Fig F. Raw frequencies for religious terms using the German Google Ngram corpus.

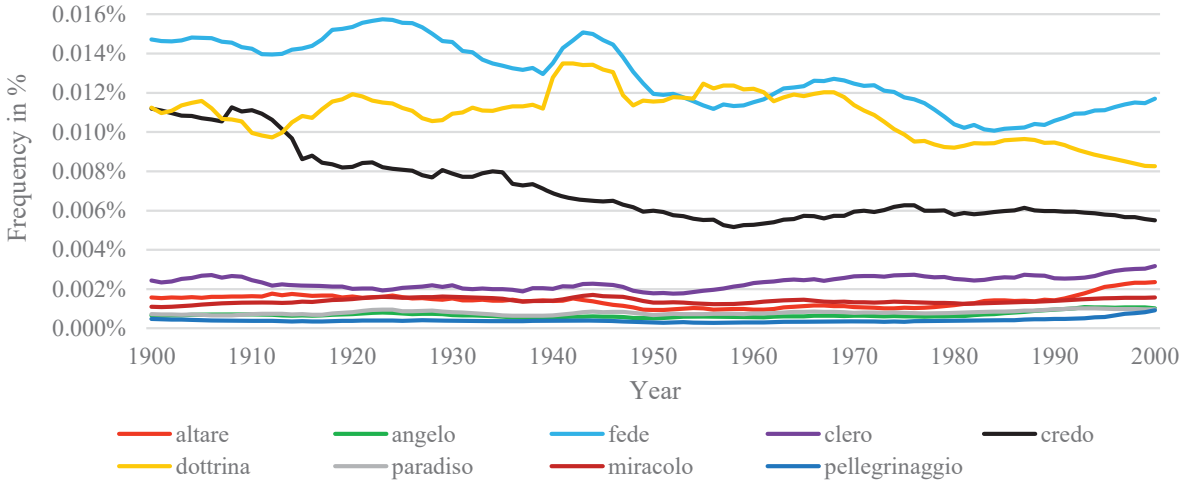

Fig G. Raw frequencies for religious terms using the Italian Google Ngram corpus.

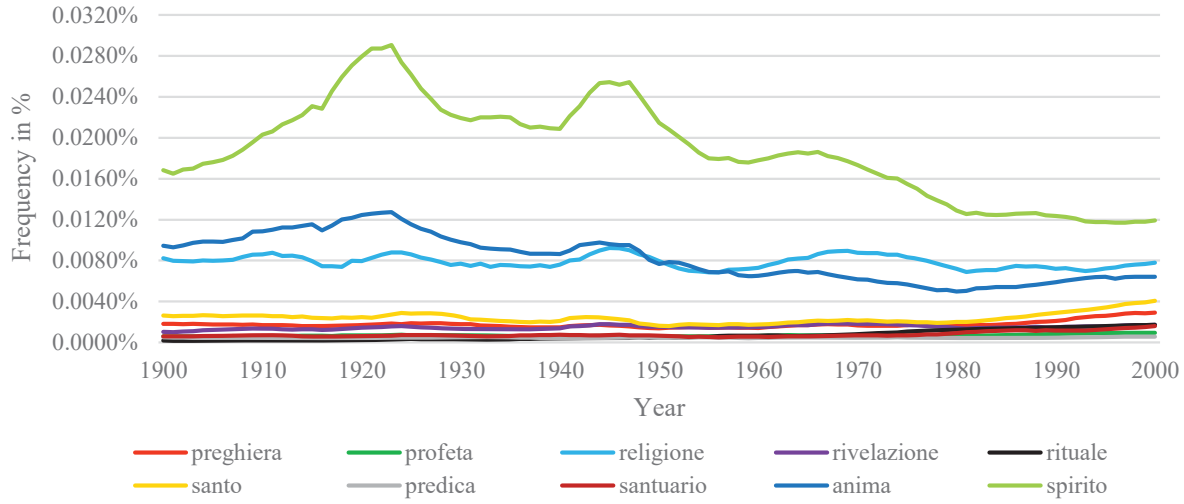

Fig H. Raw frequencies for religious terms using the Italian Google Ngram corpus.

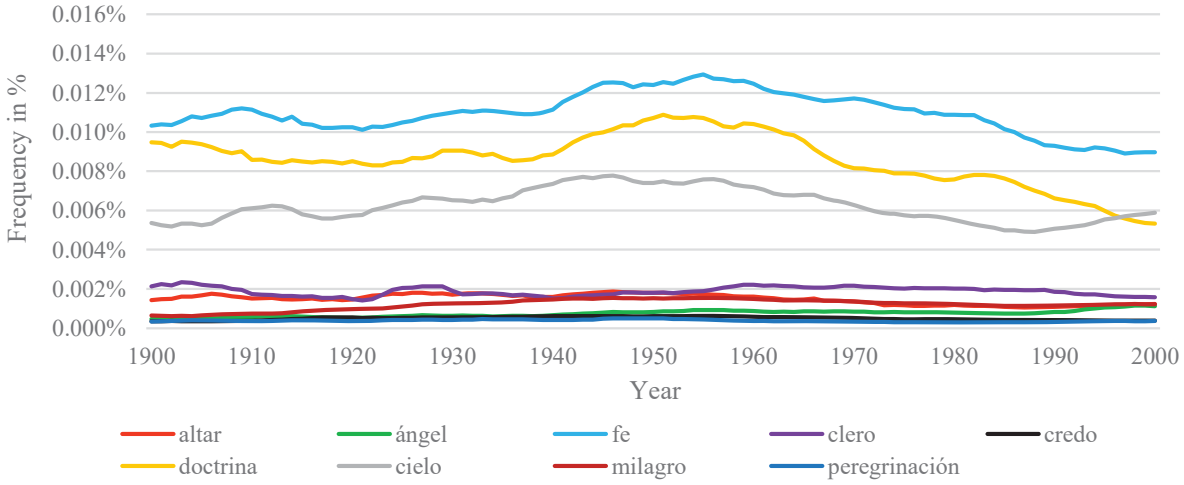

Fig I. Raw frequencies for religious terms using the Spanish Google Ngram corpus.

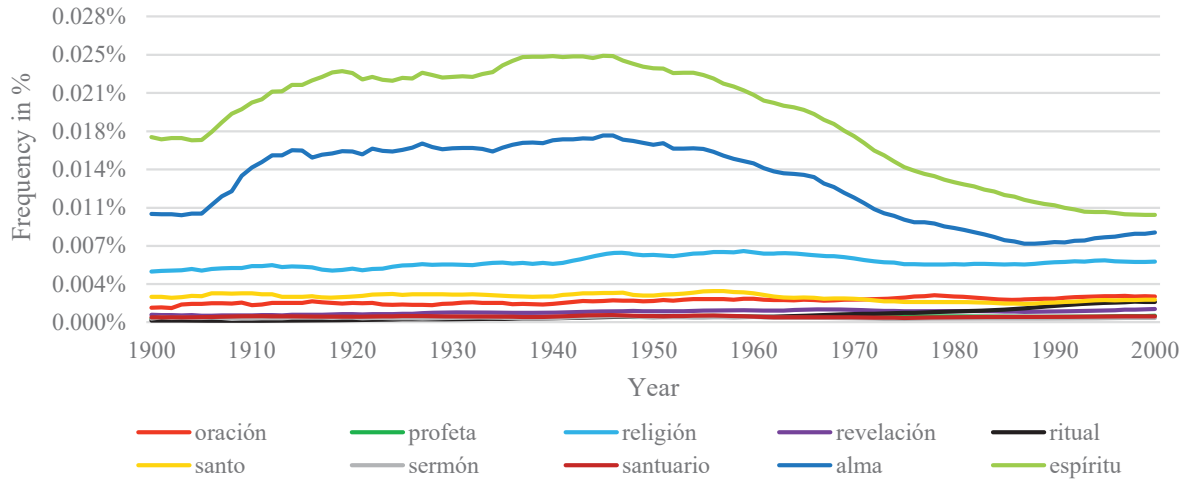

Fig J. Raw frequencies for religious terms using the Spanish Google Ngram corpus.

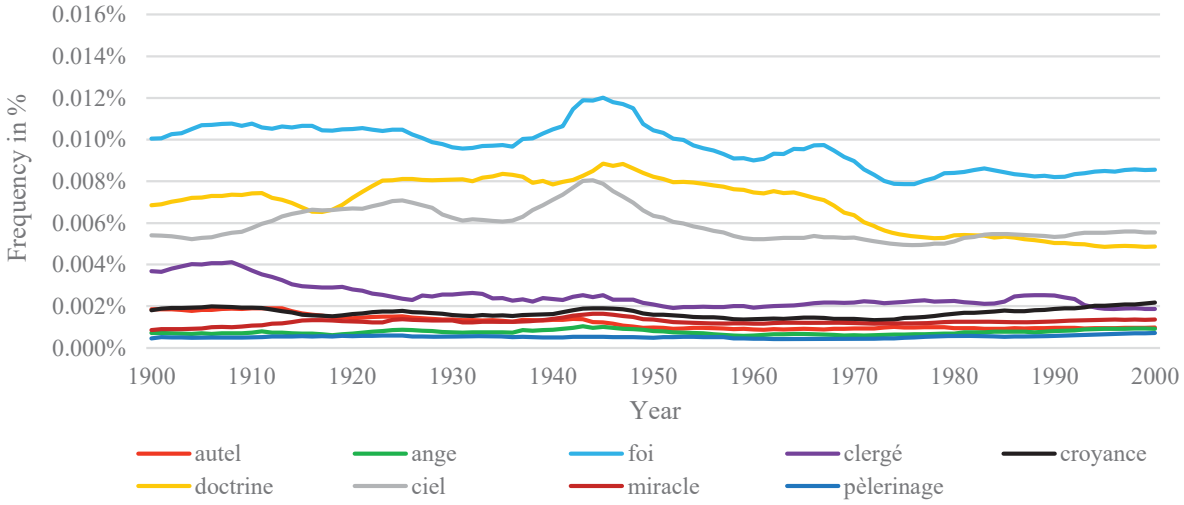

Fig K. Raw frequencies for religious terms using the French Google Ngram corpus.

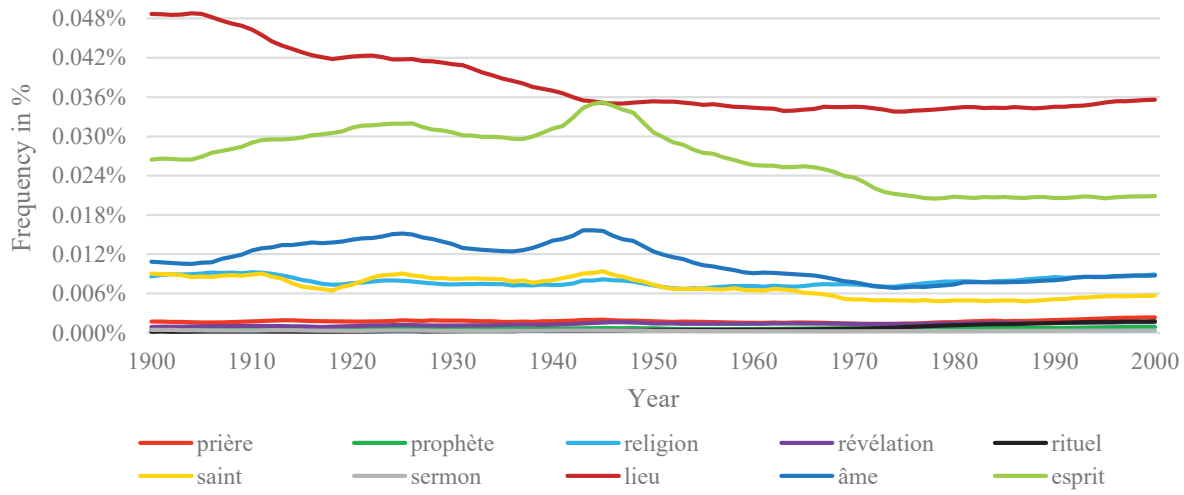

Fig L. Raw frequencies for religious terms using the French Google Ngram corpus.

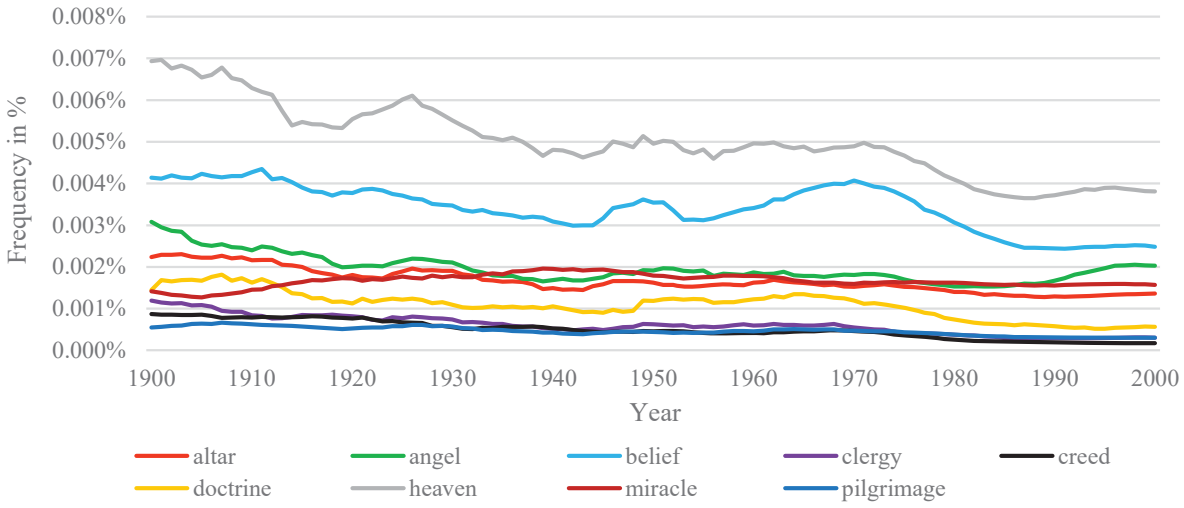

Fig M. Raw frequencies for religious terms using the Fiction English Google Ngram corpus.

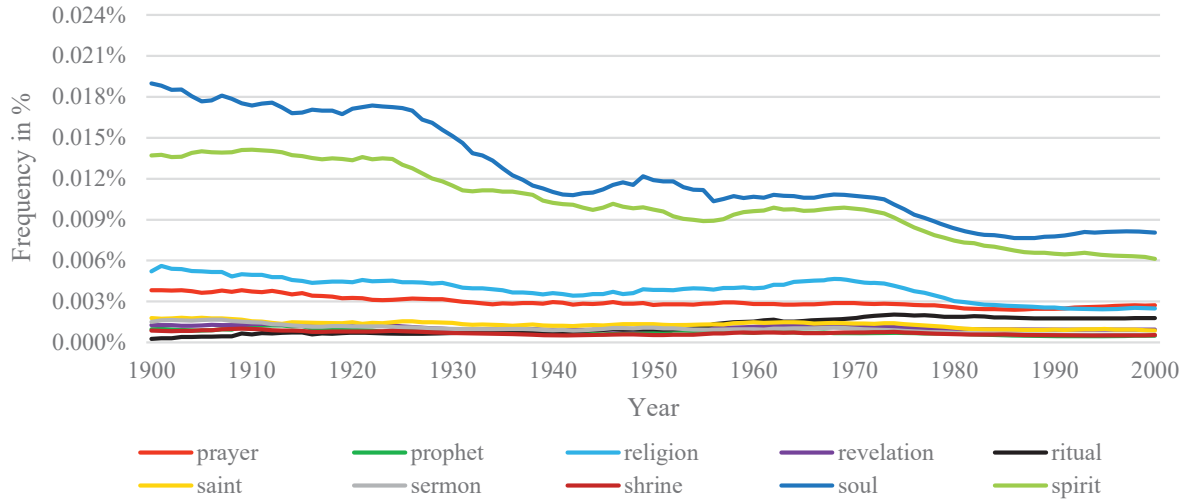

Fig N. Raw frequencies for religious terms using the Fiction English Google Ngram corpus.

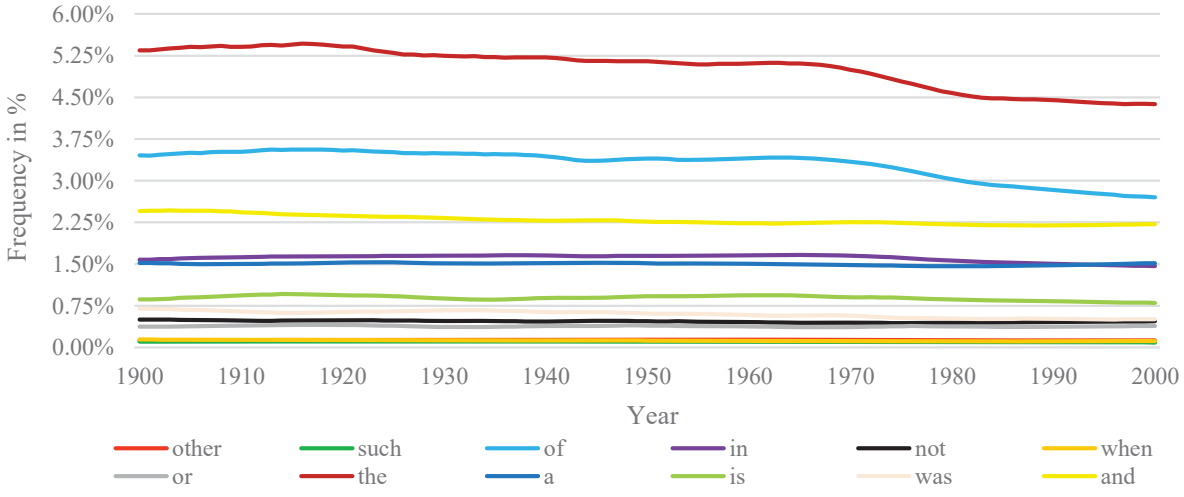

Fig O. Raw frequencies for common words using the American English Google Ngram corpus.

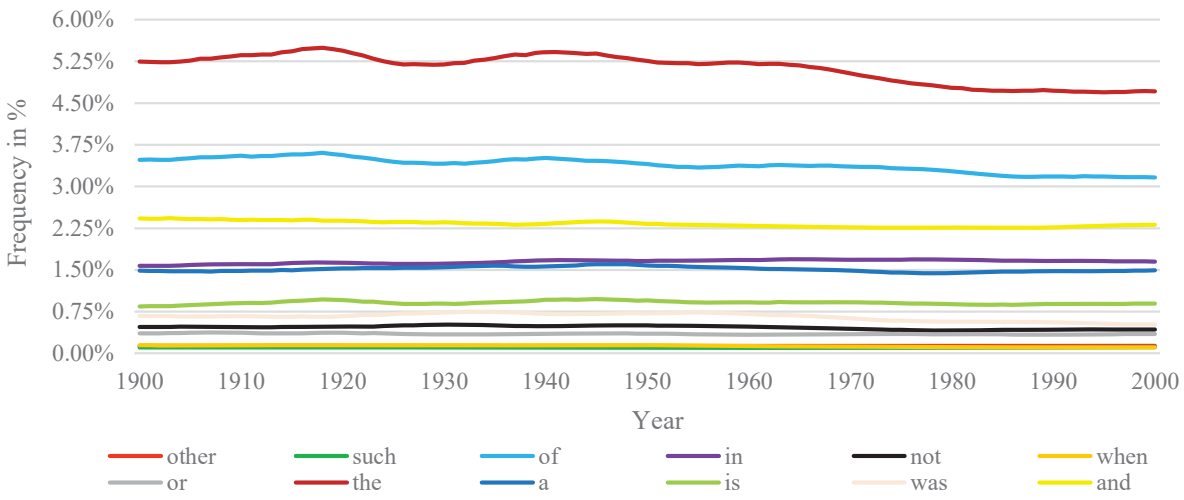

Fig P. Raw frequencies for common words using the British English Google Ngram corpus.

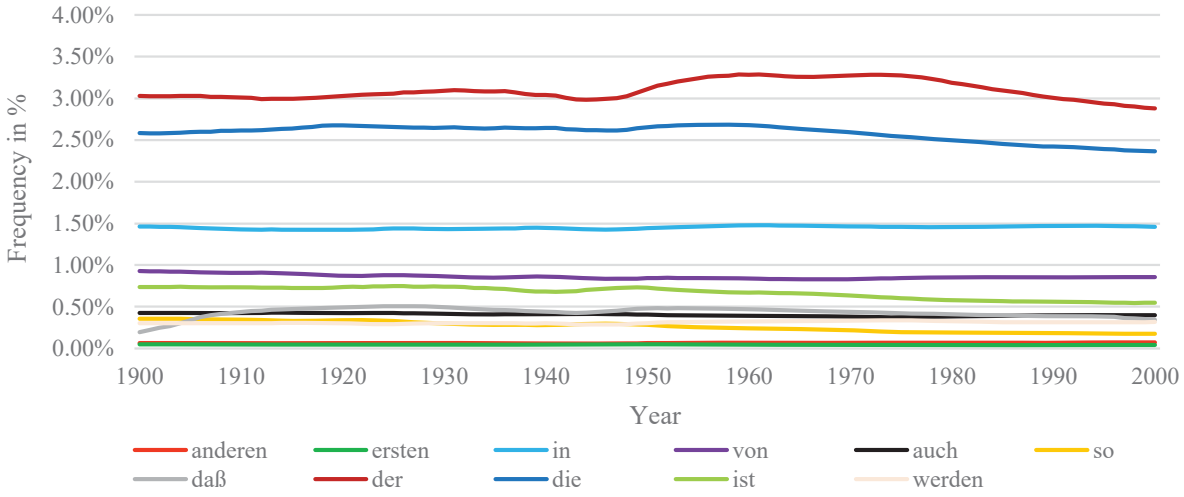

Fig Q. Raw frequencies for common words using the German Google Ngram corpus.

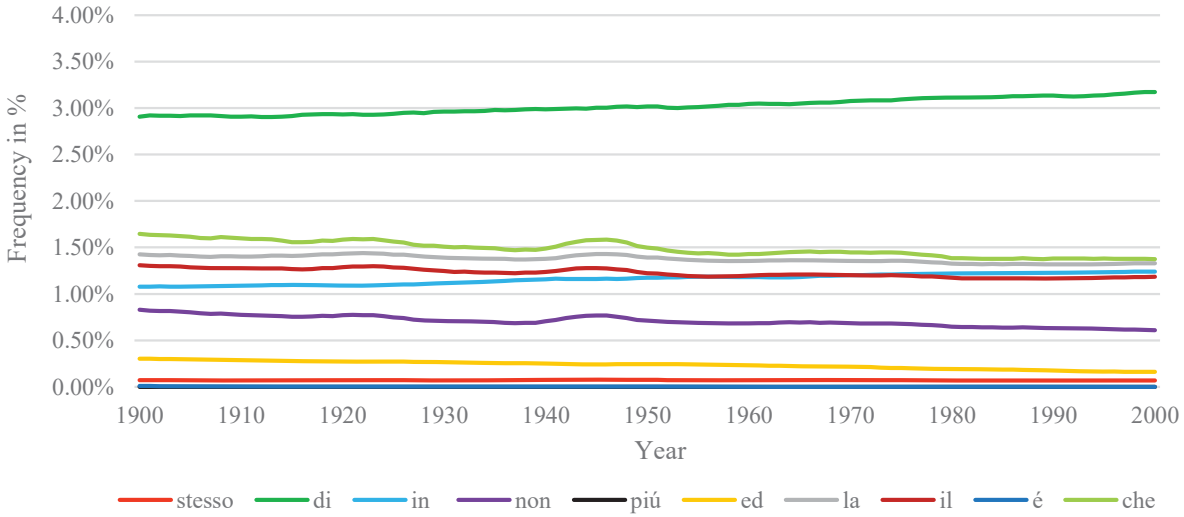

Fig R. Raw frequencies for common words using the Italian Google Ngram corpus.

Table A. Psychological Google Ngram research published between 2010 and 2018.

Journal	Article	Language	Subject
American Journal of Clinical Hypnosis	[46]	English	Other
Archives of Sexual Behavior	[17]	English	Gender
Cognition	[26]	English	Cognition
Cognition and Emotion	[22]	English	Emotions
Current Psychology	[24]	English	Personality
Decision Support Systems	[21]	English	Emotions
Frontiers in Psychology	[12]	Chinese	I/C
International Journal of Psychology	[13]	Chinese	I/C
	[14]	German	I/C
	[16]	Russian	I/C
	[51]	Various	I/C
Journal of Cross-Cultural Psychology	[7]	English	I/C
	[11]	Chinese	I/C
	[15]	Russian	I/C
	[37,38]	Various	I/C
Journal of Personality and Social Psychology	[47]	English	Other
Journal of Positive Psychology	[8]	English	I/C
Journal of Psychoeducational Assessment	[48]	English	Other
Journal of Research in Personality	[19]	English	Gender
	[23]	English	Personality
Personality and Social Psychology Bulletin	[36]	English	Emotions
Perspectives of Psychological Science	[49]	English	Other
PLOS ONE	[6]	English	I/C
	[20]	English	Emotions
	[25]	English	Cognition
Psychological Reports	[50]	English	Other
Psychological Science	[10,52]	English	I/C
Qualitative Inquiry	[45]	Various	Emotions
SAGE Open	[9]	English	I/C
Science	[4]	Various	Other
Sex Roles	[18]	English	Gender
The Psychological Record	[27]	English	Cognition

Note. I/C indicates "Individualism/Collectivism". Literature was collected by conducting a Google Scholar search. The first 50 pages covering the years 2010 to 2018 were taken into account. The search terms were "google ngram books". Only psychological literature was considered.

Table B. Overview of religious terms in Spanish and French.

Original	Spanish	French	Original	Spanish	French
altar	altar	autel	prayer	oración	prière
angel	ángel	ange	prophet	profeta	prophète
belief	fe	foi	religion	religión	religion
clergy	clero	clergé	revelation	revelación	révélation
creed	credo	croyance	ritual	ritual	rituel
doctrine	doctrina	doctrine	saint	santo	saint
God	Dios	Dieu	sermon	sermón	sermon
heaven	cielo	ciel	shrine	santuario	lieu
miracle	milagro	miracle	soul	alma	âme
pilgrimage	peregrinación	pèlerinage	spirit	espíritu	esprit

Table C. Overview of synonyms.

Original	Synonyms English			Synonyms German			Synonyms Italian		
altar	sacrarium	sanctuary	prothesis	Opferstätte	Brandopferstätte	X	ara	X	X
angel	archangel	seraph	cherub	Cherub	Himmelsbote	Himmelswächter	X	X	X
belief	deism	theism	animism	Frömmigkeit	Glaubensüberzeugung	Gläubigkeit	religione	professione	credo
clergy	pastor	shepherd	vicar	X	X	X	ecclesiastici	sacerdoti	religiosi
creed	credo	dogma	orthodoxy	Anschauung	Ansicht	Betrachtungsweise	dogma	fede	X
doctrine	theology	divinity	scholasticism	Dogma	Glaubenssatz	Lehre	principi	teorici	dogma
God	Creator	Preserver	Allah	Allwissender	Gottvater	Herr	Iddio	Domineddio	Creatore
heaven	paradise	nirvana	zion	Jenseits	Paradies	Himmelreich	eden	cielo	empireo
miracle	prodigy	portent	sign	Ausnahmeerscheinung	Geheimnis	Hexenwerk	portento	prodigio	X
pilgrimage	rites	mysteries	laud	Wallfahrt	Hadsch	X	viaggio	X	X
prayer	orison	devotion	retreat	Bitte	Bittgebet	Dankgebet	orazione	prece	invocazione
prophet	priest	archpriest	hierophant	Hellseher	Mahner	Rufer	annunziatore	vate	X
religion	creed	dogma	piety	Bekenntnis	Glaube	Glaubenslehre	fede	credo	confessione
revelation	divine	apocalypse	disclosure	X	X	X	teofania	X	X
ritual	rites	cult	institution	Kult	Ritus	Brauch	cerimoniale	cerimonia	X
saint	patron	X	X	Legende	Mythos	Patron	martire	patrono	X
sermon	oration	speech	effusion	Kanzelrede	Kanzelwort	Sermon	sermone	omelia	predicazione
shrine	temple	fane	pantheon	Heiligenschrein	Reliquiar	Reliquienschrein	tempio	chiesa	sacrario
soul	spirit	geist	mind	Empfindungsleben	Gefühlsleben	Gemüt	spirito	X	X
spirit	ghost	shades	visitant	Dämon	Erscheinung	Gespenst	anima	X	X

Note. X indicates that no synonym was available. In case singular and plural forms were listed we only considered one form. For the German word "Gott" we did not conisder the pronoun "Er" (he) as a synonym. For the English word "spirit" we did not consider the words "Manes", "lemures", and "zombie" as synonyms.

Table D. Overview of original words and their higher frequency inflections.

American English		British English		German		Italian	
Original	High	Original	High	Original	High	Original	High
orison	orisons	orison	orisons	Cherub	Cherubim	principi	principio
seraph	seraphim	mysteries	mystery	Allwissender	Allwissenden	martire	martiri
mysteries	mystery	shades	shade	Dämon	Dämonen	prece	preci
shades	shade	institution	institutions	Glaubenssatz	Glaubenssätze	omelia	omelie
institution	institutions			Glaube	Glauben	sacerdoti	sacerdote

Note. As noted in the main text, (higher frequency) inflections may change over time.

Table E. Correlation coefficients for WWII using words of highest frequency.

Language	Original Terms	Synonyms	Original Terms + Synonyms
American English	r=-0.88, p<0.001	r=-0.89, p<0.001	r=-0.88, p<0.001
British English	r=-0.95, p<0.001	r=-0.93, p<0.001	r=-0.93, p<0.001
German	r= 0.88, p<0.001	r= 0.81, p<0.001	r= 0.85, p<0.001
Italian	r= 0.41, p>0.1	r= 0.87, p<0.001	r= 0.79, p<0.01

Note. Correlation coefficients are calculated by subtracting the summed z-scores of raw frequencies of various very common words from the summed z-scores of raw frequencies of the original terms and/or synonyms.