Oneline Supporting Information for:

Literary Fiction Indicates Early Modernization in China Prior to Western Influence

Authors: Ying Zhong[1]*, Valentin Thouzeau[1], Nicolas Baumard[1]*

Affiliations:

[1] Institut Jean Nicod, Département d'études cognitives, École normale supérieure, Université PSL, EHESS, CNRS; Paris, 75005, France.

This file includes:

Supporting text

Figures S1 to S39

Tables S1 to S16

SI References

Supplementary Text

Study 1: Evolution of Chinese fiction in history

Data Collection

Inclusion criteria. We have adopted a broad definition of fiction to include Tang Dynasty's short stories (*chuanqi*), Song Dynasty's tales (*huaben*), Yuan Dynasty's plays, novels from the Ming and Qing Dynasties onwards, movies from the 20th century and beyond, and fictional TV dramas.

Exclusion criteria. We do not collect samples for:

- **Non-fiction**: historical narrative texts (such as *The Records of the Three Kingdoms* 三国志 written by Chen Shou)*,* note prose, essays, or narrative poems (such as *Chang hen ge* 长恨歌).

- **Fiction with only oral tradition**: opera (such as peking opera, *kunqu…*) texts that have been lost; folklore and their adaptations because their dating is difficult (for example: *Seven Fairies* are adapted into many narrative stories, but they lack traceable written texts). Exceptionally, We collected folklore fiction that are summarized by *A bibliography of vernacular novels* (Jiangsu Academy of Social Sciences 1990) which is based on written texts, such as *The Legend of the Cowherd and the Weaver Girl* （牛郎织女传 1910 version).

To collect the most comprehensive and systematic meta-data available (summary of plot, or descriptions of the fiction), one issue is that information on Chinese fiction is scattered across various platforms. For instance, while Wikipedia hosts an extensive repository of data on Chinese modern fiction, facilitating automated extraction, its utilization is notably less prevalent within mainland China. Baidu Baike is more commonly used in mainland China, but it lacks indexing and an open API for retrieving information. Besides, there are fictions that haven't been recorded on both Wikipedia and Baidu Baike, yet they have been documented in bibliographies with summaries written by scholars.

Sampling

We employed both manual and automated methods to gather metadata from three primary sources: Baidu Baike, Wikipedia, and a scholarly bibliography of vernacular novels compiled by the Jiangsu Provincial Academy of Social Sciences.

For the manual collection, due to the absence of query indexing on Baidu Baike, we relied on literary history books and Chinese online commentary sites as references. From these sources,

we identified works of fiction, then manually retrieved summaries from Wikipedia and Baidu Baike, where available. This process is illustrated in Figure S1, which details how we extracted names of literary works from historical books and then accessed the corresponding Baidu Baike and Wikipedia pages to collect detailed information about these novels. This manual approach covered a comprehensive set of textual fiction from Tang dynasty (600 CE) up to modern (2022 CE).
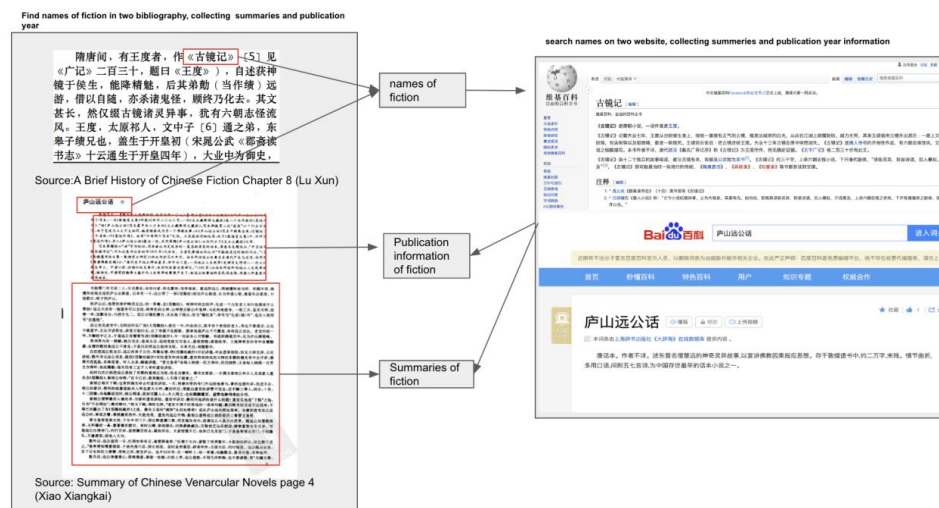
Fig. S1. Process of data collection. The left side shows two fiction examples in literary historical books. One is 古镜记 (*Biography of Old Mirror*) in chapter eight of *A Brief History of Chinese Fiction* (Lu 2000). Based on the information and name of this fiction, we were able to trace the page of this fiction on Wikipedia (see right figure). Then, we collected the summary from Wikipedia. Another example is 庐山远公话 *Lu Shan Yuan Gong Hua* in *A bibliography of vernacular novels* (Jiangsu Academy of Social Sciences 1990), there are name of fiction, publication info and summaries of fiction. We went to Baidu and Wikipedia to collect summaries based on the info. Summaries are also collected from the book written by scholars (see left figure).

For modern fiction, including films, TV shows, and other written works post-1912, we adopted an automated approach due to the massive scale of fiction published annually. We utilized Wikidata's query services, a multilingual knowledge graph managed by the Wikimedia Foundation, to conduct extensive searches by country (China) and category (e.g., novels, movies).

Data sources for literary fiction: manual collection

We relied on the following history books for extracting names of written fiction. Based on the names provided by the following sources, we collected summaries from Baidu Baike and Wikipedia.

- Lu Xun's book *A brief history of Chinese fiction* (Lu 2000) , spanning from 400s to 1912s

Lu Xun's bibliography is organized by dynasty and genre, which gives names and version information about the most important and famous literary fiction in history before the Republic of China. We started the collection after chapter 8 (Tang *chuanqi* as the beginning). One reason is that Tang *chuanqi* is widely accepted as the starting point of the form of fiction in history. Another reason is there were not enough stories before. One of the examples in chapter 8 is the earliest Tang fiction mentioned by Lu Xun - *The tales of mirror* (古镜记). It was created around 612 CE, adapted from the Jin dynasty 400 CE *Anecdotes About Spirits and Immortals* (搜神记). Because these two are classified as fictional stories in both literature and online information, we collected them into the database.

Accordingly, we followed the same process below: for each name mentioned by Lu Xun, we manually check whether this work is fictional and has a narrative structure. If it is, we then go to the Wikipedia and Baidu Baike pages to manually extract the description of the work, the summaries of the plot and publication year. We stopped the collection at chapter 28, the end of this book.

- *A bibliography of vernacular novels* (Jiangsu Academy of Social Sciences 1990), spanning from 972s to 1912s.

The book *A bibliography of vernacular novels* contains vernacular novels from Song to the Qing Dynasty. It does not include fiction written in classical Chinese (*wen yan* 文言 novel), such as Tang *chuanqi*. Each section of the book includes the title of fiction, author name, edition information of each fiction, summary of story written by groups of scholars, and chapter table

of contents. For each fiction mentioned in this book, summaries are collected from Baidu Baike and Wikipedia (if available) based on name index.

- *Thirty Years of Modern Chinese Literature* 中国现代文学三十年, modern novels (Qian, Wen, and Wu 1998), spanning from 1917 to 1949.

The novel section of the book "*Thirty Years of Modern Chinese Literature*" reviews the main literary trends from 1917 to 1949, covering wide literary samples during the Republic of China, May Fourth Movement novels, autobiographical lyrical novels, the *Mandarin Duck and Butterfly School* represented by Zhang Henshui, and the history of martial arts novels. Chapter 3, 4, 14, 15, 23 and 24 were selected, with *Appendix Chronology* at the end of each chapter, which gives the name and year of the novel published. We referred to *A History of Modern Chinese Fiction* (*中国当代小说史*) written by Hsia (Hsia 1999) to confirm the unbiased selection by the author. Then based on the names mentioned in the book's index, we manually assess whether they are fictional novels. If so, we proceed to Baidu Baike and Wikipedia.

- *A History of Contemporary Chinese Literature* (Hong 2007) modern novels, spanning from 1949 - 2000.

We used *A History of Contemporary Chinese Literature (中国当代文学史)*, extracting the *Appendix Chronology (年表)* after page 395, which gives literary publication information, for example, "Sun Li published his novel *Zhu Fu 嘱咐* in March 1949". We collected the data following the same process before based on names of fictions provided by this book. Non-fiction is removed during sampling.

To sample popular literary fiction that are not collected in literary history such as 21st century online literature, we rely on popular website because there is not yet a systematic history review on fiction published in 21st century China:

- Douban website and DangDang website (for popular modern novels)

Douban is China's largest book review website, which, based on user data, annually provides a list of the 250 most popular books[1]. Typically, this list sees little fluctuation within the year, reflecting some bestseller indicators (for example, *Dream of the Red Chamber, To Live* and *The Three-Body Problem* are perennial bestsellers). We also referenced Dangdang Wang, a website that is China's equivalent to Amazon and is the most used site by consumers. It also offers a list of best-selling books, based on the rankings from 2018, 2020, and 2021 (2019 is currently unavailable)[2]. From these, we were able to obtain names of best-selling books. Using these four

---

[1] Top Douban books: https://book.douban.com/top250

[2] http://bang.dangdang.com/books/bestsellers/01.03.00.00.00.00-year-2021-0-1-1

lists from the two websites, we collected names of contemporary fictional novels and then repeated the previous process of collecting summaries and publication years from Baidu Baike and Wikipedia.

We have collected not only summaries from Baidu Baike and Wikipedia but also summaries written by scholars, especially those of ancient literary fiction from:

- *A bibliography of vernacular novels* (Jiangsu Academy of Social Sciences 1990), spanning from 972s to 1912s.

This book offers summaries written by groups of scholars from Jiangsu Provincial Academy of Social Sciences. These summaries objectively narrate the story without incorporating literary criticism. We use OCR (Optical Character Recognition) to convert pdf to text and then integrate the texts into the csv. When dealing with older PDF versions, OCR may inaccurately interpret individual words, leading to some errors or typos in the text.

Table S1 summarizes the sources of literary fiction and the number of summaries we extract based on these sources.

Table S1. Sources of fiction name and sources of summaries for manual collection

Source of names of fiction	Source of summaries	Info extracted from source of summaries	Type	Number of summaries
A brief history of Chinese fiction (Lu 2000)	Baidu Baike, Wikipedia	Description, publication year, type of fiction	Ancient literary	579
Bibliography of vernacular novels (Jiangsu Academy of Social Sciences 1990)		Description, publication year, type of fiction	Ancient literary	
Bibliography of vernacular novels (Jiangsu Academy of Social Sciences 1990)	*Bibliography of vernacular novels* (Jiangsu Academy of Social Sciences 1990)	Description, publication year, type of fiction	Ancient literary	921
Thirty Years of Modern Chinese Literature 1917 - 1949 (Qian et al. 1998)	Baidu Baike, Wikipedia	Description, publication year, type of fiction	Modern literary	574
A History of Contemporary Chinese Literature 1949 - 2000 (Hong 2007)		Description, publication year, type of fiction	Modern literary	
Douban website and DangDang website		Description, publication year, type of fiction	Modern literary	
Total				2074

NB. When we move from sampling fictional names to extract corresponding data online, we encounter a loss of data points due to the absence of pages. For example, we got 687 modern novel names from history books. However, only 574 of these fictions had corresponding pages on Baidu Baike.

Data sources for modern fiction: automatic sampling based on Wikidata

Considering that the selection of both ancient and contemporary novel materials is influenced by the preferences of literary historians and acknowledging the modern shift towards consumers favoring video formats of fiction over textual books, we implemented an automated method to sample all fictional works from China listed on Wikipedia.

We queried literary works, movies, TV series, anime and games, etc., through SPARQL services[3]. Then based on URLs and names of fiction in query services list, we used Python to scrape the description of fiction on Wikipedia pages.

---

[3] https://query.wikidata.org/

Table S2. **Sources and categories of summaries for automatic collection.**

Category	Wikidata id	Source of summaries	Number of summaries
Film	Q11424	Wikipedia	1313
Novels	Q47461344, Q7725634	Wikipedia	71
Tv series and program	Q5398426, Q15416	Wikipedia	66
Video game	Q7889	Wikipedia	63
			1513

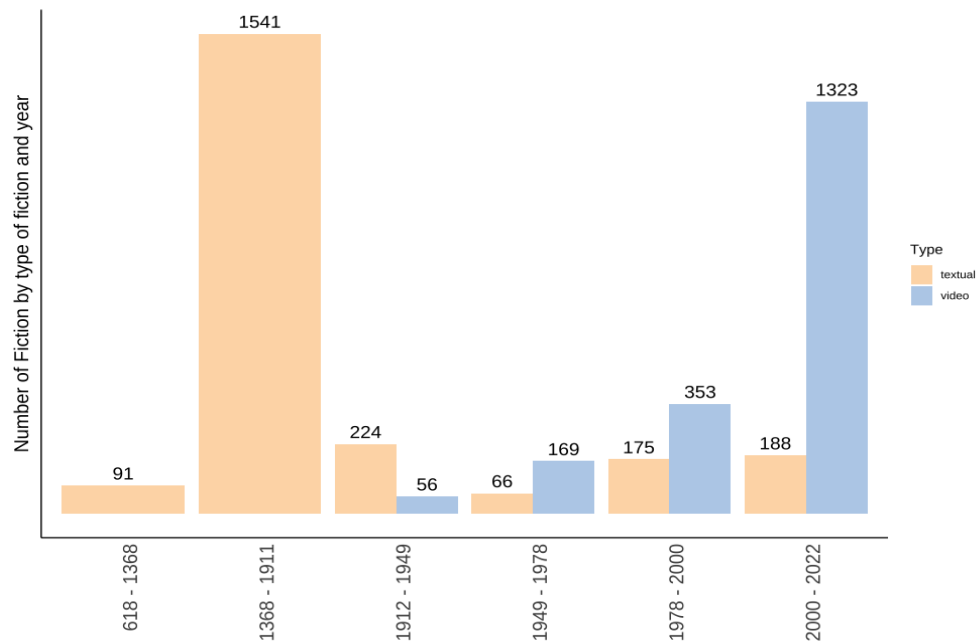Note: under the category as China (Q148)

Data distribution

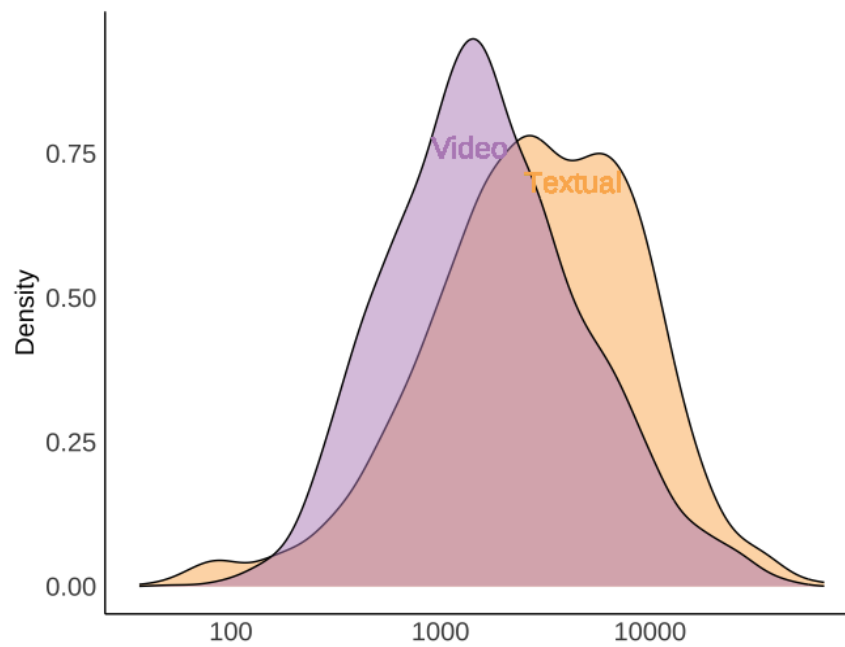**Fig. S1.** The numbers of fiction in each year intervals for textual and video type of fiction

Fig. S2. The proportion of fiction distribution against the total number of tokens in summary by type of fiction. Textual type of fiction contains longer length of tokens in summaries.

Environmental data collection

We used GDP per capita as a proxy for historical living standards, with data sourced from the Maddison Project 2020 Database (Bolt and Van Zanden 2020). To address missing data points, particularly from the imperial era, we employed different methods to process the GDP data. For periods with extended gaps over 10 years between data points, we used linear interpolation to estimate GDP pc. For shorter gaps, such as those occurring every ten years, we utilized the nearest available GDP data for each missing year.

Here are the details: 1 AD ~ 1000 AD: GDP assumed constant; 1000 AD ~ 1400 AD: Yearly linear interpolation; 1400 AD ~ 1620 AD: Use nearest available GDP data per year; 1620 AD ~ 1661 AD: Yearly linear interpolation; 1661 AD ~ 1911AD: Use nearest available GDP data per year. For 1912: Calculate the average of 1911 and 1913; 1913 ~ 2023: Apply linear interpolation for missing annual GDP data; 2019 ~ 2022: Incorporate World Bank data adjusted to 2011 USD values, with transformation ratio of 9,905 / 13,102. Specific data for these years: 13418 (2019), 13769 (2020), 16690 (2021), 16826 (2022).

Log_GDP is computed based on the logarithm of GDP pc.

ΔGDP (by decade) is computed to trace the relative changes of GDP in history, as following:

- Calculate averages of GDP pc for recent 10 years:

$$recent_{10avg} = \frac{\sum_{i=n-9}^{n} GDP_i}{10}$$

- Calculate averages of GDP pc for previous 10 years:

$$previous_{10avg} = \frac{\sum_{i=n-19}^{n-10} GDP_i}{10}$$

- Calculate the ΔGDP:

$$\Delta\text{GDP} = recent_{10avg} - previous_{10avg}$$

- ΔGDP_rate has been calculated (calculated by annual):

$$\Delta\text{GDPrate} = \left(\frac{recent\_10\_avg - previous\_10\_avg}{previous\_10\_avg} \right) \times \frac{100}{10}$$
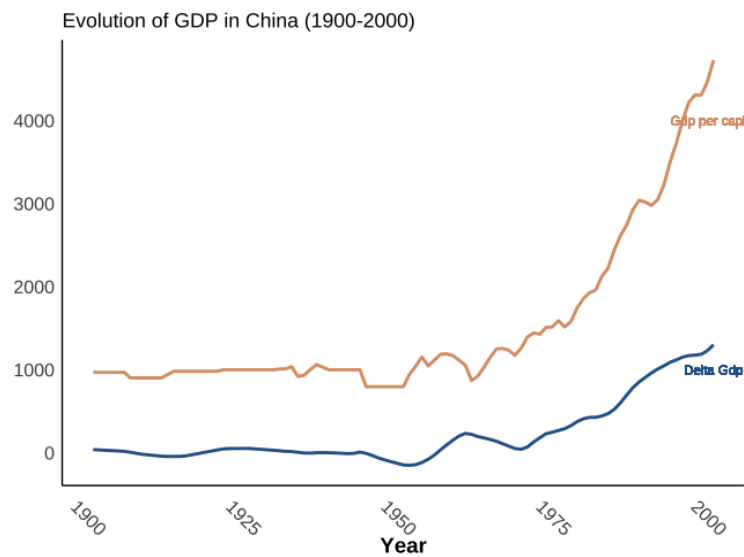
Fig. S3. Evolution of GDP pc and ΔGDP pc in history from 1900 to 2022. This is to present the differences between two treatments of GDP pc.
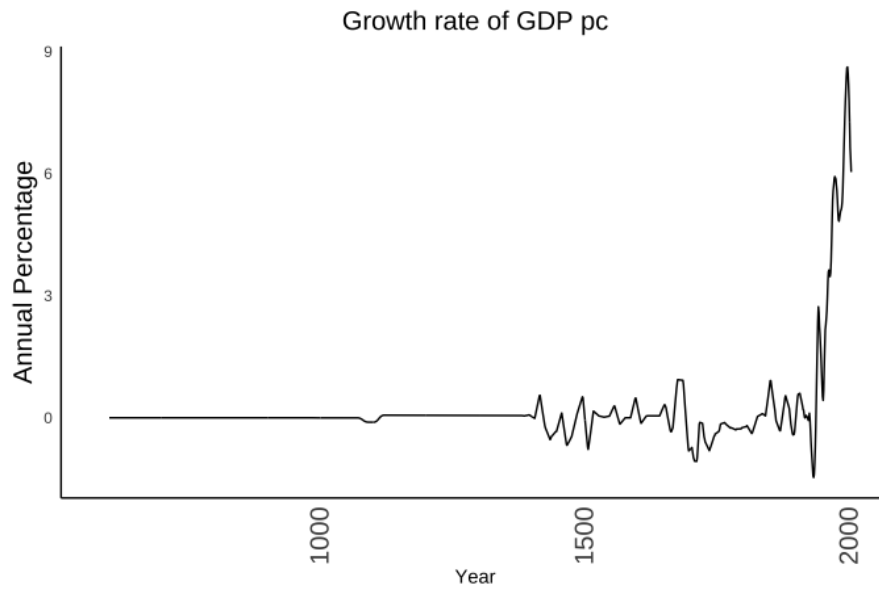
Growth rate of GDP pc

Fig. S4. Evolution of ΔGDP_rate in history from 600 to 2022.

Due to the current controversy in measuring GDP pc in imperial China, we have also considered other existing indicators that can reflect living standards in imperial China.

Urbanization ratio. According to Rozman (1974), the urbanization ratio (percentage of total population) increased from the mid Tang period (800) to the early Qing period (1650 CE), from 4.7 to 6.8. It then decreased to 5.9 in 1820. This trend is further investigated by Xu (Xu, van Leeuwen, and van Zanden 2018), which reveals that the urbanization ratio in inner China peaked in 1630 and then declined, with a recovery in inner China in 1918. Also see Figure S5.

Book production. According to the studies of T. Xu (Xu 2013), the number of books per million inhabitants in China around 1501-1550 was 19. This indicator increased to 236 during 1651-1700, but then declined to 176 during 1751-1800. Specifically, the number of books related to science, technology, and geography increased within the year interval.

Education. According to Elman (2000) and T. Xu (Xu 2013), the number of private academies per one million inhabitants has increased from the Song (960 CE) to the Qing (1911 CE). The indicator rose from less than 1 in the Song Dynasty to 6 in the Ming Dynasty and 16 in the Qing Dynasty.
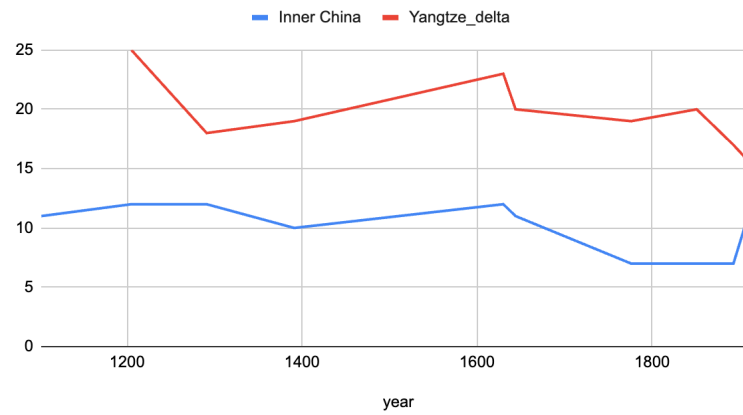
Fig. S5. Evolution of urbanization ratio (percentage of total population). Data source see (**Xu et al. 2018**).

Generating dictionaries for domains

Operationalization and Extraction of Seed Words for Each Value

To detect the presence of values in texts, it is necessary to establish a working definition for the values and then operationalize them into measurable concepts. For example, the nature of *love* is described as a commitment device characterized by passion, intimacy, and caregiving (Fletcher et al. 2015). Proximally, it manifests through emotional, behavioral, hormonal, and neuropsychological signatures (Bode and Kushnick 2021; Fisher, Aron, and Brown 2006; Fletcher et al. 2015). Mechanisms such as partner idealization, jealousy, passion are detectable cues in literary texts. They are embedded in fictional plots of "love at first sight 一见钟情," "jealousy 嫉妒," "love suicide 殉情," "falling in love 坠入爱河," and "happy endings 大团圆."

For each value, we aimed to identify traceable literary topics to measure them in texts. We pre-registered thirteen values along with fifteen corresponding topics. Out of these, nine values with their respective topics passed the validity test.

Modern related values

We predict that alongside economic development, modern values such as *romantic love*, *reciprocal cooperation*, *open-mindedness*, *self-development*, *friendship, caring, fairness, and parenting* will gain importance.

Romantic Love. Romantic love is a "commitment device" for motivating pair-bonding in humans (Fletcher et al. 2015) and was studied both in western countries and China (Gao 2001). Romantic love, here, refers to the components of strong feelings of lovers and long-term love bonds. In Chinese, the most representative keywords are "情 feelings of love" (Pan 2015). But this term in modern times might also be related to various kinds of love such as parental love (*qin qing*), and friendship love (*you qing)*. We kept the words 爱情 romantic love, not 情 qing in general. Love has multiple components related, including passion, intimacy or emotion (Sternberg 1986). The following seed words were thus selected: ['爱情 romantic love', '恋爱 love, '激情 passion', '亲密 intimacy'].

Open-mindedness. Open mindedness is the motivation of considering ideas and opinions that are new or different. It correlates with the motivation toward exploration, novelty or creativity, as examples in the imaginary world (Dubourg and Baumard 2021). We take topics of fictionality as proxies for *open-mindedness*. In Chinese fiction, narratives based on historical truth mixed with imaginary elements are quite common in history (Ma 2019; Zhang 2004). But the emergence of *gods and demon* novels later show a rising purpose for

entertainment-oriented fantasy. In the cover comments added to the Ming Dynasty editions of *The Journey to the West* around the 1600s[4] It is commented that: "*Journey to the West* speaks of non-existent things as if they exist and speaks of falsehoods as if they are true. This is the foremost trickery. 西游记说无事真如有事，说假事真如真事。此第一秒手也", which reflect a great emphasis on novelty and fictionality. Based on these features, we select seed words: ['虚构 imaginary', '幻 unreal', '想象 imagine*', '杜撰 coin'].

Self-development

Self-cultivation (did not pass the validity checks). In existing literature, self-development can be measured using various physical indicators, such as body condition, body mass, height, and other indexes (Glazier 1999). In cultural domains, self-cultivation is an important aspect of religious practices of self-development in pre-modern and modern society. This element can be found in ancient themes of *gods and demons*, and *martial arts* fiction, as well as in modern web-novels of *xiuzhen* fantasy 修真 (achieving immortality or becoming a god through cultivation) (Ni 2020). We assume this literary topic reflects well the motivation of somatic investment and discipline. Seed words are chosen as: ['成长 growth','修炼 cultivate','学习 study','自律 self-discipline','修身养性 cultivation','自强 self-improvement'].

Learning. We choose learning as another different literary topic to measure self-development. Learning mechanisms facilitate the children to gain capabilities, and to be more competent in a given environment. It also foils the dispositions toward exploration and play (Bjorklund 2022). Among all, this domain should be prevalent to audiences with higher openness (Rawlings, Flynn, and Kendal 2022), younger age and a favorable environment (Chen et al. 2018). In fiction, it appears in cultural content like Japanese shonen manga and Chinese martial arts stories. We formulate a dictionary for related concepts around gaining knowledge, skills, learning, breakthrough. Seed words: ['技能 skill','训练 training','培养 cultivate','学习 learn','进步 progress','挑战 challenge','成长 growth','能力 ability','技巧 skill','知识 knowledge','新知 new knowledge','才智 intelligence','提高 improve','掌握 command','突破 breakthrough']

Friendship. Friendship is typically a long term reciprocal cooperation compared to kinship which is based on genetic relatedness. Engagement in interdependent relationships is associated with a cognitive mechanism whose sensitivity varies with personality traits: people

---

[4] 英国国家图书馆藏《新刻全像批评西游记》，残存四卷二十回
 Newly Engraved Complete Annotated Edition of Journey to the West, preserved in the British Library, consisting of the remaining four volumes and twenty chapters.

higher in agreeableness, higher in extraversion, and lower in neuroticism are more motivated to form friendships in real life (Ashton et al. 1998; Oda et al. 2014). In traditional culture for example, the "friendship" is a social relationship distinguished from other relationships, as the "Fifth Relationship" proposed by Confucius. The word "friend 友" has been used in Chinese language since the Spring and Autumn period (770-476 BCE). The semantic stability of this domain aligns well with our methodology. We selected seed words based on the literature (Huang 2007; Shields 2020), including key concepts "友 you" and "信 trust". Seed words were [友 you, 知己 best friends, 知音 friends, 信任 trust, 交友 make friends].

Reciprocal Cooperation. Rising living standards are thought to shift value orientations toward reciprocal cooperation (Inglehart 1997; Martins and Baumard 2020). And these tendencies of cooperation are highly correlated with agreeableness and emotional stability (Ashton et al. 1998). Here we choose "mutual 互" related words as a proxy for reciprocal cooperation to demonstrate any relationship mutually: ['互相 mutual','互助 mutual aid','彼此 each other','互动 interact"].

Care (did not pass the validity checks). Caring, helping and comforting have been important traits in prosocial behavior (Dunfield et al. 2011). It relates to the concepts also as: *sharing, informing* (Warneken and Tomasello 2009)*,* has related emotion domains - *sympathy* and *empathy* (Eisenberg, Eggum, and Spinrad 2015; Eisenberg and Miller 1987; Vaish, Carpenter, and Tomasello 2009). In literary fiction, "侠 Chivalrous altruism" demonstrate well this domain. It is defined as "possessing a strong sense of social responsibility, patriotic and people-loving, clever and brave; assisting the poor and needy, punishing evil and promoting good; upholding justice and willing to help others"[5]. Related to this, seed words are selected as: ['安慰 comfort', '关心 care', '帮助 help', '朋友 friend', '友谊 friendship', '捐赠 donate', '慈善 charity', '协助 assist', '善良 kind', '同情 empathy']

Fairness (did not pass the validity checks). An intrinsic sense of *fairness* (morality) helps to secure reputation and facilitate cooperation (Baumard, André, and Sperber 2013). Its emotional components include *guilt* and *indignation* (Fitouchi, André, and Baumard 2021). Despite various cultural concepts around morality, such as the five Moral Foundation Theory (Haidt and Graham 2007), here we only take fairness or justice as proxy (e.g., good, bad,

---

[5] Baidu Baike description:
https://baike.baidu.com/item/%E4%BE%A0%E4%B9%89/1124832#:~:text=%E4%BE%A0%E4%B9%89%EF%BC%8C%E6%8C%87%E6%8A%B1%E6%9C%89%E5%BC%BA%E7%83%88,%E7%88%B1%E5%A6%BE%E7%9A%84%E6%80%A7%E5%91%BD%E6%96%AD%E9%80%81%E3%80%82

moral, fairness, guilt). External morality will not be included in dictionaries, such as 公允 public sense of fair, 公德 public morality, 仁义 ren. They are the moral contents emphasized by the community instead of internal sense. Seed words selected: ['道德 morality','良知 innate knowledge and sense of right and wrong','公平 fair','愧疚 guilt','愤慨 indignant']

Parenting (did not pass the validity checks). This involves the concept of a narrow scale of kinship which is related to parenting values. It evolves together with motivation of romantic love which provides better resources for the offsprings (Fletcher et al. 2015). More importantly, the motivation of parenting is sensitive to ecological cues of resources. A favorable environment increases the motivation of parental investment (Gibson and Lawson 2011; Szepsenwol et al. 2015). Related to this, we select concepts that better represent the behaviors or terms of the nuclear family. Seed words are: ["家庭 family","children 孩子"," 儿童 child", "爸 father","妈 mother","父 father","母 mother","养育 raise","哺育 raise","教 teach","亲蜜 closeness","贤妻良母 good wife and mother","天伦之乐 happy in family","丈夫 husband","妻子 wife"]

Tradition related values

We predict that alongside economic development, traditional values, which are *honor*, *authority*, *intensive kinship*, *decency*, and *tightness* will decrease in their importance.

Honor. Honor (or pride) and shame are distinct components of a culturally universal status-management system. Pride or honor track status gains, while shame tracked status losses (Durkee, Lukaszewski, and Buss 2019). The cultural representation among these two domains are well documented (see Lewis (2020)). One interesting case is "face culture 面子", a concept to demonstrate one's honor-shame toward others. We hypothesize the domain closely related to status-seeking and the importance people give to hierarchy and signaling effect. Based on the words provided by cultural studies of Lewis (2020), we select seed words around: [荣 honor, 辱 shame, 光耀 prominent in status, 贵 high status，贱 low status，尊 respect，敬 respect，卑 low status，面子 face，廉 purity，羞耻 shame].

Authority (did not pass the validity checks). "Dominance captures behavioral patterns found in social hierarchies that arise from agonistic interactions in which some individuals coercively exploit their control over costs and benefits to extract different from others, often through aggression, threats and/or intimidation." (Zeng, Cheng, and Henrich 2022). Usually, it is also described as *"rank", "permit", "forbid" "obligated"* (Cummins 2016). While pride

and shame are particular emotional reactions to status changing (Durkee et al. 2019). This domain has been tested in Europe context (Martins and Baumard 2020). We refer to these studies and selected: ['地位 status','品第 status','身分 status','荣誉 honor','骄傲 pride','服从 obey','顺从 obidience','权威 authority','实力 strenth','羞耻 shame','允许 permit','禁止 forbid',' 义务 obligate', '冲突 conflict', '战争 war', '叛乱 rebellion']

Tightness. Cultural tightness refers to the degree to which a society is characterized by rules and norms and the extent to which people are punished or sanctioned when they deviate from these rules and norms Comment (Gelfand, Nishii, and Raver 2006). We use the literary topics of punishment and discipline as proxies for cultural tightness.

Punishment (did not pass validity test). People react differently to moral wrongdoing considering punishment, attributions or judgment (Feinberg et al. 2018; Kwan 2016; Miller and Bersoff 1992). Specifically, people from a threatened environment are more likely to favor authoritarian leaders and harsh punishment (Nettle and Saxe 2021). China boasts a long-standing tradition of literary genres centered around moral punishment, such as *gongan* novels. These literary works primarily aimed at promoting virtue and condemning the immorality "惩恶扬善". This dictionary aims to measure the motivation of social control. We selected the following seed words: ['惩恶扬善 punish the evil, and praise the good','惩恶 punish the evil','责怪 blame','归咎 blame to','谴责 condemn','报应 retribution'].

Discipline. There are well-established questionnaires measuring related traits for discipline such as HEXACO-PI-R (Ashton and Lee 2009), Self-Control Scale (Tangney, Baumeister, and Boone 2004), Delay gratification scale (Mischel, Shoda, and Rodriguez 1989). This domain might be well presented in sport films that feature training, planning and self-control. And empirical findings suggest this genre is well interested in people who are higher in conscientiousness (Nave, Rentfrow, and Bhatia 2020). It might underscore how tightly people stick to norms, rules and emphasize hard work. Seed words are: ['抵制诱惑','戒除','懒','口无 遮拦','失控','不计后果','按部就班','计划','婉拒','善变','见异思迁','口快心快','冲动','规避','挥 霍','井井有条','放纵','自律','可靠','理智','感情用事','保守','秘密','临时抱佛脚','气馁','三思而 后行','养生','修身','清淡','玩忽职守','消极怠工','分心','远谋','长久之计','深谋远虑','长远','禁 不住','忍耐','全面考虑','发脾气','打岔','酗酒','滥用','按时','平心静气','远志','目标长远','迫不 及待','有始有终','耐心','固执己见','自发','失言','玩忽','定心','厚积薄发','深思熟虑','一心二 用','思考','全神贯注','自制力','意志力','目光远大','眼前利益','一丝不苟', '勤奋', '严格', '勤 快', '刻苦', '严谨', '认真', '努力', '勤劳', '锲而不舍', '进取', '坚韧不拔', '懒散', '懒洋洋', '三天

打鱼两天晒网', '懒惰', '散漫', '懒虫', '粗心', '马虎', '坚持', '仔细', '考虑', '慎重', '计划', '有计划', '倾尽全力', '马马虎虎', '条理', '冲动', '谨慎', '细心', '自律', '放纵', '准时', '分心', '分神', '气馁', '放弃', '注意力', '有始有终', '兴趣', '不放弃', '挫折', '坚毅', '泄气','读书','知识','学习','能力','技能','理想','事业','梦想','念书','好学','训练','徒','师','弟子','学生','修炼','修行','锻炼','修身','修己'].

English translation offered: ['resist temptation', 'abstain', 'laziness', 'lounging', 'out of control', 'reckless', 'in order', 'planning', 'politely refusing', 'fickle', 'seeing things differently', 'fast-talking', 'impulsive', 'circumventing', 'profligate', 'well-organized', 'indulgent', 'self-disciplined', 'reliable', 'sensible', 'emotional', 'conservative', 'secretive'. 'improvise', 'discourage', 'think twice before you act', 'nurture', 'cultivate', 'light', 'neglect', 'negligence', 'negligent', 'distraction', 'far-seeing', 'long-range', 'far-sighted', 'far-sighted', 'long-term', 'forbidden', 'patience', 'full consideration', 'tantrums', 'interruptions', 'alcoholism', 'abuse', 'on time', 'calmness'. 'far-sighted', 'aiming long term', 'impatient', 'beginning and ending', 'patient', 'stubborn', 'spontaneous', 'slip of the tongue', 'negligent', 'fixed mind', 'thick-headed', 'thoughtful', 'double-minded', 'thinking', 'full attention', 'self-control', 'willpower', 'far-sighted', 'immediate interest', 'meticulous', 'diligent', ' rigorous', 'diligent', 'painstaking', 'rigorous', 'serious', 'hardworking', 'diligent', 'persevering', 'enterprising', 'persevering', 'lazy', 'lazy', 'lazy', 'lazy', 'scatterbrained', 'lazy', 'careless', 'sloppy', 'persistent', 'careful', 'considerate', 'prudent', 'plan', 'planned' , 'give your best', 'sloppy', 'organized', 'impulsive', 'cautious', 'careful', 'self-discipline', 'indulgent', 'punctual', 'distracted', 'distracted', 'discouraged', 'give up', 'attention', 'start and finish', 'interest', 'don't give up', 'frustration', 'perseverance', 'discouragement', 'reading', 'knowledge', 'learning', 'ability', ' skill', 'ideal', 'career', 'dream', 'study', 'good learning', 'training', 'apprentice', 'teacher', 'disciple', 'student', 'cultivate', 'practice', 'exercise', 'cultivate body', 'cultivate self'].

Decency. Discipline measures the extent to which an individual restrains and defers immediate gratification in pursuit of long-term goal, during which might involve grit and effortless. In contrast, decency focuses on how an individual condemns, rejects, and judges the harmless actions of others, particularly on sexual, bodily pleasure, and habits. It is also related to moral discipline (Fitouchi, André, and Baumard 2023). Historically, it might be relatively scarce in hunter-gatherer societies, increasing in large traditional societies, and decreasing in modern societies due to the rise of social trust in some societies (Fitouchi et al. 2023). Decency have been promoted in traditional Chinese religion, such as Chan Buddhism "苦行 asceticism", Taoism "无为 Doing Nothing" or the Confucian school of thought

advocates "women's chastity" or "a gentleman's integrity and impeccability". Particularly, these values are deeply engraved in the fiction in imperial, including the restraint on female character chastity, promoting modesty, and decent manners, as well as the condemnation of excessive bodily pleasures and personal gains in sentiment novels (世情小说) during Ming and Qing dynasty. Based on the work of (Fitouchi et al. 2023) and Yü (2021) and moral foundation dictionary of puritanism (Frimer et al. 2019), we select seed words:['享乐', '贪色', '贪吃', '诱惑', '纵欲过度', '声色之欲', '穷奢极欲', '淫', '秽', '淫秽', '污秽','淫荡', '暴饮暴食', '奢', '私利', '利欲熏心', '贪图', '贪', '私欲', '自私自利', '一己私利', '公益', '靡靡之音', '低级趣味', '糜烂', '堕落', '不文明', '不卫生', '不干净', '不得体', '不道德', '失节', '不守妇道', '礼仪', '文明', '卫生', '干净', '得体', '礼节', '廉洁', '纯', '贞洁', '贞节', '妇道', '正直', '廉正', '懒惰', '好逸恶劳', '坐享其成', '偷懒', '不务正业','寡欲', '戒色', '清心寡欲', '禁欲', '禁欲主义', '戒除', '摒弃', '淡泊名利', '朴素', '克己', '戒奢崇俭','节俭', '节制', '勤俭节约','高尚','修身养性', '静坐', '困苦', '贫苦'] ['pleasure', 'lust', 'gluttony', 'temptation', 'excessive indulgence', 'lust for sound and sex', 'extravagant desire', 'obscenity', 'obscenity', 'obscenity', 'filth', 'lasciviousness', 'overeating', 'extravagance', 'self-interest', 'lust for profit', 'covetousness', 'greed', 'selfishness', 'selfishness', 'self-interest', 'public good', 'decadence', ' low-class', 'rotten', 'depraved', 'uncivilized', 'unhygienic', 'unclean', 'indecent', 'immoral', 'out of order', 'ungentlemanly', 'ungentlemanly', 'ungentlemanly', 'etiquette', 'civilized', 'hygienic', 'clean', 'decent', 'etiquette', 'incorruptible', 'pure', 'chaste', 'chastity', 'womanly', 'upright', 'incorruptible', 'lazy', ' good for nothing', 'sitting on one's laurels', 'lazy', 'not doing one's job properly', 'lust', 'abstinence', 'abstinence', 'asceticism', 'abstinence from lust', 'renunciation', 'indifference to fame and fortune', 'simplicity', 'self-restraint', 'abstinence from luxury and frugality', 'frugality', 'moderation', 'diligence and thrift', 'nobility', 'cultivating one's body and mind', 'meditation', 'hardship', 'poverty ']

Intensive Kinship. Here the intensive kinship involve concept of: wider kinship networks; Stem families; Indirect parental investment (lineage status, inheritance, clan persistence); Respect for kins and lineage (孝); Cousin marriage. This domain targets the cooperation and techniques around the kinship that are formed to ingroup obedience and solidarity in face of threats and outgroup competitions. Seed words are: ["亲戚 relatives", "家族 expanded family", "血缘 blood-bond","宗族 clan","祖宗 ancestors","表妹 cousin","孝 filial piety", "继承 inheritance","尊老爱幼 respect the old and love the young","四世同堂 four generation in one house", "姻亲 affinity"]

Expand list of words through synonyms tools

To broaden the range of words by incorporating more related semantics, synonyms for each word are identified. *Wantwords* is a language tool developed by Tsinghua University (Zhang et al. 2019) for finding relevant words. Example of the generated relevant words for "爱情 romantic love" is shown in Figure S6. We use this tool to find relevant words for each seed words, removing ambiguous words, for instance, the ninth word - "情谊 friendship". A list of words is selected for further processing for each topic from wantwords results.

1. 恋情	21. 男欢女爱	41. 风情	61. 暗恋
2. 情爱	22. 儿女之情	42. 多情	62. 厚爱
3. 有情	23. 相思	43. 隆情	63. 爱河
4. 痴情	24. 情深一往	44. 高情	64. 情场
5. 情意	25. 儿女情	45. 一见钟情	65. 泛爱
6. 私情	26. 爱欲	46. 儿女情多	66. 性爱
7. 爱恋	27. 挚情	47. 迷恋	67. 眷爱
8. 恋爱	28. 春心	48. 感情	68. 系恋
9. 情谊	29. 有意思	49. 你恩我爱	69. 花心
10. 情丝	30. 痴爱	50. 见爱	70. 一往深情
11. 恩爱	31. 钟情	51. 一见倾心	71. 情义
12. 情海	32. 单相思	52. 你贪我爱	72. 亲爱
13. 心爱	33. 一往情深	53. 炽情	73. 殷
14. 挚爱	34. 热爱	54. 盛情	74. 风
15. 热恋	35. 恩恩爱爱	55. 相亲相爱	75. 恩
16. 有意	36. 艳情	56. 倾心	76. 怜
17. 春情	37. 笃爱	57. 求爱	77. 恋慕
18. 情痴	38. 情分	58. 儿女情长	78. 高情厚谊
19. 动情	39. 纯情	59. 深情	79. 爱慕
20. 恋	40. 眷眷之心	60. 深情厚意	80. 罗曼蒂克

Fig. S6. Example of output in Wantwords using "爱情 romantic love". The blue color shows the high level of association of the term with "爱情 romantic love": "love 恋情, love 情爱, affection 有情, infatuation 痴情, affection 情意, affair 私情…"

Expand list of words through word embedding

To supplement the drawbacks of manually selecting vocabulary, we also conducted automated dictionary generating based on Word2vec embeddings (see the path of method 2 dictionary in figure S7)

Get synonyms for seed words. For each word in the seed words list, we get 10 most similar words output using python package "synonyms[6]".

Get Baidu Baike Corpus. *Wantwords* **is trained by Wikipedia's corpus**. But a large amount of our dataset is collected from Baidu Baike. Wikipedia and Baidu Baike might perform differently in certain entries. For instance, Wikipedia is more accurate and informative than Baidu Baike in historical entries (Shang 2018). We alternatively trained a vector space using a corpus collected from Baidu Baike. The collecting script can be accessed on *baibubaike_collector.ipynb*[7].

Train vector space using Word2vec. A function is used to construct a vector space model using Word2Vec from the Baidu Baike corpus, with a minimum count threshold set for including words. The resulting word embeddings effectively capture semantic relationships between words in the corpus, thereby facilitating tasks such as word similarity detection.

Using vector space to evaluate and expand the synonyms. For each term in the synonym's lists, we calculate the 10 most similar words based on their embeddings in the vector space defined by the Word2Vec model. We evaluate the words in the lists, keeping relevant words, removing non-relevant words. For example, "honor 荣耀" is closely distributed with "morality 道德" in the vector space, but we treat them as two different domains in our analysis. So, we remove "morality 道德" from dictionary for the value *honor*. Similarly, for love dictionary, "友情 friendship", "亲情 family love", "青春 youth" is closely distributed with "love 爱情", yet are removed by us from the love dictionary. Details of codes are available on the script of *Generate_dictionaries_chinese_fiction.ipynb*[8]

Approximate translation. We add a proximate translation for the dictionaries in English (see pre-registration documents).

---

[6] https://github.com/chatopera/Synonyms/
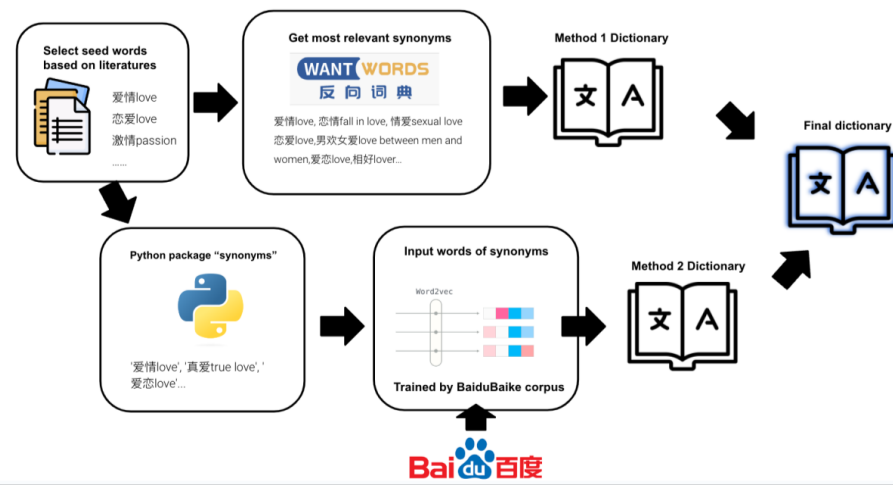
[7] https://osf.io/9nxpe/

[8] https://osf.io/9nxpe/

Fig. S7. Workflow of generating dictionaries for domains. We started from seed words and expanded the list of words through two pipelines. And in the end, for the dictionary that passed the validity test, we merge them into a bigger final dictionary.

The final dictionaries for each value are available at: https://osf.io/ed93s

The lists of most frequently occurring words (only show the top 10) for each domain are available in Appendix 1.

Validity test

Internal validity

Movie genres reflect the thematic focus of movies. For example, a romantic comedy in Wikipedia is described as: "In a typical romantic comedy, the two lovers tend to be young, likable, and seemingly meant for each other, yet they are kept apart by some complicating circumstance (e.g., class differences, parental interference, a previous girlfriend or boyfriend) until, surmounting all obstacles, they are finally united." In contrast, genres such as fantasy, historical, or martial arts films focus on markedly different themes. These distinct thematic concentrations have been used to assess the internal validity of our dictionary. For example, we hypothesize that romantic comedy films should exhibit a higher frequency of words related to romantic love compared to historical films. We chose film data because the genres are well-documented in Wikidata. Based on URLs links extracted from Wikidata, we collected summaries of films from Wikipedia. The sample sizes for each genre, identified by Wikidata IDs, are detailed in Table S3.

Table S3. Wikidata query films sample size for validity test.

Genre	Sample size	Wikidata id	Langugaes
Romantic comedy films	657	Q860626	
Historical films	161	Q17013749	
Chinese drama films	447	Q130232	
Fantasy films	682	Q132311	
Martial art films	423	Q1033891	Chinese
Coming of age films	174	Q2975633	
Adventure films	199	Q319221	
Drama films	91	Q130232	

Results of validity test

Preprocessing, including tokenization and removal of punctuation, is conducted on summaries. Subsequently, the ratio of values is calculated for each sample. An independent t-test is performed to compare the topic ratios across two film genres. Figure S8 illustrates three examples of analysis for *love*, *open-mindedness*, and *reciprocal cooperation*. Table S4 present the statistic results of group comparation. Through this procedure, we removed the topics that did not pass the validity test, which were *self-cultivation*, *punishment*, and *parenting*.
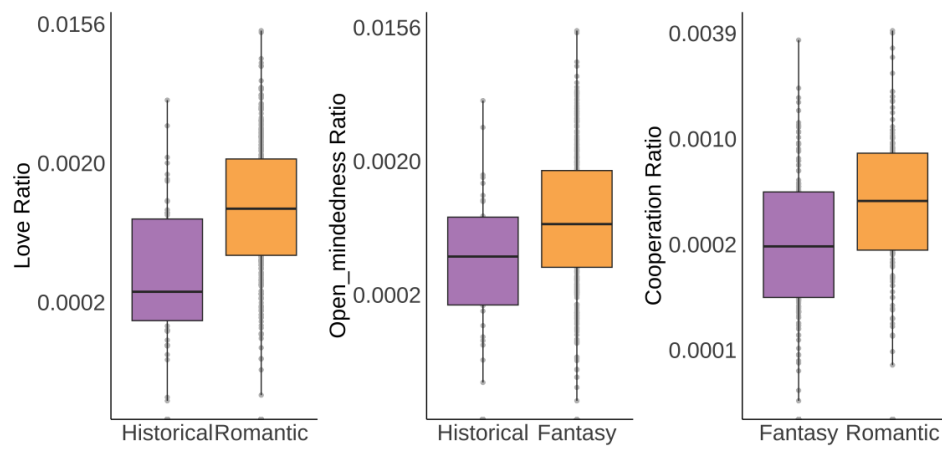
Fig. S8. Box plot of ratio of topics between summaries of genres of films for *love*, *open-mindedness* and *reciprocal cooperation*.

Table S4. Results. T-test for 15 pre-registered topics developed for Chinese fiction

Values	Genre for this domain	Genre to compare	Method 1 dictionary	Method 2 dictionary	Combo of two dictionaries
Love √	Romantic comedy films	Historical films	mean: 0.0005 >0.0001 t = 3.633602 p < 0.001***	mean: 0.0007 >0.0001 t = 3.560278 p < 0.001***	mean: 0.0007 > 0.0002 t = 4.452001 p < 0.001***
Open_mind edness √	Fantasy films	Historical films	mean: 0.0004 > 0.00003 t = 5.819146 p < 0.001***	mean: 0.0007 > 0.0001 t = 4.748734 p < 0.001***	mean: 0.0007 > 0.0001 t = 4.748734 p < 0.001***
Self cultivation ×	Martial art films	Romantic comedy films	mean: 1.6517e-05 > 1.5297e-05 t = 0.093079 p = 0.92^{NS}	mean: 0.00016 < 0.00020 t = -0.866056 p = 0.38^{NS}	mean: 0.0001 < 0.0002 t = -1.31211 p = 0.18^{NS}
	Martial art films	Non martial art films1	mean: 1.6517e-05 > 5.3714e-06 t =-0.638725 p = 0.52^{NS}	mean: 0.0001 > 0.00009 t = 2.582005 p < 0.01**	mean: 0.00016 > 0.00010 t = 2.207526 p < 0.05*
Sel-developmen t √	martial art films	romantic comedy	mean: 0.0002 >9.6925e-05 t = 4.9155 p < 0.001***	mean:7.5218e-05 >4.4239e-05 t = 1.7276 p = 0.08	mean:0.0002>0.0001 t = 5.1708 p < 0.001***
Friendship √	Romantic comedy films	Historical films	mean: 0.0008>0.0001 t = 6.224382 p < 0.001***	mean:0.0005>7.7541e-05 t = 6.370506 p < 0.001***	mean: 0.0008 > 0.0001 t = 6.311280 p < 0.001***
Reciprocal Cooperatio n √	Romantic comedy films	fantasy	mean: 0.0001 >8.0451e-05 t = 2.7998 p < 0.01**	mean: 4.3474e-05 >2.1146e-05 t = 2.4364 p < 0.01**	mean: 0.0001 >8.3216e-05 t = 2.6530 p < 0.01**

Honor√	Historical films	Romantic comedy films	mean:9.4084e-05<9.9367e-05 t = -0.167504 p = 0.86[NS]	mean:0.0002>0.0001 t = 2.843761 p < 0.001***	mean: 0.0002>0.0001 t = 2.355191 p < 0.05*
Fairness √	Martial art films	Non martial art films	mean: 9.7926e-06 >4.8793e-05e-06 t = 1.891742 p =0.05[NS]	mean: 3.5463e-05 > 7.4153e-06 t = 5.879730 p < 0.001***	mean: 3.4191e-05 > 8.0483e-06 t = 5.807632 p < 0.001***
Caring √	Martial art films	Non martial art films	mean: 0.0006 > 0.0001 t =15.124905 p < 0.001****	mean: 0.0003 >0.0002 t = 5.317419 p < 0.001***	mean: 0.0007>0.0002 t = 15.388176 p < 0.001***
Authority √	Historical films	Romantic comedy	mean: 9.9069e-05 > 7.3638e-05 t = 0.897323 p = 0.36[NS]	mean: 0.0006 > 0.0001 t = 7.243049 p < 0.001***	mean: 0.0006 > 0.0002 t = 7.769134 p < 0.001***
Punishment ×	martial art films drama	romantic comedy	mean: 2.8365e-05 >1.8581e-05 t = 0.6190 p = 0.53[NS]	mean: 2.6755e-05 >8.6807e-06 t =1.8777 p = 0.06[NS]	mean: 4.6697e-05 >2.4810e-05 t =1.1951 p = 0.23[NS]
Intensive kinship √	Historical films	Romantic comedy	mean: 0.0003 >0.0002 t = 2.563560 p < 0.05*	mean: 0.0001 > 0.00007 t = 3.204965 p < 0.01**	mean: 0.0004 > 0.0002 t = 3.838280 p < 0.001***
Parenting ×	Romantic comedy	Historical films	mean: 0.00096 >0.00093 t = 0.233420 p = 0.81[NS]	mean: 0.0008 > 0.0005 t = 2.357898 p < 0.05*	mean: 0.0015 > 0.0012 t = 1.527864 p = 0.12[NS]
Tightness √	martial art films	romantic comedy films	mean: 0.00044 > 0.00014 t = 6.495550 p < 0.001***	mean: 0.0003 > 0.0001 t = 5.082463 p < 0.001***	mean: 0.00049 >0.00016 t = 6.747365 p < 0.001***

Decency √	drama film	adventure film	mean: 4.3495e-05 > 1.9049e-05 t = 1.411987 p = 0.15[NS]	mean: 5.8689e-05 > 9.4375e-06 t = 2.655590 p < 0.01**	mean: 0.00029 >0.00013 t =3.083797 p < 0.01**

Non martial art movies: combined film genre of romantic comedy film, fantasy film, coming of age film, historical film, adventure film; mean: the mean ratio of each dictionary in one description; [NS]: Non-significant.

External validity

To assess whether values measured in fiction align with empirical surveyed values. Samples span from 1978 to 2022 are chosen since most surveys fall within this timeframe. We modeled the correlation between occurrences of words with years, controlling for length of summaries, type of the fiction, and sources of the summary. The coefficient (β2) in the following model represents the estimated changing trends for the last 44 years. We summarized the literature around these topics in Table S5.

$$\text{Frequency}_i \sim \text{Poisson}(\lambda_i)$$

$$\log(\lambda_i) = \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{Year}_i + \beta_3 \cdot \text{Source}_i + \beta_4 \cdot \text{Type}_i + \epsilon_i$$

$$, \text{where Year} > 1978$$

Table S5. Summaries of literatures documenting value changes in modern China.

Topics	Traits related	Trends
Love	The importance of love and intimacy in courtship from 1949 - 1999 (Yan 2003)	Increased
Open-mindedness	Perception of exploration, curiosity across generations. Personality of openness-to-experience across generations (Peng and Luo 2021; Zhou et al. 2018)	Increased
Friendship	The importance of friendship in life across generations (Sun and Wang 2010)	Increased
Reciprocal Cooperation	No direct studies on this topic.	Increased (Cui 2017)
	Related topics: perceived social trust and interpersonal trust; Social support	Decreased (for interpersonal trust, see (Xin and Xin 2016, 2017)
Self-development	Value of self-actualization, self-esteems, and narcissism (Sun and Wang 2010); Importance of academic achievement and self-development across generations (Zhou et al. 2018)	Increased
Caring	Prosocial attitudes, tendencies, and behaviors (Chen and Yang 2024)	Increased (~ 2014) then decreased since

Fairness	Folk beliefs of fairness and equality (Xu and Hamamura 2014)	Increased
Authority	-	-
Honor	Culture of mianzi (面子 face) (Shi, Wu, and Kuang 2018)	Still important
Tightness	Value toward effortness (Zeng and Greenfield 2015); Within China: Tightness is related to urbanization, economic growth, better health, greater tolerance (Chua, Huang, and Jin 2019)	Decreased
Decency	Values toward traditional ethics (Schwartz 2006; Xu and Hamamura 2014)	Decreased
Intensive Kinship	Values toward traditional family values (Xu and Hamamura 2014); Collectivism (Ogihara 2023)	Decreased

Figure S9 reports the estimated coefficient of β2 (years). No significant changes were found in the domain of *authority* (which we had hypothesized would decline), *caring* (which we had hypothesized would rise) and *fairness* (which we had hypothesized would rise), which are removed in the formal analysis.
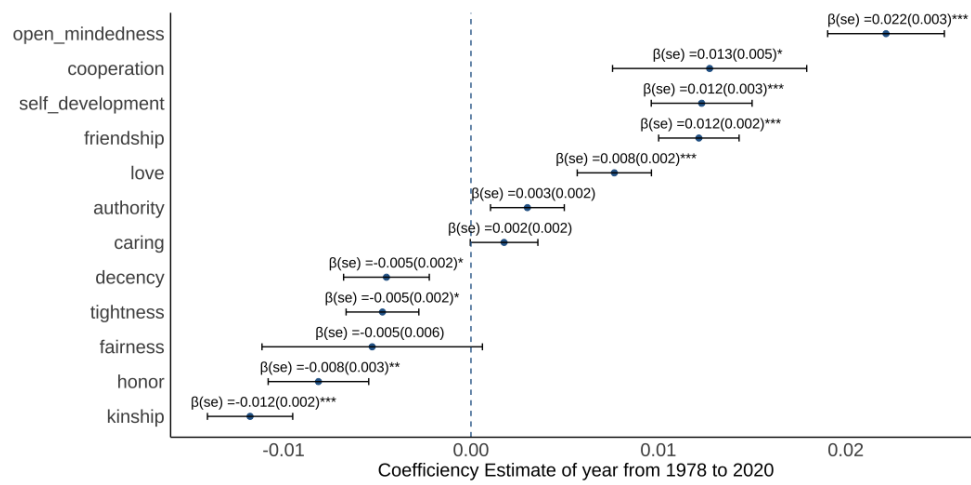
Fig. S9. Coefficient estimation of years from 1978 to 2022 for 11 topics.

Comparing full texts and descriptions over *romantic love*

We aim to assess how well summary data captures the original information in full texts. The analysis here focuses on *romantic love*, a prominent topic in Chinese fiction since the Tang Dynasty. The semantics of words related to *romantic love* are highly concentrated, making them easily traceable in original texts using a dictionary approach, in contrast to themes like *cooperation* or *fairness*. To do so, we build a new dataset containing 802 full texts of ancient Chinese novels (pre-1912) and try to investigate the correlation of importance of romantic love in original texts and in summaries. The full texts are collected from online open source website:

- 汉典古籍 Han canon (https://gj.zdic.net/);
- 古诗文网 Gushiwen (https://so.gushiwen.cn/guwen/);
- Wikisouce (https://zh.wikisource.org/wiki/Portal:%E5%B0%8F%E8%AF%B4);
- Github project (https://github.com/luoxuhai/chinese-novel).

Through crawling and organizing, we stored each novel in the format of texts.

An issue we encounter with original ancient texts is languages. For instance, much of the fiction during the Tang Dynasty was written in classical Chinese, where a single character is a word. However, by the Yuan Dynasty and onwards, fictional writing gradually transitioned to vernacular Chinese and spoken language. To tackle this challenge, we first manually collect online translations for classical languages and replace the classical languages texts. Notably, most Tang *chuanqi* texts have modern translations available. Figure S10 to Figure S11 illustrates the distribution before and after adding translation.
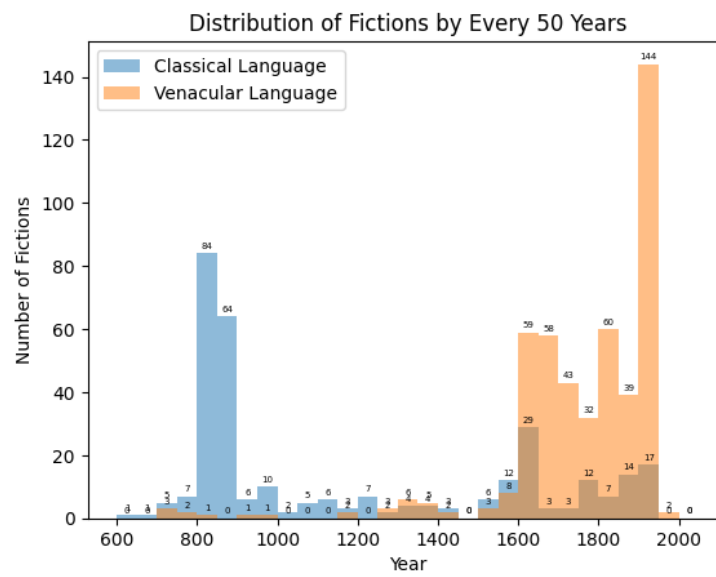
Fig. S10. The distribution of full text fiction by type of languages (before adding translations).
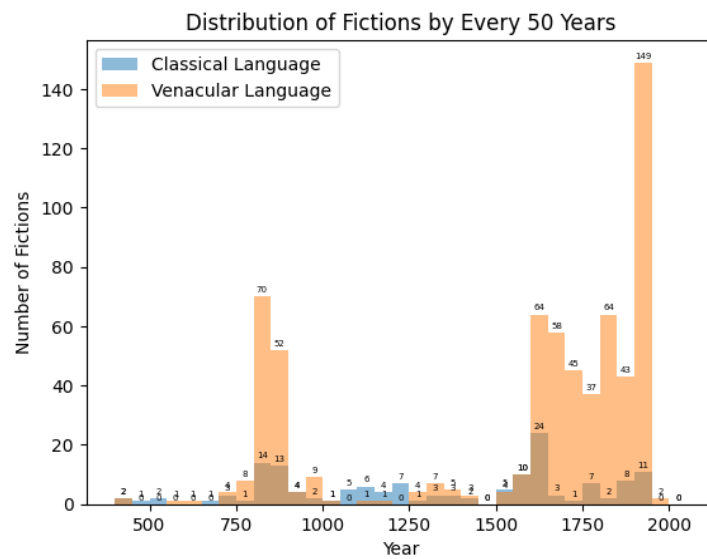
Fig. S11. The distribution of full text fiction by type of languages (after adding translations).

Then for each text, we went through the following procedure:

Classification: For 802 fictions (after adding translations), we classify the texts into two groups: vernacular or classical. This is done by automatically identifying texts based on the language feature.

Tokenization of texts: for classical texts, we tokenize the sentences into "tokens" using Jiayan[9]. For vernacular texts, we tokenize the sentences using Jieba[10].

Select seed words for full texts: We chose seed words for romantic love based on the dictionaries generated previously. Additionally, we manually included some words closely associated with romantic love in ancient contexts. For instance, "连理枝" (intertwined branches)" is a metaphor symbolizing long-term romantic companionship in Chinese culture.

Find neighbors' words in full texts: for each seed word (91 seed words are chosen), we extract 10 most frequently tokens around the seed words separately in vernacular and classical language texts. We call them neighbors' words. For instance, in Figure S12 line 1, for the seed word "两情知" (mutual affection), its high-frequency neighboring words include "芳情" (fragrant affection), "欢情" (joyous affection), and "韶光" (beautiful time).

- **Select relevant Neighbor words into lists**: for all neighboring words, we select those related to love and remove other unrelated words (for example, words of descriptions of the environment, character facial descriptions).
- **Using word2vec embedding to evaluate and expand list**: we take the same procedure as in previous methods, to generate vector space using input of ancient full texts as corpus. And based on the vector space, we evaluate and expand the lists of words. See **Figure S14.**
- **Manual selection and construct the final dictionaries for vernacular and classical dictionaries.**

---

[9] https://github.com/jiaeyan/Jiayan

[10] https://github.com/fxsjy/jieba

	seed word	neighbors
0	情牵	[依依, 尽日, 梦魂, 移步, 夜夜, 珠泪, 那堪, 泪眼, 绿杨, 幽恨, 玉颜]
1	两心知	[芳情, 几经, 欢情, 半吐, 韶光, 响夜蛩, 秋景, 艳阳, 忆旧游, 易老, 晓日]
2	依依	[亲亲, 妹妹, ：, 怜悯, 好生, 负国, 爹娘, 哥哥, 教人, 奉承, 悌]
3	儿女之情	[邂逅, 相逢, 晤, 今宵, 与君, 姻缘, 久别, 因缘, 相爱, 异乡, 多情]
4	情生	[欢喜, 说话, 咁, 咯, 烦恼, 暗想, 呀, 老婆, 妹妹, 哩, 温存]
...	...	...
87	情许	[云雨, 几回, 夜夜, 临风, 无限, 懒, 对月, 凄凉, 恼人, 含情, 情话]
88	真情	[风月, 春色, 桃李, 临风, 花柳, 秋月, 玉树, 春光, 繁华, 芳草, 杏花]
89	与子偕老	[合巹, 亲迎, 花烛, 涓吉, 成礼, 择吉, 傧, 沐浴, 设席, 辞去, 设筵]
90	连理枝	[合欢, 钿, 雕阑, 镜里, 阑干, 翡翠, 云鬓, 芍药, 枝上, 翠羽, 露滴]
91	相爱	[缱绻, 绸缪, 缠绵, 情意, 琴瑟, 欢娱, 情好, 眷恋, 臻至, 曲尽, 欢谑]

Fig. S12. Lists of neighbors' words for seed words in full texts.

```
1 list_of_love_w2v = get_word2vec_list(wordlist_classical, word2vec_output_without_punc_↑
2 index = 0
3 for w2v_list in list_of_love_w2v:
4     wordlist_df.loc[index, 'seed word'] = wordlist_classical[index]
5     wordlist_df.loc[index, 'neighbors'] = w2v_list
6     print(index, w2v_list)
7     index += 1
```

```
0 ['依依', '尽日', '梦魂', '移步', '夜夜', '珠泪', '那堪', '泪眼', '绿杨', '幽恨', '玉颜']
1 ['芳情', '几经', '欢情', '半吐', '韶光', '响夜蛩', '秋景', '艳阳', '忆旧游', '易老', '晓日']
2 ['亲亲', '妹妹', '：', '怜悯', '好生', '负国', '爹娘', '哥哥', '教人', '奉承', '悌']
3 ['邂逅', '相逢', '晤', '今宵', '与君', '姻缘', '久别', '因缘', '相爱', '异乡', '多情']
4 ['欢喜', '说话', '咱', '咯', '烦恼', '暗想', '呀', '老婆', '妹妹', '哩', '温存']
5 ['伴侣', '好景', '朝朝', '幽情', '游蜂', '春梦', '怅望', '别是', '苟细', '裘马', '出尘']
6 ['婉恋', '粉汗', '响夜蛩', '菲菲', '纱衫', '汀花', '潋', '半吐', '斜簪', '娇欲语', '盟鸥']
7 ['钟爱', '年长', '姊庆宜', '早亡', '溺爱', '充后宫', '早丧', '许嫁', '敬重', '直舍人院', '妒悍']
8 ['知己', '异类', '古人', '天道', '世人', '兴废', '但知', '世间', '造物', '仁人', '善人']
9 ['眷恋', '情意', '迷恋', '喜愠', '戚戚', '缱绻', '念生', '相亲', '暮景', '常存', '怜惜']
10 ['佳人', '红颜', '春风', '多情', '才子', '相思', '风流', '春光', '伤心', '春色', '浮生']
11 ['情人', '海岳', '宁知', '孤负', '舜日', '最苦', '十洲', '难可', '通儒', '怜伊', '翻成']
12 ['比翼', '连理', '和鸣', '柳絮', '交颈', '飞花', '渺渺', '玉树', '雨后', '满天', '漠莫']
13 ['情意', '情好', '相亲', '缱绻', '柔情', '眷恋', '缠绵', '绸缪', '情深', '恩爱', '欢情']
14 ['佳期', '今宵', '相思', '恩情', '春光', '红颜', '别离', '辜负', '时节', '难再', '相逢']
15 ['鸳鸯', '明月', '梨花', '梅花', '芙蓉', '玉树', '蝶', '杨柳', '嫩', '桃花', '春风']
16 ['爱人', '刚毅', '专务', '骄人', '宏远', '浮躁', '好义', '虽贵', '清慎', '狡猾', '恢弘']
17 ['相亲', '情深', '恩爱', '情意', '鱼水', '欢爱', '久别', '绸缪', '眷恋', '契合', '相知']
18 ['脉脉', '泪湿', '幽恨', '横波', '孤影', '寒灯', '眉黛', '飞絮', '舞袖', '肠断', '忆旧游']
19 ['柔情', '张绪', '临风', '情怀', '慵', '交颈', '韶光', '秋波', '娇姿', '雨打', '闲情']
20 ['恋恋', '眷恋', '遽舍', '他适', '妻儿', '戚戚', '膝下', '私爱', '能忍', '远行', '迟留']
21 ['爱惜', '怜爱', '爱姬', '修饰', '贵重', '哀怜', '刚烈', '吝啬', '护持', '爱我', '工巧']
22 ['爱慕', '汉苑', '双频', '芳姿', '披裘', '毓秀', '娇欲语', '娇憨', '相酬', '滴翠', '咏怀']
23 ['恩情', '恩爱', '佳期', '情怀', '深情', '怀抱', '红颜', '妄心', '易老', '间阻', '到底']
```

Fig. S13. Expanding lists of words through word2vec for vernacular texts.

Correlation between words occurrences in summaries and original books

Utilizing the generated dictionaries, we calculated the frequency of love-related words in the original text and subsequently computed the ratio of love. Following this, we filtered the fiction dataset, which overlaps with our description database (N = 336 fiction). The ratio of love in the description data was computed using the dictionary generated for vernacular texts, as the terminology closely aligns.

Next, we computed the correlation between the two indices. The estimated correlation was found to be significantly positive (r = 0.113, p < 0.0001***). The correlation is visualized in Figure S14.

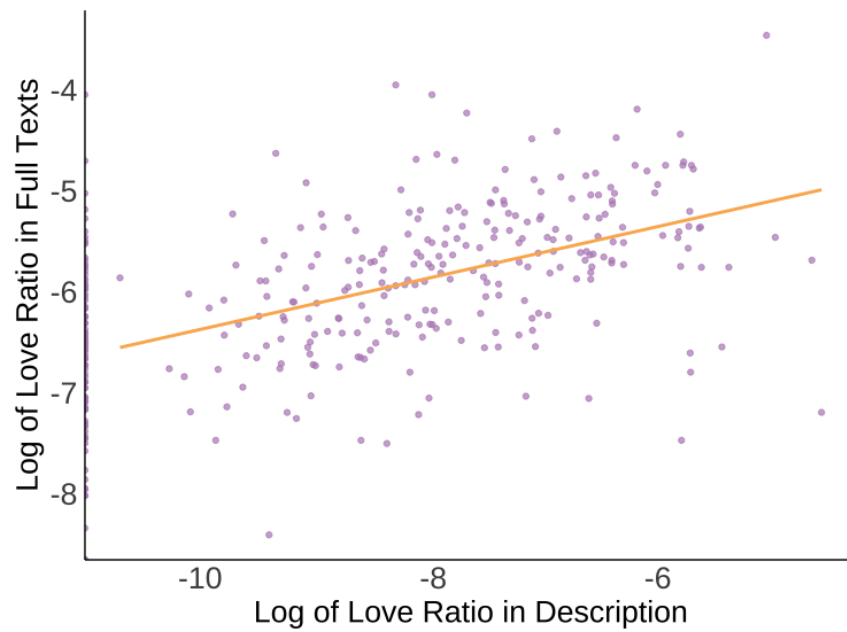Summaries for validated dictionaries are presented in Table S6.

Fig. S14. Correlation between ratio of love in full texts and ratio of love in summaries.

Table S6. **Summaries of validated values and examples.**

Ecology	Values	Proxy of concept	Examples in Chinese fiction	Pre-registration documents
Affluence	**Romantic love**	**Romantic love**	*Romance of the Western Chamber* (1300): "it is the story of a young couple consummating their love without parental approval, and has been seen both as a "lover's bible" and "potentially lethal…""	*https://osf.io/nmb73*
	Open-mindness	**Fictionality**	*Journey to the West* (1592): "it speaks of non-existent things as if they exist and speaks of falsehoods as if they are true. This is the foremost trickery. 西游记说无事真如有事，说假事真如真事。此第一秒手也[11]"	*https://osf.io/nmb73*
	Self development	**Learning**	*Journey to the West* (1592)*:* "Sun Wukong is a stone monkey born by absorbing the essence of heaven and earth, because he learned the law from Master Bodhi and attained Taoism, he is able to transform into 72 species, ride on the clouds… and he proudly calls himself the Monkey King 孙悟空是吸收天地精华而生的石猴，因为向菩提祖师学法而得道，能通地煞七十二变、乘斤斗云、使如意金箍棒，他骄傲地自称为齐天大圣"	*https://osf.io/v7yuw*
	Friendship	**Firendship**	*Lost On Journey* (2010): "The two go from being hostile to becoming friends, and Niu Geng's sincerity also changes Li Chenggong's indifferent character and evokes his return to true love. 两人从敌对到成为朋友，牛耿的真诚也改变李成功冷	*https://osf.io/yruzq*

---

[11] 英国国家图书馆藏《新刻全像批评西游记》，残存四卷二十回
 Newly Engraved Complete Annotated Edition of Journey to the West, preserved in the British Library, consisting of the remaining four volumes and twenty chapters.

			漠的性格，唤起李成功对真情的回归。"	
	Reciprocal cooperation	**Mutuality**	*Tie Mu Qian Zhuan* (1956): "Slowly, the friendship between Li Laodong and Fu Laogang became deeper and deeper, and the two helped each other and cooperated with each other. 黎老东和傅老刚的交情越来越深，两个人互帮互助，互相合作。"	*https://osf.io/v7yuw*
Harshness	**Honor**	**Honor-shame**	*Officialdom Unmasked* (1903): "Governor Huang was promoted to Commissioner, and while he was feeling proud, he was reported for a case he had handled before. Later, through connections with friends in the capital and after spending ten thousand taels of silver, the matter was finally settled. Huang remained as pleased as ever…黄知府升了道台，正在得意之时，因为前边所办的案子被人告发，后来通过京城中友人的疏通，又花了一万两银子，终于了结。黄依然春风得意。"	*https://osf.io/yruzq*
	Intensive Kinship	**Intensive kinship**	*Moment in Peking* (1939): "It tells the story of the three great families of Zeng, Yao and Niu in Beiping, from the Boxer Rebellion in 1901 to the War of Resistance against Japan, which lasted for more than thirty years. 讲述了北平曾、姚、牛三大家族从 1901 年义和团运动到抗日战争三十多年间的悲欢离合和恩怨情仇"	*https://osf.io/nmb73*
	Tightness	**Discipline**	*The Sun Also Rises* (2007): "The theme of the film is that in that era which extremely suppressed individual desires, it reflects "various extreme states of human nature and desires." 其表达的电影主题是，在那个极度压抑个人性欲的时代，每个人所反映出的人性欲望之中各种极致境界"	*https://osf.io/7j6zv*

	Decency	Puritanical morality	*Yingying's Biography* (790 CE): "But in the end, Zhang changes his heart, believing that Yingying is the most exceptional beauty in the world. He even cites the examples of the downfall of the Shang and Zhou dynasties due to their rulers' infatuations, proving that his virtue is insufficient to overcome such seductive evil, and he has no choice but to reluctantly give her up. 但张生终于变心，认为莺莺是天下之尤物，还搬出商纣周幽亡国的例子，证明自己德不足以胜妖孽，只好割爱。"	https://osf.io/7j6zv
Note: examples drawn from descriptions on Baidu Baike and Wikipedia.				

Analysis and results

Temporal modeling

We used ggplot2 in R studio to plot the relationship between Frequency (residuals of model 1) and year. It is done by *geom_smooth* function, with defaulted k terms.

Evolution from 618 CE to 2022 CE

We have set the limits of the x-axis, starting from 618 to 2022 (with the first sample being 690 CE). Summaries of GAM models results are reported in Table S7. The evolution figures are from Fig. S15 to S23.

Evolution from 1900 to 2022 CE. The evolution figures are from Fig. S24.
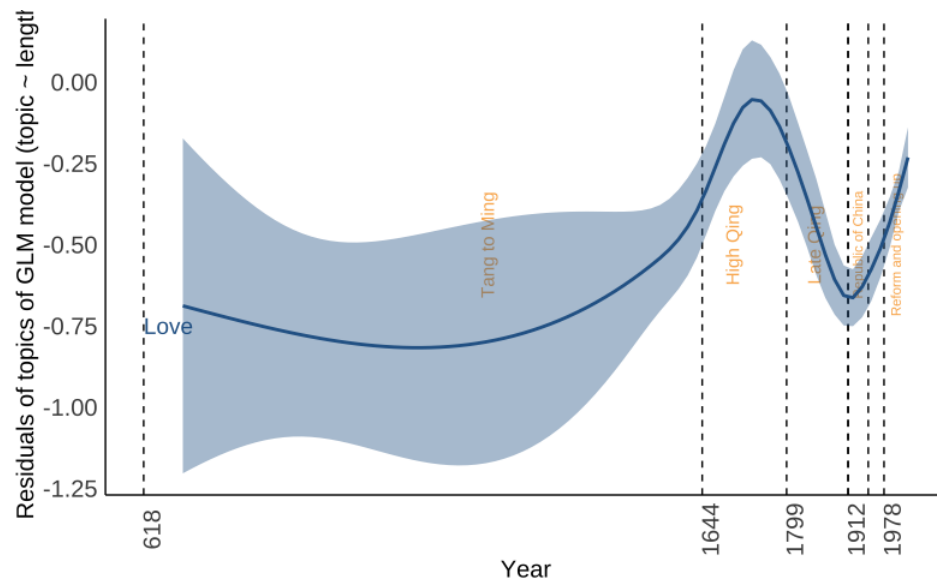
Fig. S15. Evolution of love from 690 to 2022.
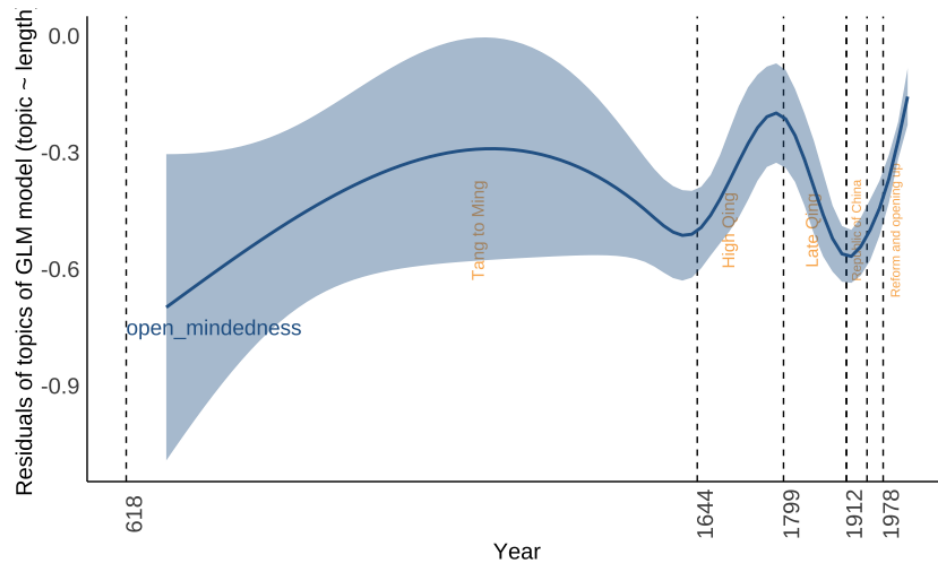
Fig. S16. Evolution of open-mindedness from 690 to 2022.
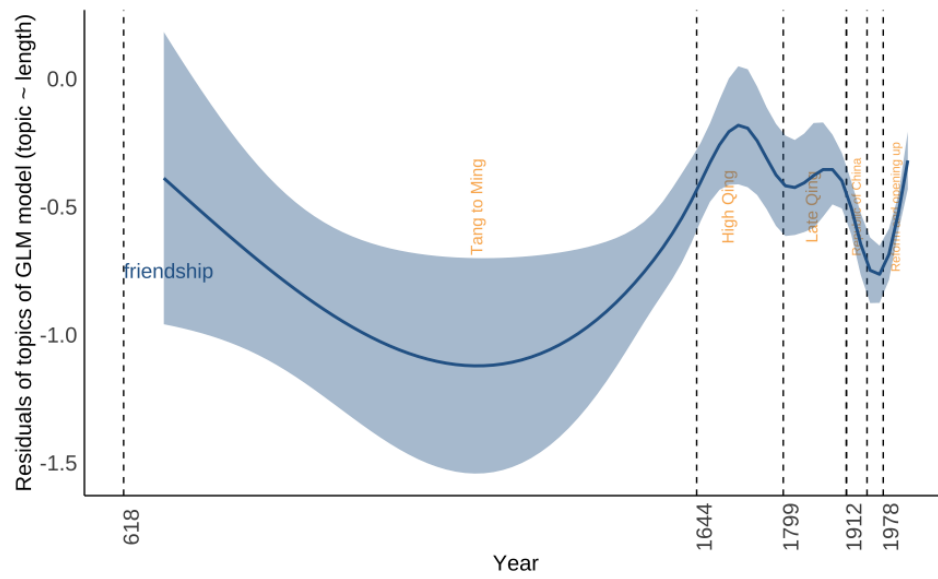
Fig. S17. Evolution of friendship from 690 to 2022.
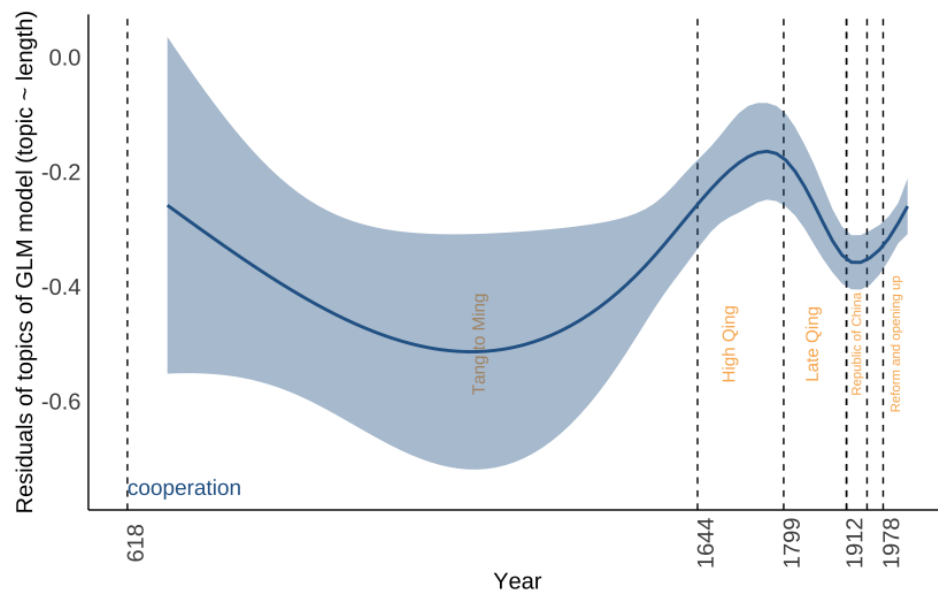
Fig. S18. Evolution of reciprocal cooperation from 690 to 2022.
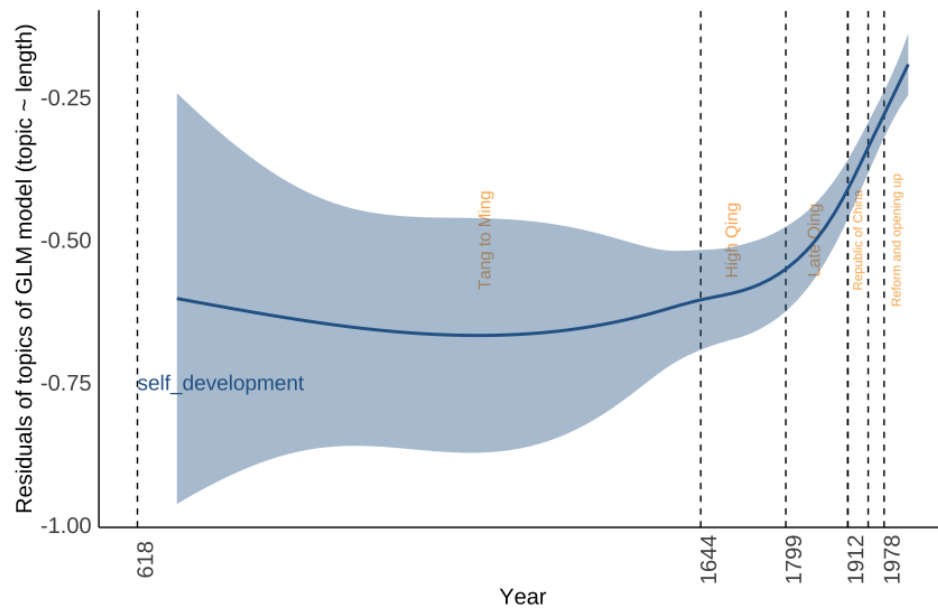
Fig. S19. Evolution of self-development from 690 to 2022.
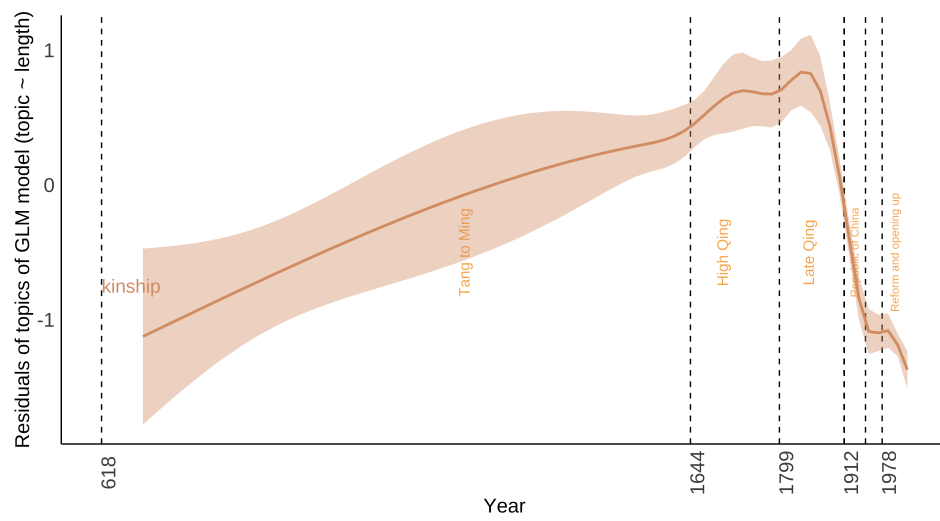
Fig. S20. Evolution of intensive kinship from 690 to 2022.
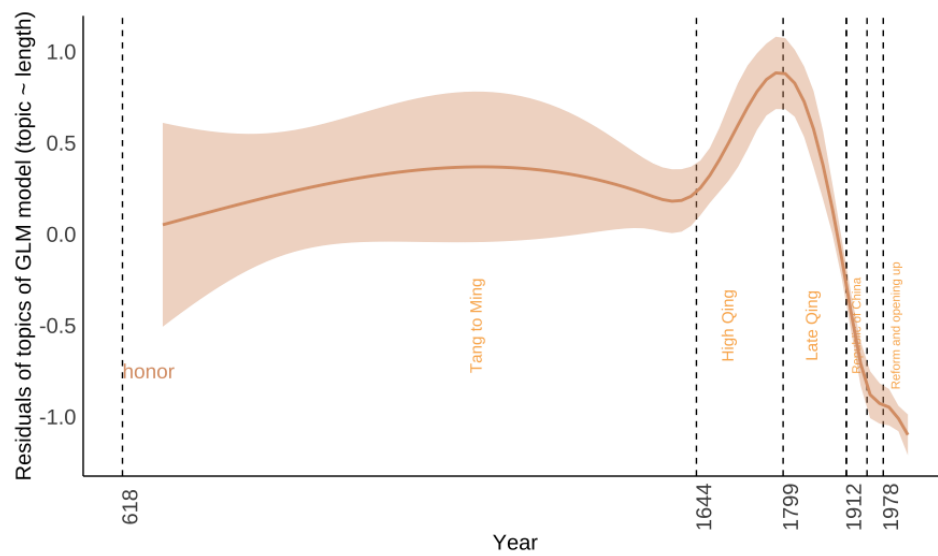
Fig. S21. Evolution of honor from 690 to 2022.
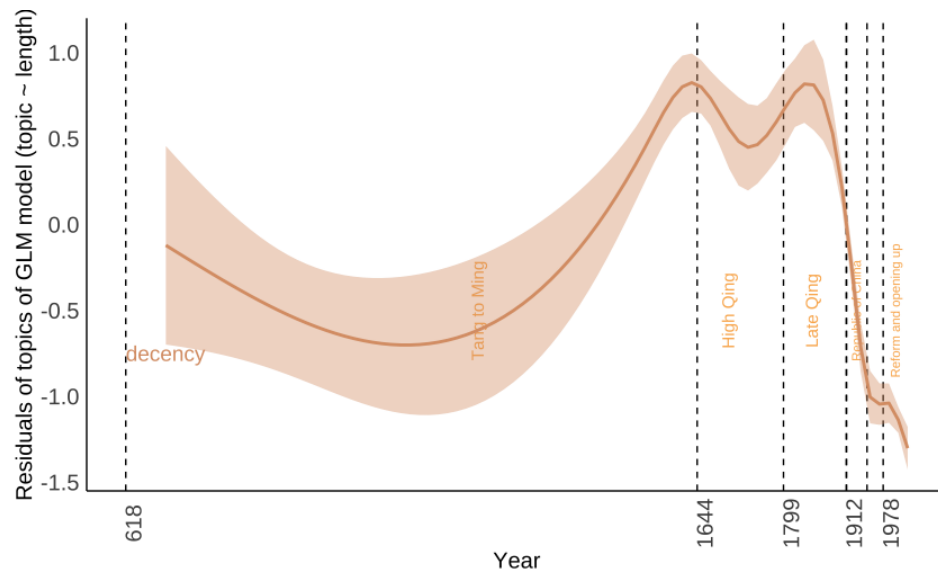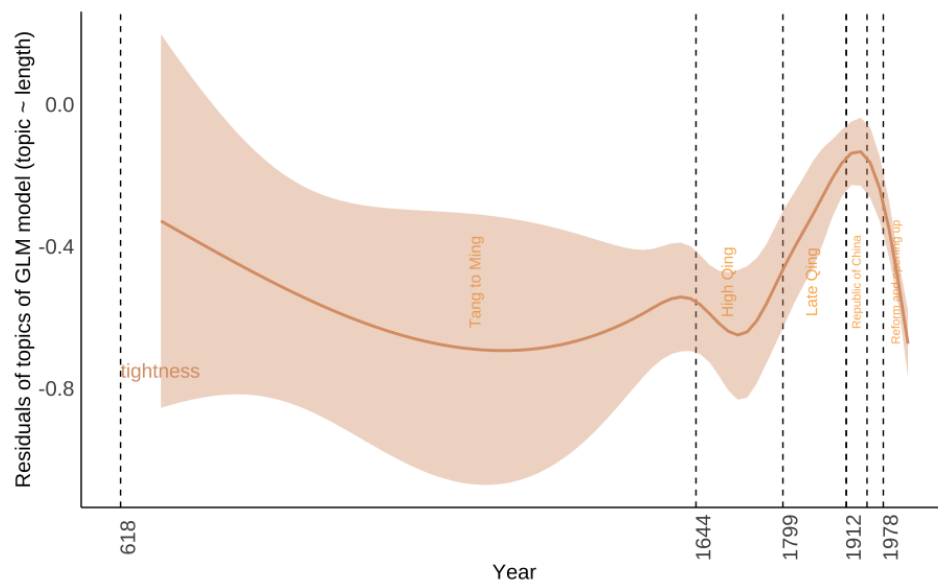
Fig. S22. Evolution of decency from 690 to 2022.

Fig. S23. Evolution of tightness from 690 to 2022.

Table S7 Summaries of GAM models for samples ranging from 690 to 2022.

Models	P value of s(Year)	edf	Deviance Explained
love ~ s(Year)	<0.001***	7.03	1.7%
open_mindedness ~ s(Year)	<0.001***	7.42	2%
friendship ~ s(Year)	<0.01**	8.244	0.69%
reciprocal cooperation ~ s(Year)	<0.01**	6.293	0.661%
self_development ~ s(Year)	<0.001***	3.788	1.94%
Intensive_kinship ~ s(Year)	<0.001***	8.032	14.4%
honor ~ s(Year)	<0.001***	7.929	14.1%
tightness ~ s(Year)	<0.001***	7.885	1.99%
decency ~ s(Year)	<0.001***	8.741	18.9%

Note: Total sample size = 4185
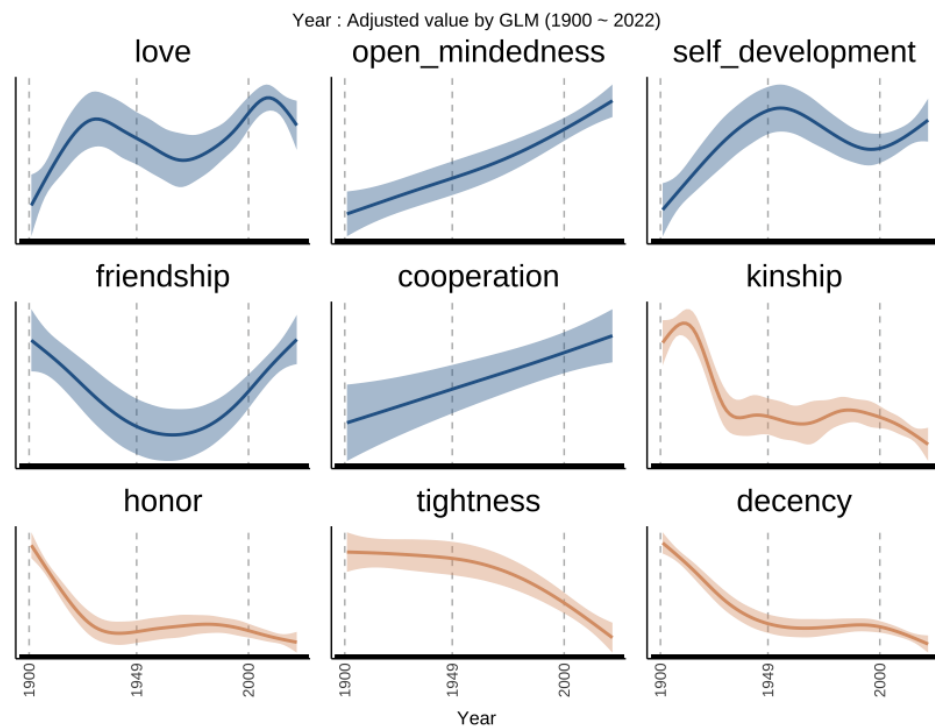
Fig. S24. Estimated variations of significance of domains in literary history from 1900 to 2022 China.

The modernization syndrome

Methods

We predict that, in typical eras with increasing living standards, people will shift their motivations to prioritize a variety of values such as *love*, *reciprocal cooperation*, and *self-development*, while giving less priority to present-oriented motivations, such as *intensive kinship*, *authority*, and *honor*. Consequently, we anticipate a modernization syndrome where increase of *love*, *open-mindedness*, *cooperation*, *friendship*, *self-development* are correlated with decline of *tightness*, *intensive kinship*, *honor*, and *decency*.

Through previous temporal modeling, we have gained initial impressions suggesting that certain domains co-vary over time. To statistically test this prediction, we used the predicted values of GAM models to run correlation tests.

Fit GAMs for each domain in R studio. We used *mgcv* packages in R studio to fit the following GAM models for each domain. Control variables are added for length of summary, source of summary and type of fiction:

$$\text{Frequency}_i \sim s(\text{Year}_i) + \beta_2 \cdot \text{Length}_i + \beta_3 \cdot \text{Source}_i + \beta_4 \cdot \text{Type}_i + \epsilon_i$$

Frequency is the occurrence of value dictionary words. The s(Year) indicates a smooth function of Year which is used to fit a non-linear relationship between Year and the outcome variable Frequency.

Extract the predicted term of s(Year). We utilized the *predict()* function to obtain the scores for the s(Year) term. This is to extract the predicted values of frequency changes over time, accounting for variations in length, source, and type of the fiction analyzed.

Plot the extracted terms with years. We plotted the predicted values against the year in Figure S25. This visualization allowed us to observe the predicted trajectory of the values over time, as estimated by the model.

Correlation test between domains. Subsequently, correlation analysis was performed on the predicted values for the nine values. This was accomplished using c*or()* in R studio, which generated a correlation matrix for the nine values.

Results

Figure 3b (main text) represents the matrix of correlation estimates between domains in response to the time changes. The mean correlation between nine domains is 0.3248. The only outlier among the modernization-affluent syndromes is *tightness*, exhibiting a positive correlation with *friendship* and *self-development*, negative correlation with *decency*,

intensive kinship and *honor*. This is because from 1800 to 1950, both *tightness*, *self-development*, and *friendship* increased, while *decency*, *intensive kinship*, and *honor* significantly declined during this period, thus exhibiting a negative correlation with *tightness*. We attribute this to the *tightness*'s unique historical evolution pattern, characterized by an inverted V-curve.

Are they statistically significant?

To determine the statistical significance of the modernization syndrome, we employed a bootstrap statistical procedure, detailed as follows:

Breaking the Year-Domain Link: We started by resampling the data within each of the nine domain columns, thereby disconnecting the association between the time and the domain estimates. This involves shuffling the data rows within each domain column independently, ensuring the temporal sequence is randomized.

Calculating Mean Correlation Scores: For each shuffled dataset, we calculated the mean correlation score across the nine domains to assess how closely related these domains are within the resampled data.

Repetition for Robustness: This shuffling and mean correlation score calculation process was repeated 10,000 times. This repetition generated a large set of mean correlation coefficients, providing a robust sample for analysis.

Analyzing the Results: We then analyzed the distribution of these 10,000 mean correlation scores to understand the variance and central tendencies.

The workflow sees Figure S26**.**

The results of variance and central tendencies are depicted in Figure S27. The null hypothesis posits that the correlation score between the values is equal to or less than 0, indicating no significant relationship between values. Under the null hypothesis, the sample distribution is expected to be centered around 0, as presented in Figure S27. Yet we observed a mean correlation of 0.3248. The calculated p-value associated with this observed correlation was found to be less than 0.0001. Given the extremely low p-value, we reject the null hypothesis that the correlation between the values is 0. This suggests that there exists a statistically significant modernization syndrome in our case.
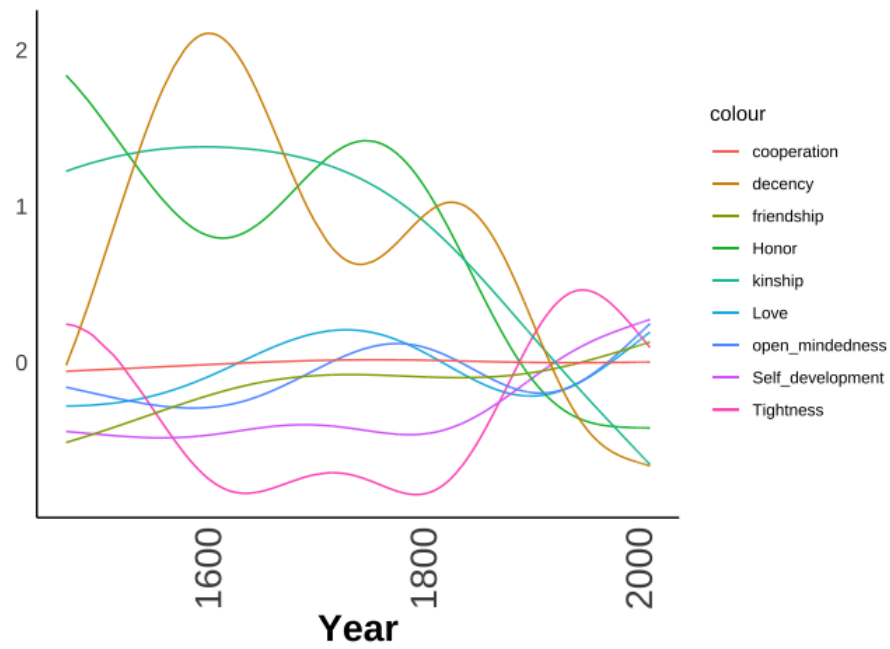
Fig. S25. Evolution of extracted terms from GAMs models with years from 1450 to 2022.
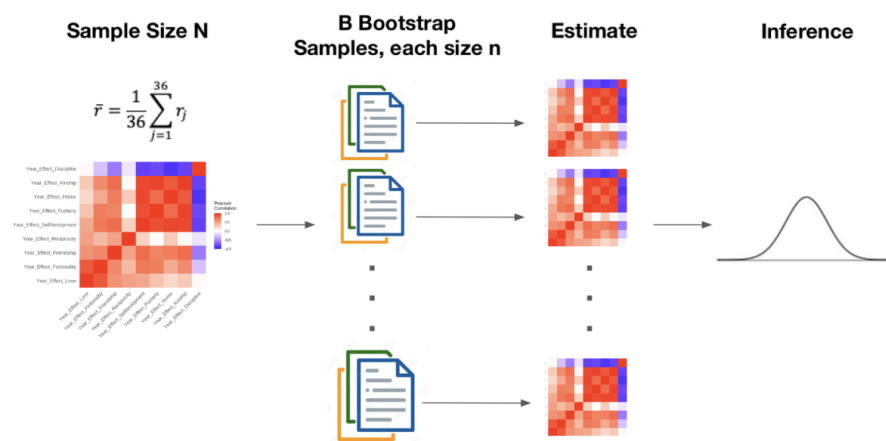
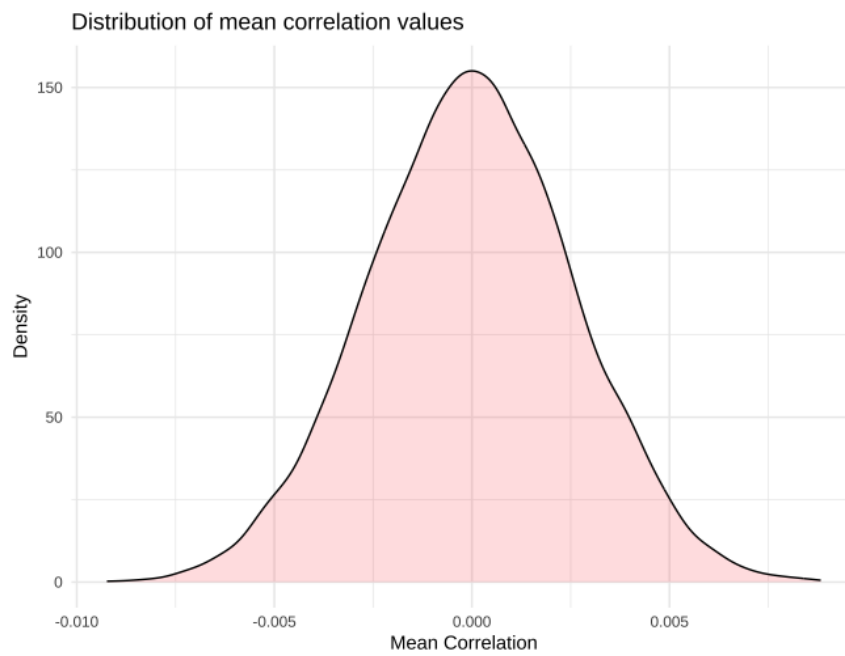**Fig. S26.** Procedure of bootstrapping analysis.

Fig. S27. The distribution of mean correlation of 10000 times of bootstrapping.

The association between economic development and variations of domains in history

Living standards in the long run

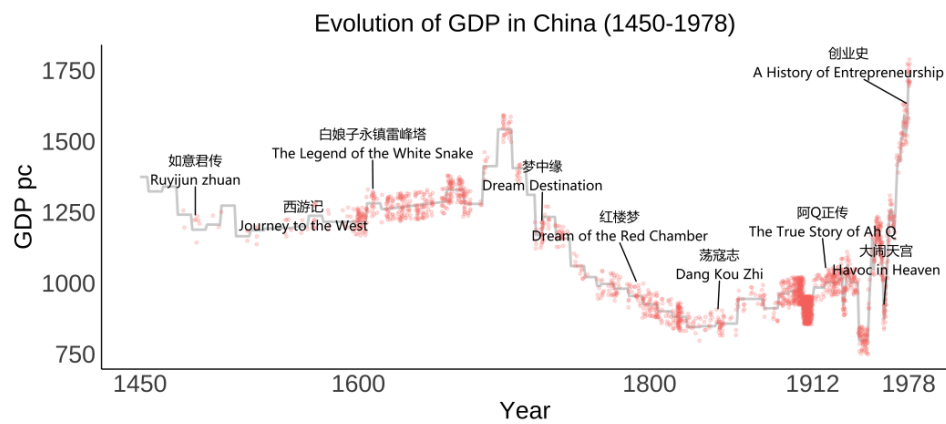The distribution of fiction samples alongside the evolution of GDP pc from 1450 to 1978 China are in Fig. S28 (more literary samples are added).

Fig. S28. The distribution of fiction samples alongside the evolution of GDP pc from 1450 to 1978 China (more literary samples are added).

Predictions and models

For each value, we make a prediction with the analysis model proposed in Table S8 (they are pre-registered in OSF before the analysis).

Table S8. Models. Main Prediction test with statistical models, modern Chinese fiction compare to ancient Chinese fiction

Domain	Prediction	Test $\text{Frequency\_domain}_i \sim \text{Poisson}(\lambda_i)$
Romantic love	H1: Higher levels of romantic love in fiction will be associated with an affluent environment.	$\begin{aligned} \log(\lambda_i) = {} & \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i \\ & + \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i \\ & + \beta_5 \cdot \text{Type}_i + \epsilon_i \end{aligned}$
Open_minded ness	H2: Higher levels of open-mindedness in fiction will be associated with an affluent environment.	$\begin{aligned} \log(\lambda_i) = {} & \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i \\ & + \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i \\ & + \beta_5 \cdot \text{Type}_i + \epsilon_i \end{aligned}$
Friendship	H3: Higher levels of friendship in fiction will be associated with an affluent environment.	$\begin{aligned} \log(\lambda_i) = {} & \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i \\ & + \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i \\ & + \beta_5 \cdot \text{Type}_i + \epsilon_i \end{aligned}$
Reciprocal Cooperation	H4: Higher levels of reciprocal cooperation in fiction will be associated with an affluent environment.	$\begin{aligned} \log(\lambda_i) = {} & \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i \\ & + \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i \\ & + \beta_5 \cdot \text{Type}_i + \epsilon_i \end{aligned}$
Self-development	H5: Higher levels of self-development in fiction will be associated with an affluent environment.	$\begin{aligned} \log(\lambda_i) = {} & \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i \\ & + \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i \\ & + \beta_5 \cdot \text{Type}_i + \epsilon_i \end{aligned}$

	H6:	
Tightness	Higher levels of tightness in fiction will be associated with a harsh environment.	$\log(\lambda_i) = \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i$ $+ \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i$ $+ \beta_5 \cdot \text{Type}_i + \epsilon_i$
Intensive kinship	H7: Higher levels of intensive kinship will be associated with a harsh and unpredictable environment.	$\log(\lambda_i) = \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i$ $+ \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i$ $+ \beta_5 \cdot \text{Type}_i + \epsilon_i$
Honor	H8: Higher levels of honor in fiction will be associated with a harsh and unpredictable environment.	$\log(\lambda_i) = \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i$ $+ \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i$ $+ \beta_5 \cdot \text{Type}_i + \epsilon_i$
Decency	H9: Higher levels of decency in fiction will be associated with a harsh and unpredictable environment.	$\log(\lambda_i) = \beta_0 + \beta_1 \cdot \text{Length}_i + \beta_2 \cdot \text{GDP}_i$ $+ \beta_3 \cdot \text{Year}_i + \beta_4 \cdot \text{Source}_i$ $+ \beta_5 \cdot \text{Type}_i + \epsilon_i$

Independent variables:

GDP: GDP pc; Log of GDP pc; ΔGDP pc; GDP pc_rate

$\varepsilon(P)$: Random error, with Poisson distribution

Dependent variables:

$\text{Frequency\_domain}_i$: the total number of words for love dictionary occurrence in a description

Control variables:

Length: Length of description for each fiction

Source: Wikipedia, Baidubake or Scholars.

Type: Textual or Video.

Year: The publication year of the fiction

Results

Here we present the full results for the main models of GDP pc, as well as estimation coefficient of logarithms of GDP pc and models of ΔGDP pc, and ΔGDP pc_rate.

Table S9. Results of regressions. Results of General linear regression models with Poisson distribution for the model

GDP pc models

Domains	β1 (L)	β2 (GDP)	β3 (Source_scholar)	β3 (Source_wiki)	β4 (Type_video)	β5 (Year)	AIC
love	9.377e-05*** (44.798)	**3.010e-05*** (6.958)**	-4.525e-01*** (-9.939)	-4.186e-01*** (-8.931)	1.317e-01* (2.492)	7.183e-05 (0.788)	13724
open_mindedness	1.013e-04*** (32.240)	**6.535e-05*** (10.043)**	-9.726e-01*** (-12.254)	-9.119e-02 (-1.237)	-6.702e-01*** (-8.208)	4.179e-04*** (2.818)	7600
self-development	1.204e-04*** (52.457)	**3.725e-05*** (6.341)**	-7.733e-01*** (-11.731)	-2.710e-01*** (-4.088)	-2.832e-01*** (-3.913)	1.313e-03*** (7.906)	8204.2
friendship	1.047e-04*** (64.927)	**4.000e-05*** (9.028)**	3.163e-01*** (7.958)	-8.647e-02 (-1.754)	-2.942e-02 (-0.528)	6.481e-07 (0.008)	15694
reciprocal	1.093e-04*** (30.441)	1.888e-05 (1.925)	2.173e-02 (0.241)	4.876e-02 (0.451)	-5.258e-02 (-0.433)	1.133e-04 (0.595)	4819.2

	cooperation						
tightness	1.079e-04*** (69.937)	**-1.047e-05*** **(-2.381)**	-5.758e-02 (-1.605)	-2.491e-01*** (-5.281)	-2.067e-01*** (-4.017)	9.524e-04*** (9.947)	14744
intensive kinship	9.434e-05*** (77.549)	**-3.833e-05*** **(-7.796)**	6.215e-01*** (23.015)	-6.013e-01*** (-12.459)	-2.545e-01*** (-4.851)	-2.782e-04*** (-4.993)	19397
honor	9.555e-05*** (69.854)	**-1.774e-05** **(-3.194)**	6.029e-01*** (19.101)	-6.100e-01*** (-11.137)	-1.813e-01** (-2.981)	-8.249e-04*** (-14.797)	15141
decency	9.384e-05*** (76.080)	**-3.161e-05*** **(-6.762)**	5.231e-01*** (19.448)	-5.759e-01*** (-12.484)	-2.149e-01*** (-4.259)	-3.847e-04*** (-7.155)	17398

Observation size = 4184.

Table S10. Results of regressions. Results of General linear regression models with Poisson distribution for the model

log_GDP models

Domains	β1 (L)	β2 (log_g d p)	β3 (Source_scholar)	β3 (Source_wiki)	β4 (Type_video)	β5 (Year)
love	9.341e-05***	**1.812e-01*****	-4.140e-01***	-4.313e-01***	7.387e-02	3.647e-05
	(44.555)	**(8.292)**	(-8.921)	(-9.212)	(0.402)	(0.402)
open-mindedness	1.011e-04***	**3.343e-01*****	-9.150e-01***	-1.080e-01	-7.114e-01***	3.696e-04*
	(32.124)	**(10.145)**	(-11.309)	(-1.473)	(-8.737)	(2.498)
self-development	1.203e-04***	**1.856e-01****	-7.444e-01***	-2.765e-01***	-3.083e-01***	1.295e-03***
	(52.354)	**(6.278)**	(-11.116)	(-4.180)	(-4.221)	(7.796)
friendship	1.045e-04***	**1.844e-01*****	3.394e-01***	-8.501e-02	-4.980e-02	2.462e-05
	(64.688)	**(8.228)**	(8.342)	(-1.729)	(-0.877)	(0.289)
reciprocal cooperation	1.095e-04***	4.964e-02	1.561e-02	6.321e-02	-2.355e-02	1.529e-04
	(30.483)	(1.013)	(0.170)	(0.585)	(-0.189)	(0.799)
tightness	1.078e-04***	-2.805e-02	-5.584e-02	-2.576e-01***	-2.201e-01***	9.278e-04***
	(69.712)	(-1.339)	(-1.531)	(-5.448)	(-4.160)	(9.765)
intensive kinship	9.465e-05***	**-1.653e-01*****	6.069e-01***	-6.063e-01***	-2.309e-01***	-3.106e-04***
	(77.544)	**(-7.599)**	(22.138)	(-12.537)	(-4.242)	(-5.650)

honor	9.580e-05***	**-9.874e-02***	5.902e-01***	-6.047e-01***	-1.459e-01*	-8.324e-04***
	(69.821)	**(-3.911)**	(18.449)	(-11.044)	(-2.318)	(-15.075)
decency	9.421e-05***	**-1.573e-01***	5.047e-01***	-5.726e-01***	-1.732e-01***	-4.040e-04***
	(76.142)	**(-7.474)**	(18.512)	(-12.403)	(-3.311)	(-7.598)

Table S11. Results of regressions. Results of General linear regression models with Poisson distribution for the model

ΔGDP pc capita (decade)

Domains	β1 (L)	β2 (ΔGDP)	β3 (Source_scholar)	β3 (Source_wiki)	β4 (Type_video)	β5 (Year)
love	9.352e-05***	**6.970e-05***	-4.582e-01***	-4.111e-01***	1.452e-01**	7.884e-05
	(45.628)	**(6.277)**	(-10.075)	(-8.780)	(2.753)	(0.860)
open-mindedness	1.014e-04***	**1.641e-04***	-9.777e-01***	-8.196e-02	-6.642e-01***	4.129e-04**
	(32.250)	**(9.795)**	(-12.336)	(-1.111)	(-8.108)	(2.774)
self-development	1.205e-04***	9.168e-05***	-7.766e-01***	-2.638e-01***	-2.762e-01***	1.314e-03***
	(52.490)	(6.081)	(-11.796)	(-3.981)	(-3.813)	(7.876)
friendship	1.047e-04***	**1.061e-04***	3.175e-01***	-8.473e-02	-3.301e-02	-1.896e-05
	(64.932)	**(9.358)**	(7.996)	(-1.719)	(-0.593)	(-0.222)

reciprocal cooperation	1.093e-04*** (30.440)	**5.189e-05*** **(2.069)**	2.339e-02 (0.259)	4.860e-02 (0.450)	-5.733e-02 (-0.472)	1.012e-04 (0.531)
tightness	1.080e-04*** (70.007)	**-3.199e-05*** **(-2.813)**	-5.967e-02 (-1.664)	-2.467e-01*** (-5.243)	-2.000e-01*** (-3.897)	9.671e-04*** (10.031)
intensive kinship	9.430e-05*** (77.559)	**-1.028e-04*** **(-8.008)**	6.198e-01*** (22.943)	-6.023e-01*** (-12.504)	-2.521e-01*** (-4.822)	-2.588e-04*** (-4.601)
honor	9.550e-05*** (69.855)	**-4.337e-05*** **(-3.014)**	6.037e-01*** (19.109)	-6.129e-01*** (-11.203)	-1.868e-01** (-3.078)	-8.207e-04*** (-14.617)
decency	9.383e-05*** (76.117)	**-8.824e-05*** **(-7.229)**	5.203e-01*** (19.341)	-5.745e-01*** (-12.483)	-2.073e-01*** (-4.123)	-3.665e-04*** (-6.755)

Table S12. Results of regressions. Results of General linear regression models with Poisson distribution for the model

ΔGDP pc decade growth rate (scale to annual)

Domains	β1 (L)	β2 (ΔGDP_rate)	β3 (Source_scholar)	β3 (Source_wiki)	β4 (Type_video)	β5 (Year)
love	9.357e-05***	**5.765e-02*****	-4.002e-01***	-4.250e-01***	7.439e-02	-4.553e-05
	(44.622)	**(7.841)**	(-8.509)	(-9.082)	(1.370)	(-0.488)
open-mindedness	1.014e-04***	**9.175e-02*****	-9.190e-01***	-8.282e-02	-6.659e-01***	2.995e-04
	(32.262)	**(8.340)**	(-11.286)	(-1.130)	(-8.141)	(1.948)
self-development	1.205e-04***	**5.938e-02*****	-7.294e-01***	-2.682e-01***	-3.030e-01***	1.185e-03***
	(52.404)	**(5.950)**	(-10.788)	(-4.060)	(-4.144)	(6.922)
friendship	1.048e-04***	**6.079e-02*****	3.578e-01***	-8.002e-02	-5.416e-02	-7.229e-05
	(64.924)	**(8.044)**	(8.624)	(-1.628)	(-0.949)	(-0.825)
reciprocal cooperation	1.097e-04***	1.246e-02	1.423e-02	6.907e-02	-1.250e-02	1.427e-04
	(30.594)	(0.755)	(0.153)	(0.641)	(-0.099)	(0.719)
tightness	1.077e-04***	**-3.079e-03**	-4.895e-02	-2.655e-01***	-2.386e-01***	9.164e-04***
	(69.628)	**(-0.439)**	(-1.320)	(-5.626)	(-4.485)	(9.267)
intensive kinship	9.451e-05***	**-5.413e-02*****	5.925e-01***	-6.119e-01***	-2.226e-01***	-2.247e-04***
	(77.580)	**(-7.574)**	(21.220)	(-12.683)	(-4.035)	(-3.893)

honor	9.574e-05***	**-3.572e-02***	5.768e-01***	-6.048e-01***	-1.271e-01*	-7.827e-04***
	(69.929)	**(-4.291)**	(17.683)	(-11.081)	(-1.989)	(-13.562)
decency	9.422e-05***	**-6.147e-02***	4.793e-01***	-5.679e-01***	-1.284e-01*	-3.039e-04***
	(76.322)	**(-8.834)**	(17.316)	(-12.352)	(-2.419)	(-5.434)

Study 2: Evolution of adaptations in history

Data collection

While gathering the names of fiction based on literary history books (refer to Study 1), we paid attention to information on adaptations of these fictional works. Due to cost manual coding, adaptations were only collected for pre-modern novels (before 1912), as well as adaptations of those adaptations, until no further adaptations were found. For instance, Lu Xun suggests that "*A Dream In A Pillow*" (circa 780 CE) is adapted from "*Anecdotes About Spirits and Immortals*" (circa 400 CE), and "*Handan Notes*" (1601 CE) is adapted from "*A Dream In A Pillow*." To record this information, each fiction was assigned a unique numeric ID. This unique ID facilitated the linking of two books in one row to record adaptation information.

Adaptation identification criteria. We collected fiction's adaptation according to the information provided by Lu Xun (Lu 2000) , Xiao Xiangkai *A bibliography of vernacular novels* (Jiangsu Academy of Social Sciences 1990), Wikipedia and Baidu Baike. If they stated that "xxx is based on xxx" or "xxx is adaptation, original, copy, blueprint, or source of xxx", then this is clear adaptation information. We will record the adaptation information, as well as collect summaries for the adaptation from Baidu Baike and Wikipedia (if available). We do not include them in this case, even if these fictions have the same name and that seem to have the same story.

Procedure of collection of adaptations. We collected the names of the adapted fictions according to adaptation information. Then we searched on Wikipedia and Baidu Baike to see if these adaptations have entries. If so, we extract their descriptions, summaries and publication date. If this adaptation also has adaptation, we repeat the procedure to collect another entry. We stopped when there was no more adaptation.

Description of adaptation dataset

According to the links we record during coding, we plot the network of adaptations in R studio using igraph (see Figure 5 in main text).

We are curious about why these labeled fictions are adapted most frequently in history (at least they are popular considering the quantity of adaptations). A descriptive content analysis is conducted through word cloud analysis (for all the summaries of original works, we remove stopwords, such as none, location, geographic names. Then occurrences of words are calculated). As Figure S29 show, the most frequently present words are around the topics of supernatural gods and demons (such as 妖怪 monsters, 人间 Mortal World, 收伏 subdue, 仙

人 immortals), kins (父母 parents, 姐妹 sister, 夫妇 couple, 父亲 father) , love (鸳鸯 lover, 爱情 love, 夫妻 couple, 成亲 get married), altruistic cooperation (rescue 救出, 侠义 xiayi, 救出 save, 侠客 knight, 联合 unit). This implies the significance of the literary topics of *love*, *intensive kinship*, *fantasy* in history.
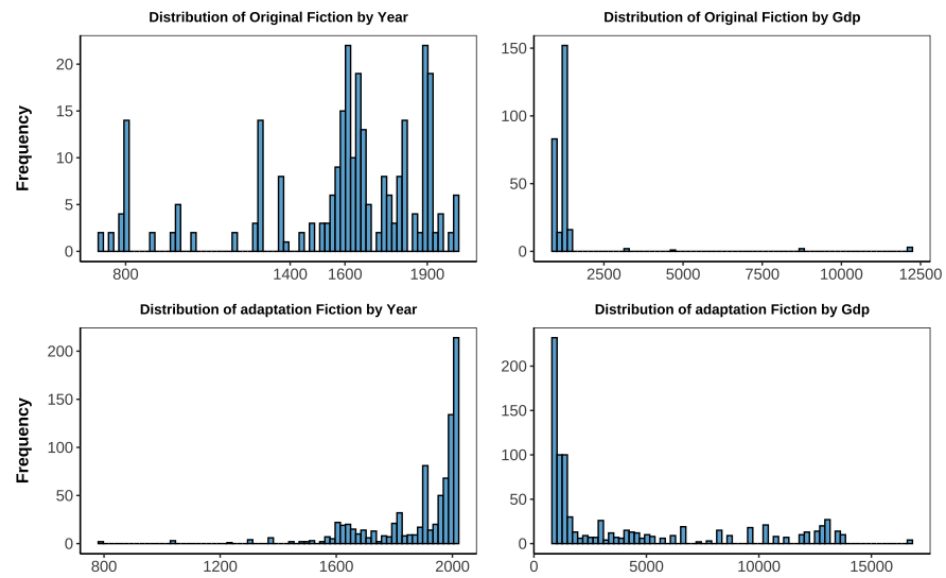
Fig. S29. The word cloud of the most popular original works. It is computed by an online platform https://www.weiciyun.com/ of word cloud in Chinese.
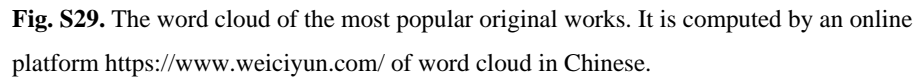
Fig. S30. Distribution of adaptations and original works by years and by GDP pc.

Most adaptations are created after the 1900s.

Model

We employ a general linear regression model as follows to test how changes in domain significance are associated with changes with the economic environment or media transformation. Three variables are computed: changes in value frequencies, changes in GDP pc (use logarithm of GDP pc), changes in length of the summary (as control variable), distance of time, changes of type. The model is as follow:

$$\Delta \text{Frequency}_i = \beta_0 + \beta_1 \cdot \Delta \text{length}_i + \beta_2 \cdot \Delta \text{GDP}_i + \beta_3 \cdot \Delta \text{Year}_i + \beta_4 \cdot \Delta \text{Type}_i + \epsilon_i$$

Results

Statistics results are reported in Table S13.

Table S13. Results of regressions. Results of Linear regression models for the model

Domain	β1(Δlength)	**β2(ΔGDP)**	β3(ΔType: text to video)	β4(ΔYear)
Romantic love	0.0001447 (13.781)	**0.1399048 (1.607)**	0.7008815*** (3.659)	0.0003480 (0.906)
Open-mindedness	8.653e-05*** (11.995)	**2.027e-01*** (3.391)**	-3.042e-01* (-2.312)	8.926e-04*** (3.381)
Friendship	2.823e-04*** (29.522)	**2.579e-01** (3.254)**	5.333e-01** (3.057)	-1.278e-03*** (-3.654)
Reciprocal cooperation	9.186e-05*** (25.182)	**-4.877e-02 (-1.613)**	-1.179e-01 (-1.772)	-2.165e-04 (-1.622)
Self-development	8.220e-05*** (20.393)	**1.313e-01*** (3.929)**	-1.120e-01 (-1.523)	7.783e-05 (0.528)
Tightness	3.564e-04*** (39.469)	**-2.252e-01** (-3.009)**	-3.754e-01* (-2.279)	-1.789e-03*** (-5.415)
Intensive Kinship	8.159e-04*** (43.773)	**1.186e-01 (0.768)**	4.343e-01 (1.277)	3.170e-03 (4.646)
Honor	7.703e-04*** (52.075)	**-8.566e-02 (-0.699)**	2.706e-01 (1.003)	1.908e-03*** (3.525)
Decency	8.060e-04*** (46.826)	**6.023e-02 (0.422)**	8.608e-01** (2.741)	2.229e-03*** (3.538)

We hypothesize that the effects of GDP per capita (GDP pc) on adaptations may vary over time. For example, much of the observed effect might be driven by adaptations from modern times, especially after 1980 CE, a period marked by significant global cultural transmission. To explore this, we construct period-specific indicators (as dummy variables) based on the publication dates of the adaptations and interact these indicators with GDP changes. Periods are calculated based on centuries. The model is written as:

$$\Delta\text{Frequency}_i = \beta_0 + \beta_1 \cdot \Delta\text{length}_i + \beta_2 \cdot (\text{Period} * \Delta\text{GDP}_i) + \beta_3 \cdot \Delta\text{Type}_i + \epsilon_i$$

The results for the interaction effects (β_2) are presented in S14. They indicate that the effect of GDP changes on value-related words varies across periods. For example, the positive influence of ΔGDP on the increase in love and friendship in adaptations is significantly observed in the 17th and 19th centuries. In contrast, the association between ΔGDP increases and the decline frequencies of kinship and honor in adaptations is significantly evident in the 19th century, but not for other periods. These results suggest that changes in economic conditions do not produce uniform patterns of change in adaptation content, highlighting the complexity of mechanisms at play during transmission chains. Data points from the 21st century is excluded due to high collinearity among the factors.

Table S14. Results of interaction effects for periods and GDP pc changes (reports of co-efficiency).

Periods* ΔGDP$_i$	Romantic love	Open-mindedness	Friendship	Reciprocal cooperation	Self development
1300-1399: ΔGDP	9.461e-01	2.132e+01	-5.963e+00	-9.067e+00	-1.108e+01
1400-1499: ΔGDP	**-3.961e+01***	-5.151e-01	6.393e+00	-2.855e-01	-5.183e-01
1500-1599: ΔGDP	4.110e-01	-2.807e+00	-9.837e+00	-3.184e+00	2.616e+00
1600-1699: ΔGDP	**2.512e+01*****	3.133e+00	**8.366e+00***	-1.512e+00	**-5.046e+00****
1700-1799: ΔGDP	-5.150e-01	1.008e+00	-2.277e+00	**-1.623e+00****	**-1.432e+00***
1800-1899: ΔGDP	**2.972e+00***	-7.293e-01	-3.093e-01	-3.232e-01	-5.876e-01
1900-1999: ΔGDP	1.273e-01	-2.554e-02	-1.172e-01	5.286e-02	5.188e-03

| 2000-2020: ΔGDP | NA | NA | NA | NA | NA |

Periods* ΔGDP$_i$	Tightness	Intensive kinship	Honor	Reciprocal cooperation	
1300-1399: ΔGDP	-8.808e+00	2.848e+01	6.071e+01	-8.904e+00	
1400-1499: ΔGDP	3.884e+00	-6.946e+00	3.322e+01	1.172e+01	
1500-1599: ΔGDP	**-1.458e+01***	-4.918e+00	-7.461e+00	2.460e+01	
1600-1699: ΔGDP	-5.546e+00	4.647e+00	7.085e+00	**1.988e+01****	
1700-1799: ΔGDP	**3.147e+00***	-3.448e+00	3.817e+00	**7.893e+00****	
1800-1899: ΔGDP	1.659e+00	**-5.987e+00****	**-4.445e+00***	2.048e+00	
1900-1999: ΔGDP	-9.537e-03	**-1.004e+00***	6.750e-02	-5.590e-01	
2000-2020: ΔGDP	NA	NA	NA	NA	

Study 3: Evolution of fiction in Hong Kong, Singapore and Taiwan

Data

Our data collection aimed to assemble comprehensive and representative information on modern novels and films from Hong Kong, Taiwan, Singapore, and mainland China, for potential use by other scholars. For analysis, we utilized data solely from Hong Kong, Taiwan, and Singapore.

Data collection procedure

Collecting novels

- We referenced Wikipedia pages listing novelists, playwrights, or screenwriters. Name lists can be checked in: *hk_novelists.xlsx*, *Taiwan_novelists.xlsx*, *singapore_writers.xlsx*, *mainland_novelists.xlsx.* The sources of Wikipedia are from English and Chinese Wikipedia. See Hong Kong novelists (https://en.wikipedia.org/wiki/Category:Hong_Kong_novelists); Taiwan male and woman novelists (https://en.wikipedia.org/wiki/Category:Taiwanese_male_novelists, https://en.wikipedia.org/wiki/Category:Taiwanese_women_novelists); Singapore novelist, screenwriter, Chinese writer: https://zh.wikipedia.org/zh-sg/Category:%E6%96%B0%E5%8A%A0%E5%9D%A1%E5%89%A7%E4%BD%9C%E5%AE%B6 https://zh.wikipedia.org/zh-sg/Category:%E6%96%B0%E5%8A%A0%E5%9D%A1%E5%8D%8E%E4%BA%BA%E4%BD%9C%E5%AE%B6 https://zh.wikipedia.org/zh-sg/Category:%E6%96%B0%E5%8A%A0%E5%9D%A1%E5%B0%8F%E8%AF%B4%E5%AE%B6 https://zh.wikipedia.org/zh-sg/Category:%E6%96%B0%E5%8A%A0%E5%9D%A1%E7%BC%96%E5%89%A7

- We checked the Baidu Baike website to see if there were entries about these writers, and if so, we collected the URLs link for each author (figure S31).

- For each available author page, we extracted the information about the author's publication: especially the urls links of the fictions (figure S32).

- Based on the URLs link of each fiction, we extracted the description (the first paragraph of the entry site and the main story, usually under some specific tags); and date of publication (the first recorded date of publication for each work). See the following figure (Fig. S33).

Get the list of writers from Wikipedia, for example here, we search to get the list of writers in Hong Kong.

Extract names and links of author

Get the final info with names and Baidu Baike links

Manually check the entries in Baidu

Fig. S31. Find author page in Baidu Baike

For each author page's html

Get the works of "Major works" of each author

Save the html files in corresponding folders

Fig. S32. Find fiction page in Baidu Baike

Get the works after "Published books"

Obtain detailed information about the work

Store information in an excel spreadsheet

Fig. S33. Extract fiction summaries and year based on fiction page

Collecting films

Using the Wikidata Query Service, we queried films (Q11424) from Hong Kong Q8646, Taiwan Q865, and Singapore Q334. We queried information such as movie name, wiki link, and earliest release year. Descriptions were obtained based on the Wikipedia page of each movie, focusing on the first paragraph and main story (figure S35).

Exclusion criteria. We got some irrelevant information during data collection. We manually remove:

Non-fiction entries: people (such as actors, directors), places (such as Guangzhou, Chongqing…), publishing house (such as People's Publishing House 人民出版社) and literary awards (such as Mao Dun Literature Prize 茅盾文学奖).

Works before 1900.

Duplicate work: For some famous authors (e.g. Jin Yong 金庸), we can get their collections (containing the main works of the author) in "Published Books", and we can also get individual information for each work. Then we keep only the information about the works in the collection.

NB: Our list from Wikipedia contains not only novelists, but also playwrights, screenwriters, and directors, so our collection might also include some TV series, movies and anime in this data.

The complete code can be found at:

https://colab.research.google.com/drive/1s_ocvZuRvLzsxRuZYE6_b2yluJEg9oiZ?usp=sharing

Fig. S34. Query on Wikidata obtaining link of films.

Fig. S35. Collecting plot summaries and descriptions from page of fiction on Wikipedia

Description of Fiction Data for the four regions

Our database contains a total number of works N = 2191 (from Hong Kong) + 1005 (from Taiwan) + 142 (from Singapore) + 2895 (from Mainland China) = 6233

All the data is stored in the csv file: *df_film_novel_all.csv* on OSF platform.

Gdp per capita for three regions

We use 2020 Maddison project data also for the three regions (Bolt and Van Zanden 2020). For Singapore: from 1900 to 1939, every point per year; from 1950 to 2018, every point per year. For Taiwan: from 1901 to 1940, every point per year; from 1950 to 2018, every point per year. For Hong Kong: benchmark year 1913, Gdp per capita 2354$. From 1950 to 2018, every year one data point.

Description of fiction data and GDP pc

The following figure shows the evolution of GDP pc of the three regions with fiction data points distributed around. Most fiction is distributed in Hong Kong and Taiwan. The total number of fictions from these three regions is 3338.
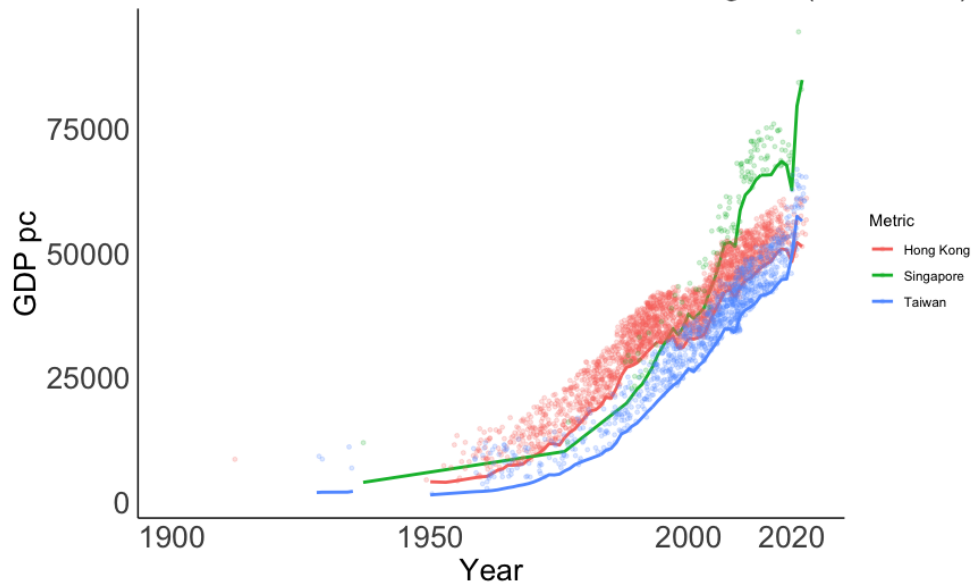
Fig. S36. Evolution of fiction and GDP pc for three regions from 1900 to 2023. GDP pc source: **(Bolt and Van Zanden 2020)** and World Bank Data for 2020, 2021, 2022, 2023.
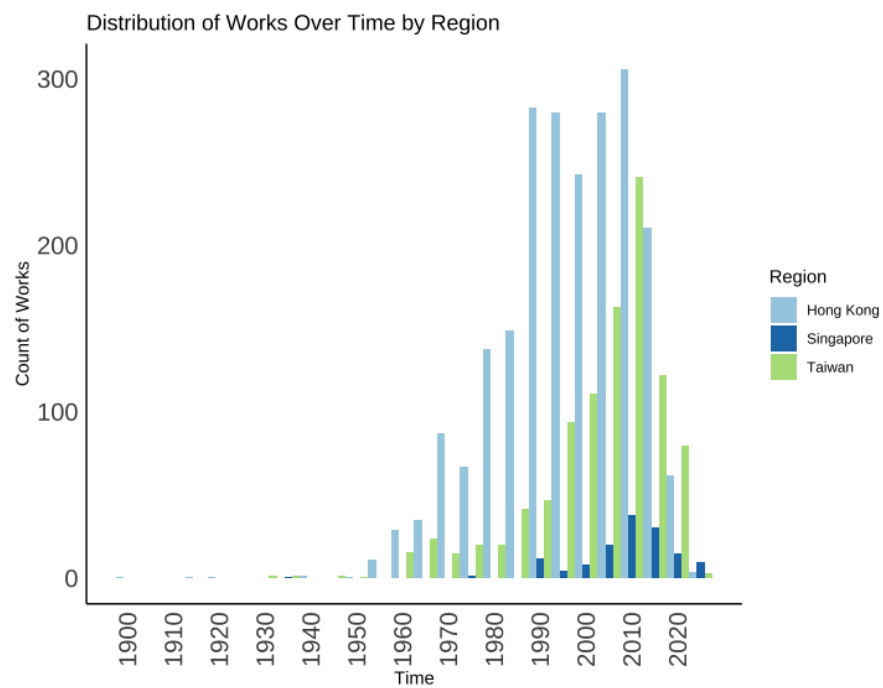
Fig. S37. Evolution of fiction by regions and years

Results

Evolution of nine domains in history in three regions

We replicate the same methodology of temporal evolution of nine domains for three regions. Because there are not enough data points before the 1960s, we formulate our GAM temporal analysis for the time span from 1950 to 2023.
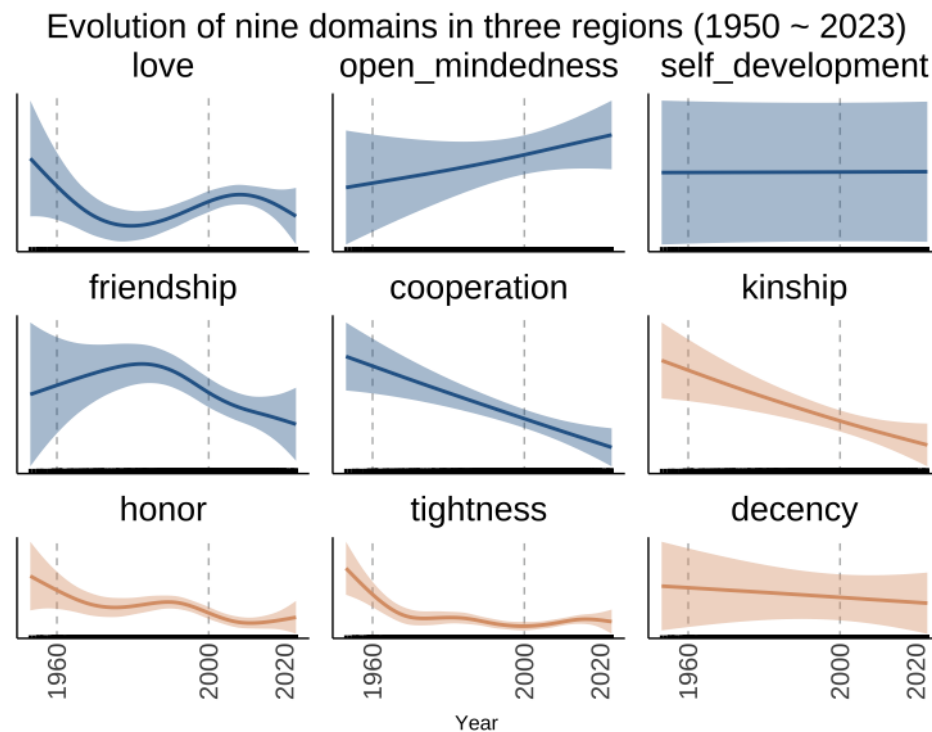
Fig. S38. Estimated changing of significance of domains in literary history from 1950 to 2023 in Hong Kong, Taiwan and Singapore.

We consider these evolution patterns might reflect a different evolution pattern compared to mainland China. An explanation might be that our measurements targeted *interdependence*, as these three regions simulate a hyper-modern (compared to mainland) future. It's possible that with economic growth, individuals' independence increases, leading to a decline in related values, such as mutual aid or reciprocity. For other domains with no apparent changes, such as *decency* or *self-development*, we attribute this to the relatively lower frequency of these themes within these three regions of fiction.

Association between economic development and domains evolution

We replicate the main hypothesis and method used in the main study. In the model here, the type of the fiction, length of the summaries, and regions from where the fiction is published are controlled. We did not add Year into the model, because it has collinearity issue (after

adding year as control variable, the GVIF_year = 18.0602, which suggests a high collinearity problem).

The results are reported in the following table and Figure S39.
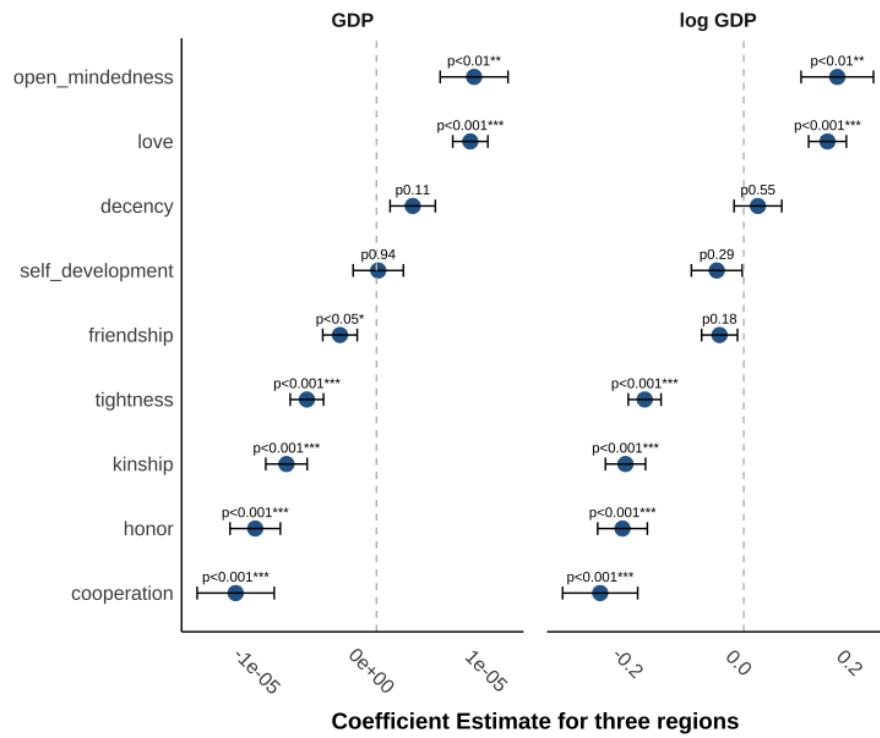
Fig. S39. The estimated coefficient of GDP pc and log of GDP pc with frequencies of domain words.

Table S15. Results of regressions. Results of General Linear Regression models for the GDP model of three regions

GDP pc model

Domain	β1(Length)	β2(Gdp)	β3(type_novel)	β4(Regions_sing)	β4(Regions_taiwan)
Romantic love	7.301e-05*** (30.473)	**8.179e-06*** (5.346)**	2.283e-01*** (5.833)	-3.979e-01*** (-3.857)	-3.146e-01*** (-7.087)
Open-mindedness	6.668e-05*** (13.614)	**8.510e-06*** (2.878)**	1.758e-01* (2.325)	-4.387e-01* (-2.065)	3.645e-02 (0.454)
Friendship	8.525e-05*** (39.253)	**-3.182e-06* (-2.117)**	-3.313e-01*** (-7.547)	3.227e-01*** (3.448)	-3.276e-01*** (-6.883)
Reciprocal cooperation	9.185e-05*** (18.020)	**-1.228e-05*** (-3.655)**	-5.905e-01*** (-5.482)	1.297e-01 (0.522)	-6.242e-01*** (-5.291)
Self-development	8.233e-05*** (28.063)	**1.556e-07 (0.071)**	3.818e-02 (0.639)	3.879e-01** (3.104)	-2.318e-01*** (-3.486)
Tightness	8.981e-05*** (48.708)	**-6.069e-06*** (-4.176)**	-2.548e-01*** (-6.010)	6.757e-01*** (8.204)	-2.626e-01*** (-5.684)
Intensive Kinship	8.838e-05*** (33.505)	**-7.848e-06*** (-4.345)**	-6.278e-01*** (-11.022)	-1.541e-01 (-1.041)	-3.009e-01*** (-5.252)
Honor	9.287e-05***	**-1.057e-05*****	-6.282e-01*** (-8.879)	-1.213e-01 (-0.682)	-6.019e-01*** (-7.901)

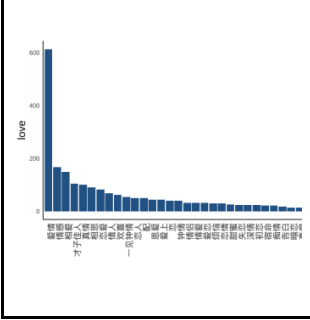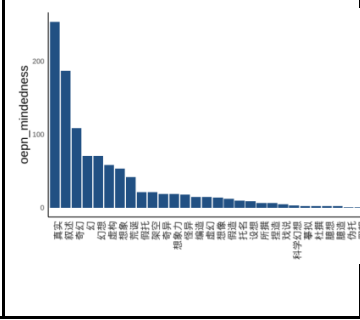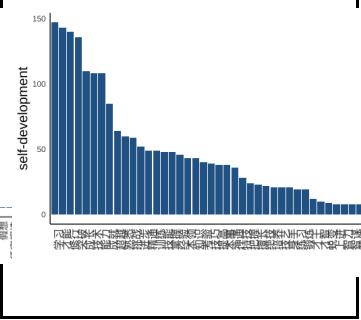|					
	(28.630)	**(-4.811)**			
Decency	8.924e-05*** (33.514)	**3.158e-06 (1.600)**	-4.824e-01*** (-8.295)	-2.701e-01 (-1.849)	-3.379e-01*** (-5.531)

Table 16. Results of regressions. Results of General Linear Regression models for the Log_GDP model of three regions

LOG_GDP model

Domain	β1(Length)	β2(log_GDP)	β3(type_novel)	β4(Regions_sing)	β4(Regions_taiwan)
Romantic love	7.266e-05*** (30.401)	**1.510e-01*** (4.444)**	2.318e-01*** (5.909)	-3.079e-01** (-3.091)	-3.165e-01*** (-7.115)
Open-mindedness	6.634e-05*** (13.585)	**1.687e-01*** (2.585)**	1.774e-01* (2.340)	-3.492e-01 (-1.697)	3.674e-02 (0.457)
Friendship	8.538e-05*** (39.295)	**-4.366e-02 (-1.350)**	-3.315e-01*** (-7.559)	2.781e-01** (3.104)	-3.278e-01*** (-6.882)
Reciprocal cooperation	9.261e-05*** (18.065)	**-2.589e-01*** (-3.824)**	-6.015e-01*** (-5.595)	1.687e-02 (0.070)	-6.385e-01*** (-5.414)
Self-development	8.240e-05*** (28.024)	-4.861e-02 (-1.062)	3.612e-02 (0.604)	4.183e-01*** (3.514)	-2.373e-01*** (-3.563)
Tightness	9.033e-05*** (48.753)	**-1.787e-01*** (-6.054)**	-2.607e-01*** (-6.152)	6.498e-01*** (8.300)	-2.739e-01*** (-5.924)
Intensive Kinship	8.901e-05*** (33.556)	**-2.133e-01*** (-5.901)**	-6.362e-01*** (-11.171)	-2.001e-01 (-1.388)	-3.153e-01*** (-5.500)
Honor	9.352e-05***	**-2.186e-01*****	-6.375e-01*** (-9.023)	-2.236e-01 (-1.289)	-6.130e-01*** (-8.049)

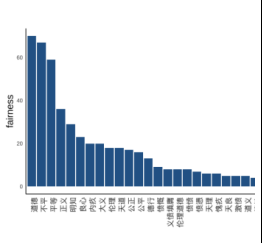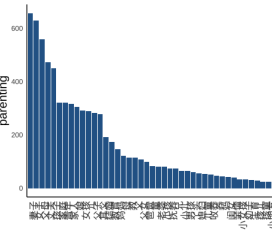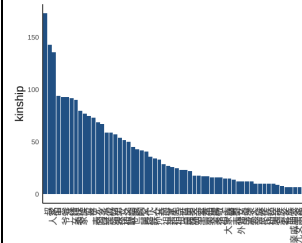|					
	(28.687)	**(-4.880)**			
Decency	8.913e-05***	2.556e-02	-4.825e-01***	-2.132e-01	-3.390e-01***
	(33.456)	(0.595)	(-8.283)	(-1.509)	(-5.538)

Appendix 1: Top frequencies of words of each domain

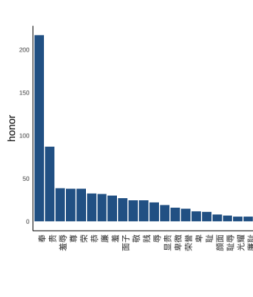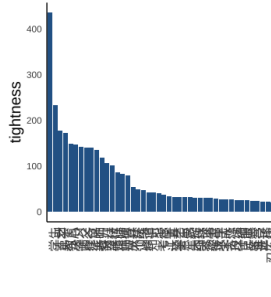**Table appendix 1.** Top frequencies of words for 15 values.
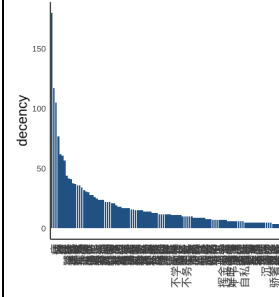
Love	*Open_mindedness*	*Self-development*
爱情 romantic love 614	真实 truth 254	学习 learn 147
情感 affect 168	叙述 storytelling187	才能 ability/talent 143
相爱 fall in love 150	奇幻 fantasy109	修行 practice (spiritual) 140
才子佳人 scholar-beauty 105	幻 unreal/imaginary 71	突破 breakthrough 136
真情 true feelings 100	幻想 fantasy 71	成长 growing/develop 100
相思 lovesickness 90	虚构 imaginary/fictional 58	技术 skill/ technique 108
恋爱 love 83	想象 Imagination 53	能力 ability 108
情人 lover 63	荒诞 odd 42	成就 achievement 85
一见钟情 love at first sight 54	假托 feign 21	超越 transcend/exceed64

Reciprocal Cooperation	*Friendship*	*Caring (did not pass*

		validity test)
合作 cooperation 258	朋友 friend 413	兄弟 brothers 755
相助 to help one another 206	信任 trust100	救 save 571
彼此 each other 161	同行 peers 96	朋友 friends 413
互相 each other 160	友人 friends 87	帮助 help 412
相与 mutual aid 37	友情 friendship 77	好友 friends 268
互动 interaction 22	结交 make friends73	相助 help 206
协同 collaboration 18	知己 best fiends 67	同学 peers 155
协力 cooperation17	友 friend 64	老师 teachers 136
相帮 to help one another10	诚 honest/sincere 38	支持 surpport 112

Fairness (did not pass validity test)	*Parenting* (did not pass validity test)	*Intensive Kinship*
道德 morality 70	妻子 wife 657	叔 uncle 173
不平 unfair 67	女子 women 629	人家 folk 143
平等 equality 59	父母 parents 560	伯 uncle 136
正义 justice 36	丈夫 husband 472	忠 loyal 94
明知 knowingly 29	孩子 child 450	爷爷 grandfather 93
良心 conscience 23	家庭 family 321	桃 ancestral hall 93
内疚 guilty 20	青年 youth 320	女婿 son in law 92
大义 righteousness 20	家人 family 316	表妹 cousin 90
伦理 ethics 18	娘 mom 305	家族 clan/family 80

Honor	*Tightness*	*Decency*
奉 to respect 217	学生 students 436	病 Disease
贵 noble 87	计划 planning 233	妓 prostitute

羞辱 humiliation39	弟子 disciple 178	奸 treacherous
尊 respect 38	教育 education 173	贞 Chastity
荣 honor 38	决心 determination 149	疾 Disease
恭 respectful 33	学习 learn 147	淫 kinky
廉 cheap 32	师父 master 142	饕餮 gluttony
羞 shame 30	修行 cultivate 140	淫乱 promiscuous
面子 face 27	努力 effortness 119	诱惑 temptation

Authority (did not pass validity test)	Punishment (did not pass validity test)	Self-cultivation (did not pass validity test)
官 official 393	惩罚 punishment 53	读书 read/learning at school288
身份 identity/status 263	谴责 reprimand 49	学 learn 157
战斗 fighting 209	指责 accuse 46	学习 study 147

战争 war 198	怨恨 resentment 41	修行 practice(spiritual) 140
出身 born social status175	惩治 punish 38	努力 effort 119
势力 power 159	斥责 scold 38	成长 growing up 110
矛盾 contradiction 138	治罪 punish28	修炼 cultivate 102
命令 command/order109	报应 retribution 26	自治 autonomy 50
反抗 resist 105	责备 blame 25	修道 monasticism 48

SI References

Ashton, Michael C., and Kibeom Lee. 2009. 'The HEXACO–60: A Short Measure of the Major Dimensions of Personality'. *Journal of Personality Assessment* 91(4):340–45. doi: 10.1080/00223890902935878.

Ashton, Michael C., Sampo V. Paunonen, Edward Helmes, and Douglas N. Jackson. 1998. 'Kin Altruism, Reciprocal Altruism, and the Big Five Personality Factors'. *Evolution and Human Behavior* 19(4):243–55. doi: 10.1016/S1090-5138(98)00009-9.

Baumard, Nicolas, Jean-Baptiste André, and Dan Sperber. 2013. 'A Mutualistic Approach to Morality: The Evolution of Fairness by Partner Choice'. *Behavioral and Brain Sciences* 36(1):59–78.

Bjorklund, David F. 2022. 'Children's Evolved Learning Abilities and Their Implications for Education'. *Educational Psychology Review* 34(4):2243–73. doi: 10.1007/s10648-022-09688-z.

Bode, Adam, and Geoff Kushnick. 2021. 'Proximate and Ultimate Perspectives on Romantic Love'. *Frontiers in Psychology* 12:573123.

Bolt, Jutta, and Jan Luiten Van Zanden. 2020. 'Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update'. *Maddison-Project Working Paper WP-15, University of Groningen, Groningen, The Netherlands*.

Chen, Qishan, Yurou Kong, Wenyang Gao, and Lei Mo. 2018. 'Effects of Socioeconomic Status, Parent–Child Relationship, and Learning Motivation on Reading Ability'. *Frontiers in Psychology* 9. doi: 10.3389/fpsyg.2018.01297.

Chen, Sijing, and Shasha Yang. 2024. 'Longitudinal Changes in Chinese Prosociality'. *Personality and Social Psychology Bulletin* 01461672231225367.

Chua, Roy Y. J., Kenneth G. Huang, and Mengzi Jin. 2019. 'Mapping Cultural Tightness and Its Links to Innovation, Urbanization, and Happiness across 31 Provinces in China'. *Proceedings of the National Academy of Sciences* 116(14):6720–25. doi: 10.1073/pnas.1815723116.

Cui, Wei. 2017. 'Social Trust, Institution, and Economic Growth: Evidence from China'. *Emerging Markets Finance and Trade* 53(6):1243–61. doi: 10.1080/1540496X.2016.1264299.

Cummins, Denise D. 2016. 'Status and Dominance Hierarchies'. Pp. 1–10 in *Encyclopedia of Evolutionary Psychological Science*, edited by V. Weekes-Shackelford, T. K. Shackelford, and V. A. Weekes-Shackelford. Cham: Springer International Publishing.

Dubourg, Edgar, and Nicolas Baumard. 2021. 'Why Imaginary Worlds? The Psychological Foundations and Cultural Evolution of Fictions with Imaginary Worlds'. *The Behavioral and Brain Sciences* 45:e276. doi: 10.1017/S0140525X21000923.

Dunfield, Kristen, Valerie A. Kuhlmeier, Laura O'Connell, and Elizabeth Kelley. 2011. 'Examining the Diversity of Prosocial Behavior: Helping, Sharing, and Comforting in Infancy'. *Infancy* 16(3):227–47. doi: 10.1111/j.1532-7078.2010.00041.x.

Durkee, Patrick K., Aaron W. Lukaszewski, and David M. Buss. 2019. 'Pride and Shame: Key Components of a Culturally Universal Status Management System'. *Evolution and Human Behavior* 40(5):470–78. doi: 10.1016/j.evolhumbehav.2019.06.004.

Eisenberg, Nancy, Natalie D. Eggum, and Tracy L. Spinrad. 2015. 'The Development of Prosocial Behavior'. P. 0 in *The Oxford Handbook of Prosocial Behavior*, edited by D. A. Schroeder and W. G. Graziano. Oxford University Press.

Eisenberg, Nancy, and Paul A. Miller. 1987. 'The Relation of Empathy to Prosocial and Related Behaviors'. *Psychological Bulletin* 101(1):91–119. doi: 10.1037/0033-2909.101.1.91.

Elman, Benjamin A. 2000. *A Cultural History of Civil Examinations in Late Imperial China*.

Feinberg, Matthew, Ray Fang, Shi Liu, and Kaiping Peng. 2018. 'A World of Blame to Go Around: Cross-Cultural Determinants of Responsibility and Punishment Judgments'. *Personality and Social Psychology Bulletin* 45:014616721879463. doi: 10.1177/0146167218794631.

Fisher, Helen E., Arthur Aron, and Lucy L. Brown. 2006. 'Romantic Love: A Mammalian Brain System for Mate Choice'. *Philosophical Transactions of the Royal Society B: Biological Sciences*. doi: 10.1098/rstb.2006.1938.

Fitouchi, Léo, Jean-Baptiste André, and Nicolas Baumard. 2021. 'Are There Really so Many Moral Emotions? Carving Morality at Its Functional Joints'.

Fitouchi, Léo, Jean-Baptiste André, and Nicolas Baumard. 2023. 'Moral Disciplining: The Cognitive and Evolutionary Foundations of Puritanical Morality'. *Behavioral and Brain Sciences* 46:e293. doi: 10.1017/S0140525X22002047.

Fletcher, Garth J. O., Jeffry A. Simpson, Lorne Campbell, and Nickola C. Overall. 2015. 'Pair-Bonding, Romantic Love, and Evolution: The Curious Case of Homo Sapiens'. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 10(1):20–36. doi: 10.1177/1745691614561683.

Frimer, Jeremy A., Reihane Boghrati, Jonathan Haidt, Jesse Graham, and Morteza Dehgani. 2019. 'Moral Foundations Dictionary for Linguistic Analyses 2.0'. *Unpublished Manuscript*.

Gao, Ge. 2001. 'Intimacy, Passion, and Commitment in Chinese and US American Romantic Relationships'. *International Journal of Intercultural Relations* 25(3):329–42. doi: 10.1016/S0147-1767(01)00007-4.

Gelfand, Michele J., Lisa H. Nishii, and Jana L. Raver. 2006. 'On the Nature and Importance of Cultural Tightness-Looseness.' *The Journal of Applied Psychology* 91(6):1225–44. doi: 10.1037/0021-9010.91.6.1225.

Gibson, Mhairi A., and David W. Lawson. 2011. '"Modernization" Increases Parental Investment and Sibling Resource Competition: Evidence from a Rural Development Initiative in Ethiopia'. *Evolution and Human Behavior* 32(2):97–105.

Glazier, D. S. 1999. 'Trade-Offs between Reproductive and Somatic (Storage) Investments in Animals: A Comparative Test of the Van Noordwijk and De Jong Model'. *Evolutionary Ecology* 13(6):539–55. doi: 10.1023/A:1006793600600.

Haidt, Jonathan, and Jesse Graham. 2007. 'When Morality Opposes Justice: Conservatives Have Moral Intuitions That Liberals May Not Recognize'. *Social Justice Research* 20(1):98–116. doi: 10.1007/s11211-007-0034-z.

Hong, Zicheng. 2007. *A History of Contemporary Chinese Literature*. Vol. 1. Leiden, The Netherlands: Brill.

Hsia, Chih-tsing. 1999. 'A History of Modern Chinese Fiction'. *Indiana University Press.* Retrieved 27 May 2024.

Huang, Martin W. 2007. *Male Friendship in Ming China*. Brill Leiden.

Inglehart. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton University Press.

Jiangsu Academy of Social Sciences. 1990. *Annotated Bibliography of Chinese Popular Fiction Zhongguo Tongsu Xiaoshuo Zongmu Tiyao*. Beijing: Zhongguo wenlian chuban gongsi.

Kwan, Letty Yan-Yee. 2016. 'Anger and Perception of Unfairness and Harm: Cultural Differences in Normative Processes That Justify Sanction Assignment'. *Asian Journal of Social Psychology* 19(1):6–15. doi: 10.1111/ajsp.12119.

Lewis, Mark Edward. 2020. *Honor and Shame in Early China*. Cambridge University Press.

Lu, Hsun. 2000. *A Brief History of Chinese Fiction*. Honolulu, Hawaii: University Press of the Pacific.

Ma, Iris. 2019. 'Between Historicity and Fictionality: Xiang Kairan, Martial Arts Fiction, and Chinese Narrative Tradition'. *Modern Chinese Literature and Culture* 31(1):229–63.

Martins, Mauricio de Jesus Dias, and Nicolas Baumard. 2020. 'The Rise of Prosociality in Fiction Preceded Democratic Revolutions in Early Modern Europe'. *Proceedings of the National Academy of Sciences* 117(46):28684–91.

Miller, J. G., and D. M. Bersoff. 1992. 'Culture and Moral Judgment: How Are Conflicts between Justice and Interpersonal Responsibilities Resolved?' *Journal of Personality and Social Psychology* 62(4):541–54. doi: 10.1037//0022-3514.62.4.541.

Mischel, Walter, Yuichi Shoda, and Monica L. Rodriguez. 1989. 'Delay of Gratification in Children'. *Science* 244(4907):933–38. doi: 10.1126/science.2658056.

Nave, Gideon, Jason Rentfrow, and Sudeep Bhatia. 2020. 'We Are What We Watch: Movie Plots Predict the Personalities of Their Fans'. doi: https://doi.org/10.31234/osf.io/wsdu8.

Nettle, Daniel, and Rebecca Saxe. 2021. '"If Men Were Angels, No Government Would Be Necessary": The Intuitive Theory of Social Motivation and Preference for Authoritarian Leaders'. *Collabra: Psychology* 7(1):28105.

Ni, Zhange. 2020. 'Xiuzhen (Immortality Cultivation) Fantasy: Science, Religion, and the Novels of Magic/Superstition in Contemporary China'. *Religions* 11(1):25.

Oda, Ryo, Wataru Machii, Shinpei Takagi, Yuta Kato, Mia Takeda, Toko Kiyonari, Yasuyuki Fukukawa, and Kai Hiraishi. 2014. 'Personality and Altruism in Daily Life'. *Personality and Individual Differences* 56:206–9. doi: 10.1016/j.paid.2013.09.017.

Ogihara, Yuji. 2023. 'Chinese Culture Became More Individualistic: Evidence from Family Structure, 1953-2017'. *F1000Research* 12:10. doi: 10.12688/f1000research.128448.3.

Pan, Lynn. 2015. *When True Love Came to China*. Hong Kong University Press.

Peng, Lin, and Siyang Luo. 2021. 'Impact of Social Economic Development on Personality Traits among Chinese College Students: A Cross-Temporal Meta-Analysis, 2001–2016'. *Personality and Individual Differences* 171:110461.

Qian, Liqun, Rumin Wen, and Fuhui Wu. 1998. *Thirty Years of Chinese Modern Literary History*. Peking University Press.

Rawlings, Bruce S., Emma G. Flynn, and Rachel L. Kendal. 2022. 'Personality Predicts Innovation and Social Learning in Children: Implications for Cultural Evolution'. *Developmental Science* 25(1):e13153. doi: 10.1111/desc.13153.

Rozman, Gilbert. 1974. *Urban Networks in Ch'ing China and Tokugawa Japan*. Princeton: Princeton University Press.

Schwartz, Shalom H. 2006. 'Value Orientations: Measurement, Antecedents and Consequences across Nations'. Pp. 169–203 in *Measuring attitudes cross-nationally: Lessons from the European Social Survey*. Sage Publications.

Shang, Wenyi. 2018. 'A Comparison of the Historical Entries in Wikipedia and Baidu Baike'. Pp. 74–80 in.

Shi, Zhuomin, Lufang Wu, and Zaoying Kuang. 2018. 'How Face Consciousness Reverse Pro-Self-Behavior? A Study on Ecological Consumption from the Perspective of Social Value Orientation'. *Journal of Contemporary Marketing Science* 1(1):117–44. doi: 10.1108/JCMARS-07-2018-0004.

Shields, Anna. 2020. *One Who Knows Me: Friendship and Literary Culture in Mid-Tang China*. Brill.

Sternberg, Robert J. 1986. 'A Triangular Theory of Love'. *Psychological Review* 93(2):119. doi: https://doi.org/10.1037/0033-295X.93.2.119.

Sun, Jiaming, and Xun Wang. 2010. 'Value Differences between Generations in China: A Study in Shanghai'. *Journal of Youth Studies* 13(1):65–81. doi: https://doi.org/10.1080/13676260903173462.

Szepsenwol, Ohad, Jeffry A. Simpson, Vladas Griskevicius, and K. Lee Raby. 2015. 'The Effect of Unpredictable Early Childhood Environments on Parenting in Adulthood.' *Journal of Personality and Social Psychology* 109(6):1045.

Tangney, June P., Roy F. Baumeister, and Angie Luzio Boone. 2004. 'High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success'. *Journal of Personality* 72(2):271–324. doi: 10.1111/j.0022-3506.2004.00263.x.

Vaish, Amrisha, Malinda Carpenter, and Michael Tomasello. 2009. 'Sympathy through Affective Perspective Taking and Its Relation to Prosocial Behavior in Toddlers'. *Developmental Psychology* 45(2):534–43. doi: 10.1037/a0014322.

Warneken, Felix, and Michael Tomasello. 2009. 'The Roots of Human Altruism'. *British Journal of Psychology (London, England: 1953)* 100(Pt 3):455–71. doi: 10.1348/000712608X379061.

Xin, Sufei, and Ziqiang Xin. 2016. 'Birth Cohort Changes in Chinese College Students' Loneliness and Social Support: One up, as Another Down'. *International Journal of Behavioral Development* 40(5):398–407. doi: 10.1177/0165025415597547.

Xin, Ziqiang, and Sufei Xin. 2017. 'Marketization Process Predicts Trust Decline in China'. *Journal of Economic Psychology* 62:120–29. doi: 10.1016/j.joep.2017.07.001.

Xu, Ting. 2013. 'The Production and Circulation of Manuscripts and Printed Books in China Compared to Europe, ca. 581-1840'. *The Other Canon Foundation and Tallinn University of Technology Working Papers in Technology Governance and Economic Dynamics*.

Xu, Yi, and Takeshi Hamamura. 2014. 'Folk Beliefs of Cultural Changes in China'. *Frontiers in Psychology* 5:1066. doi: https://doi.org/10.3389/fpsyg.2014.01066.

Xu, Yi, Bas van Leeuwen, and Jan Luiten van Zanden. 2018. 'Urbanization in China, ca. 1100–1900'. *Frontiers of Economics in China* 13(3):322–68. doi: https://doi.org/10.3868/s060-007-018-0018-9.

Yan, Yunxiang. 2003. *Private Life under Socialism: Love, Intimacy, and Family Change in a Chinese Village, 1949-1999*. Stanford University Press.

Yü, Ying-shih. 2021. *The Religious Ethic and Mercantile Spirit in Early Modern China*. edited by Hoyt. Columbia University Press.

Zeng, Rong, and Patricia M. Greenfield. 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.

Zeng, TianChen, Joey T. Cheng, and Joseph Henrich. 2022. 'Dominance in Humans'. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377(1845):20200451. doi: 10.1098/rstb.2020.0451.

Zhang, Lei, Fanchao Qi, Zhiyuan Liu, Yasheng Wang, Qun Liu, and Maosong Sun. 2019. 'Multi-Channel Reverse Dictionary Model'. Pp. 312–19 in *AAAI Conference on Artificial Intelligence 2020*. Vol. 34.

Zhang, Longxi. 2004. 'History and Fictionality: Insights and Limitations of a Literary Perspective'. *Rethinking History* 8(3):387–402. doi: 10.1080/1364252042000247837.

Zhou, Chan, Wai Ying Vivien Yiu, Michael Shengtao Wu, and Patricia M. Greenfield. 2018. 'Perception of Cross-Generational Differences in Child Behavior and Parent Socialization: A Mixed-Method Interview Study with Grandmothers in China'. *Journal of Cross-Cultural Psychology* 49(1):62–81. doi: https://doi.org/10.1177/0022022117736029.