nature human behaviour

Article https://doi.org/10.1038/s41562-025-02122-8

# The conceptual structure of human relationships across modern and historical cultures

Received: 28 December 2023

Accepted: 21 January 2025

Published online: 13 March 2025

Check for updates

Xi Cheng[1,9], Haroon Popal[1,2,9], Huanqing Wang[1,9], Renfen Hu[3], Yinyin Zang[4], Mingzhe Zhang[1], Mark A. Thornton[5], Yina Ma[1], Huajian Cai[6], Yanchao Bi[1], Jamie Reilly[7], Ingrid R. Olson[8] & Yin Wang[1]

A defining characteristic of social complexity in *Homo sapiens* is the diversity of our relationships. We build connections of various types in our families, workplaces, neighbourhoods and online communities. How do we make sense of such complex systems of human relationships? The basic organization of relationships has long been studied in the social sciences, but no consensus has been reached. Here, by using online surveys, laboratory cognitive tasks and natural language processing in diverse modern cultures across the world ($n = 20{,}427$) and ancient cultures spanning 3,000 years of history, we examined universality and cultural variability in the ways that people conceptualize relationships. We discovered a universal representational space for relationship concepts, comprising five principal dimensions (formality, activeness, valence, exchange and equality) and three core categories (hostile, public and private relationships). Our work reveals the fundamental cognitive constructs and cultural principles of human relationship knowledge and advances our understanding of human sociality.

No man is an island. Human life is a process of seeking, sustaining, repairing, judging, adjusting and sometimes dissolving relationships[1]. The quality and quantity of relationships are integral not only to our survival but also to our capacity to thrive[2,3]. Social isolation and poor relationships affect an individual's cognition, behaviour, development and well-being[4,5].

Understanding the nature of human relationships lies at the heart of the social sciences. However, studying relationships is challenging for several reasons. First, human relationships are characterized by their diversity and complexity. Social structure in non-human primates is largely dominated by hierarchy and affiliation[6]. Human society, in contrast, is governed by far more diverse and complex types of relationships (for example, frenemies, godparents and online friends). Human relationships are also context-dependent and multifaceted, involving numerous factors such as time, space, emotions, communication and cultural norms[7]. These factors interact with one another in intricate ways, making it difficult to isolate and study individual components. Unravelling the underlying elements and organizational structures of such a complex relationship system thus remains a vexing problem.

Second, human relationships are subjective beliefs, experiences and practices shaped by the unique perspectives, attitudes and personalities of the individuals involved and maintained by dynamic, unwritten rules over time and across societies[8]. This subjectivity makes

[1]State Key Laboratory of Cognitive Neuroscience and Learning, and IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China. [2]Department of Psychology, University of Maryland, College Park, MD, USA. [3]School of International Chinese Language Education, Beijing Normal University, Beijing, China. [4]School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China. [5]Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA. [6]Department of Psychology, University of Oklahoma, Norman, OK, USA. [7]Department of Communication Sciences & Disorders, Temple University, Philadelphia, PA, USA. [8]Department of Psychology and Neuroscience, Temple University, Philadelphia, PA, USA. [9]These authors contributed equally: Xi Cheng, Haroon Popal, Huanqing Wang. e-mail: mirrorneuronwang@gmail.com

it difficult to establish objective and uniform measures of relationships and to compare them across individuals. The degree to which people around the world (and across the generations) share the same set of cognitive, behavioural and cultural principles of relationships has yet to be fully evaluated.

Third, human relationships are widely studied in the social sciences. A wave of enthusiasm from multiple disciplines in the 1970s–1990s led to the exploration of the internal organization of relationships, each with their own theoretical perspectives and methodological approaches. This interdisciplinary nature can make it challenging to establish a unified understanding of relationships and to compare findings across disciplines. For example, sociologists were interested in the formation and organization of social relationships and discovered a three-factor model for role-based relationships (that is, intimacy, visibility and regulation)[9], anthropologists attempted to understand the foundations of social coordination across cultures and proposed four elementary forms of social bonds (that is, communal sharing, authority ranking, expected reciprocity and market pricing)[1], cognitive psychologists studied the perception of relationships and revealed a four-dimensional framework (that is, valence, equality, activeness and formality)[10], and communication scholars focused on the communication quality in personal relationships and proposed three factors for effective relational dialogues (that is, positiveness, intimacy and control)[11]. All theories have proved insightful, as attested by their endurance in the field. However, no consensus has been reached because researchers in different disciplines have approached the problem of human relationships using their unique features of interest and thus have tapped into distinct feature spaces and relationship types.

To address the challenging questions above, here we focus on the common sense of human relationships—how ordinary people mentally conceptualize and understand human relationships (that is, relationship concepts). By building up a unified framework across multiple disciplines, we aim to clarify the underlying elements and organizational structures of the relationship concept system and reveal the similarities and differences in relationship conceptualization across different cultures and time periods.

## Results

### Study 1: a unified representational space across disciplines

In Study 1, we attempted to synthesize this cross-field literature and build a unified representational space across disciplines. On the basis of an extensive literature review (Extended Data Fig. 1), we collected and summarized 30 conceptual features of relationships from 15 prominent existing theories to encompass a composite feature space, including activeness, communality, concreteness, equality, endurance, formality, intensity, intimacy, reciprocity, societal importance, socio-emotionality, uniqueness, valence and visibility, among others (see the full list in Supplementary Table 1). These theoretical features were originally derived from dimensionality reduction or clustering techniques in each discipline, but here they were assessed together in a dimensional survey and prepared to be further reduced into higher-order components. To capture the diversity of all possible relationships, we used a naïve natural language processing (NLP) model (Methods) to generate a comprehensive list of 159 typical relationships in English, including both common (for example, siblings, friends and enemies) and uncommon ones (for example, master–servant and friends with benefits) (see the full list in Supplementary Table 2).

A diverse group of native English speakers from the USA ($n = 1{,}065$) were recruited via MTurk and completed an online survey where they rated 159 relationships on 30 theoretical features. For example, the participants were asked to rate the equality of 'between friends'. For details on feature selection and their definitions, please see Extended Data Fig. 1 and Supplementary Table 1. The next processing step involved reducing this high-dimensional feature set into a smaller number of orthogonalized latent factors via principal component analysis (PCA). The PCA extracted five latent dimensions, accounting for 82.14% of the variance of dimensionality ratings (see Methods and statistics on how to determine the optimal PCA component number). On the basis of close examination of the PCA loadings and relationship scores (Fig. 1), the first dimension was identified as 'formality'. This dimension contrasts formal, occupational and publicly visible relationships that adhere to rules and regulations (for example, co-workers and officer–soldier) with informal, socio-emotional and private relationships that exhibit a looser, more casual style (for example, parent–infant and wife–husband). The second dimension, which we termed 'activeness', loaded highly on activeness, synchronicity and spatial distance. Close relationships (for example, wife–husband and siblings) and distant relationships (for example, distant relatives and strangers) occupied the poles of this dimension. The third dimension was described as 'valence', with friendly, harmonious and high-solidarity relationships at one pole (for example, church members and writer–reader) and conflictual, hostile and antagonistic relationships at the other (for example, bully–victim and slave–master). We named the fourth dimension 'exchange', as it distinguishes between dyads exchanging concrete resources such as money, goods and services (for example, dealer–buyer and prostitute–customer) and dyads exchanging symbolic, intangible resources such as information, love and identity (for example, celebrity–haters and brother–sister). The fifth dimension was labelled 'equality', as it differentiated dyads with equal powers (for example, sports rivals and pen friends) from dyads with unequal powers (for example, man–god and politician–supporter). Other dimensionality reduction techniques (that is, independent component analysis, exploratory factor analysis and multidimensional scaling) were also evaluated to examine the robustness of the latent factor solution to different statistical algorithms, and they all yielded the same five-factor solution (Supplementary Fig. 3). We hereafter refer to this five-dimensional solution as the FAVEE model (an abbreviation for formality, activeness, valence, exchange and equality).

Categorical thinking (for example, family, friends and colleagues) is pervasive when people define and manage their social connections. We next studied how people sort relationships and how categorical representations relate to the FAVEE dimensions. Two cognitive paradigms were employed: in the multi-arrangement task[12], participants judged the similarity between the 159 relationships by arranging them on a 2D computer screen in such a way that the distance between any two relationships reflected their conceptual dissimilarity (that is, the more conceptually similar, the closer together the relationships were); in the free sorting task[13], participants classified the same set of relationships into labelled categories of their choosing.

Using a within-subject design, we recruited 60 US participants to complete three tasks in the laboratory (that is, one dimensional survey and two cognitive tasks) (Fig. 2a). Categorical representations were derived from each task by applying clustering algorithms to the dissimilarity matrix of relationship concepts. Three clusters were found in the dimensional survey, which can be labelled 'hostile, private and public' (abbreviated as the HPP model) (for optimal cluster details, see Methods). Relationships in the 'hostile' cluster featured people who are antagonistic or have negative feelings with each other, such as 'divorced spouses' and 'business rivals'. Relationships in the 'private' cluster were personal and family ties, such as 'siblings' and 'close friends'. Relationships in the 'public' cluster were formal and occupational and had impersonal ties, such as 'driver–passenger' and 'employer–employee'. In contrast, clustering on the two cognitive tasks revealed six canonical relationship types: hostile, familial, romantic, affiliative, transactional and power. Text analysis on the labels during the free sorting task further revealed that six canonical types emerged from three HPP categories (Fig. 2b): while the hostile cluster in the HPP model remained, the private cluster was divided into three distinct classes (familial, romantic and affiliative relationships), and the public

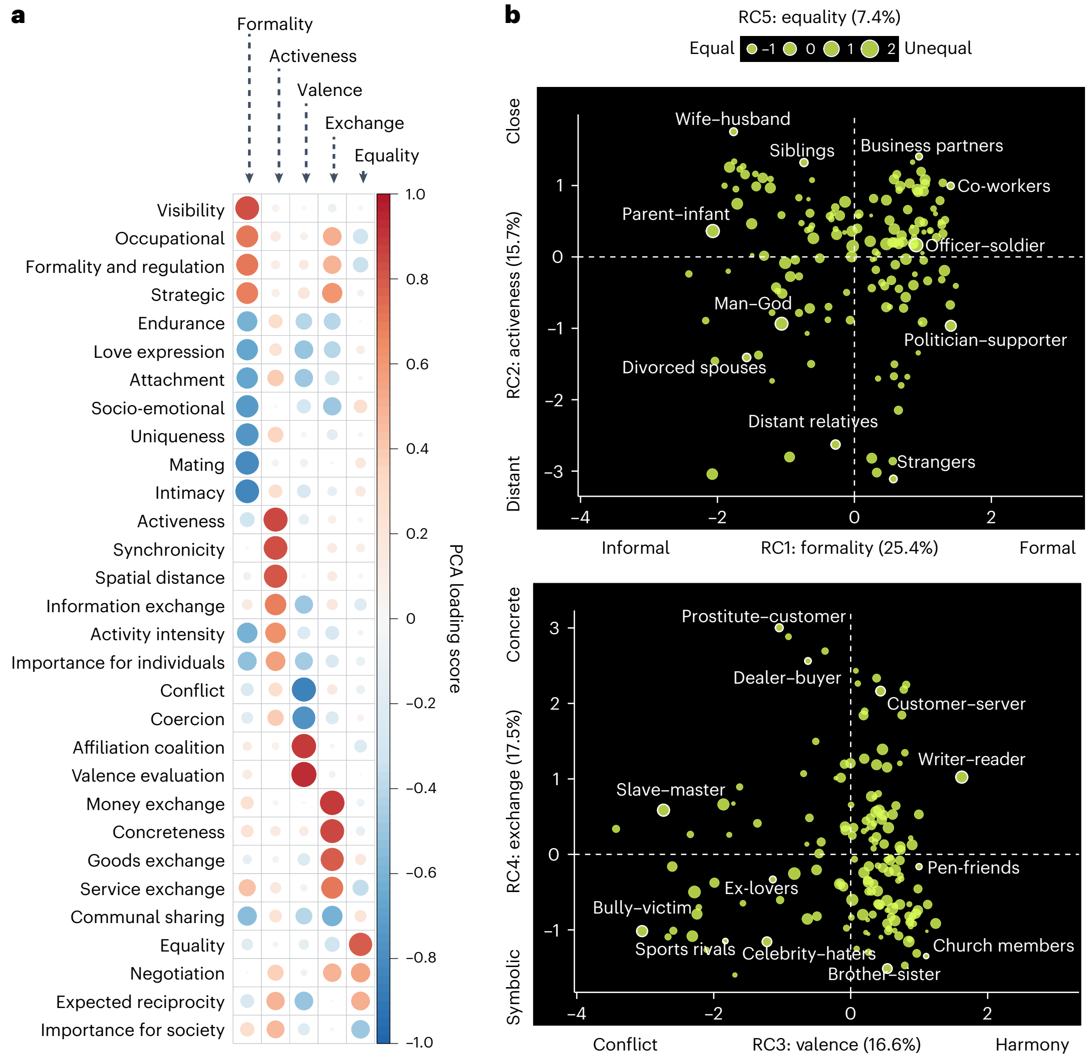



**Fig. 1 | A five-dimensional model of human relationships (FAVEE model). a**, PCA loadings on 30 theoretical features derived from multidisciplinary literature. Dark colours on the colour bar represent larger values, with blue indicating negative values and red indicating positive values. **b**, 159 relationships were plotted in the 5D space on the basis of their scores on each dimension. PCA loadings and relationship scores jointly suggested that RC1 corresponded to formality (formal versus informal), RC2 to activeness (close versus distant), RC3 to valence (harmony versus conflict), RC4 to exchange (concrete versus symbolic) and RC5 to equality (unequal versus equal). Each axis is labelled with the variance explained for the corresponding dimension. For more details about the spatial location of each relationship, please see the dynamic figures at https://bnu-wang-msn-lab.github.io/FAVEE-HPP/.

cluster was split into two classes (transactional and power relationships). To further clarify the associations between the FAVEE and HPP models, a dimension–category hybrid representation was evaluated where clustering techniques were applied to the relationships in the FAVEE embedding space. Again, three HPP clusters were identified (Fig. 2c), and each was embedded in a unique location in the 5D space: the private and public clusters were located separately at the two ends of the formality dimension, and the relationships that were low on the valence dimension formed the hostile cluster. This implied that HPP categories could originate from the FAVEE dimensions.

In sum, Study 1 revealed that when people think about social relationships, they attend to five key features. We demonstrated that all relationship concepts are mentally represented in a high-order FAVEE space, and the conceptual similarity between each pair of relationships can be represented as the distance in the 5D space. Once the spatial proximity among relationships is close enough along certain featural dimensions, they can be self-clustered into meaningful categories (for example, three HPP clusters or six canonical types). Relationship categories thus emerge from uneven distributions along the FAVEE dimensions, and relationship taxonomies can be understood as discrete sets of categories living in a continuous multidimensional space (see an illustrative flow chart in Fig. 2d).

### Study 2: universality and variability across modern cultures

All human cultures have rich vocabularies devoted to describing human relationships. Translation dictionaries, for example, suggest that the English word 'neighbours' can be equated with the Chinese word '邻居' and the Hebrew word 'שכנים'. However, does this mean that the concept of 'neighbours' is the same in the USA, China and Israel? In Study 2, we explored this question by examining representations of relationship concepts across 19 global regions and 10 languages. We aimed to reveal the cross-cultural similarity and differences and their underlying cultural mechanisms.

Study 2 was preregistered on the Open Science Framework (https://osf.io/swr2c) on 13 June 2022. We report deviations from preregistration in Supplementary Method 3. A large sample of online participants ($n$ = 17,686) were recruited from 19 global regions with diverse ecological (geography, climate and subsistence), biological (genetics and disease prevalence) and sociocultural backgrounds (language, ethnicity, education, religion, politics, wealth and urbanization) (see Supplementary Fig. 13 for the details). The dimensional survey approach was adopted due to its higher within-culture stability over cognitive tasks (Supplementary Fig. 2). For each region, three types of representational geometries were generated on the basis of representational dissimilarity matrices (RDMs)[14,15]: the full-feature model (that is, RDMs based on the original data on all evaluative features without applying any dimensionality reduction or clustering techniques), a dimensional model (that is, RDMs based on FAVEE) and a categorical model (that is, RDMs based on HPP). The degree of cross-cultural concordance in relationship concepts was assessed on the basis of these region-specific representational geometry models.

Consistent with Study 1, we identified the 5D FAVEE space and three HPP categories in both globally aggregated data (Extended Data Fig. 2) and regional data (Supplementary Figs. 4 and 5). Using leave-one-region-out cross-validation, each region's unique representational geometries were accurately predicted by the left-out globally aggregated data (Fig. 3a). The ability of the FAVEE and HPP models to consistently predict relationship representations across regions suggests that they might be universal structures of relationship concepts that can be generalized across the world. In addition, to examine how well the five FAVEE dimensions represent all theoretical relationship features, we performed model comparison analysis between the FAVEE model and other existing theories. We found that the FAVEE model outperformed 15 other theories in data fitting and explained variance across global regions (Extended Data Fig. 3). Therefore, although past theories all attempted to reduce numerous relationship features into fewer components, FAVEE is the most representative, parsimonious and consistent model across cultures.

Although the basic organization of relationship concepts was found to be globally shared, there was also rich cultural variation. For example, people around the world seemed to have a different understanding of public relationships but held similar views on familial and romantic relationships (Extended Data Fig. 4). To further explore these findings, we implemented representational similarity analysis (RSA) to quantitatively model the cross-region variability of representational geometries on the basis of regressions of a variety of ecological, biological and sociocultural variables (Fig. 3b). Religion and modernization were the only two factors that significantly predicted cross-region variability in representational geometries (see the detailed statistics in Extended Data Table 1), and regions with similar religions and modernization levels were found to have similar representational geometries of relationships (Fig. 3c). Here modernization refers to a composite metric based on the education, urbanization and wealth of a country[16], and religion estimates the percentages of adherents of 21 religious denominations (Supplementary Table 3). Follow-up RSAs revealed that the two factors exerted predictive power on distinct dimensions and categories (Supplementary Fig. 7).

To further delineate and elaborate the fine-grained cultural differences, we collected additional data in China ($n$ = 6,128) (Supplementary Fig. 8) and directly compared it with the USA at a finer scale (Fig. 4). To rule out the effects of language and translation, two rounds of data collection were conducted. In the first round, 159 relationships directly translated from the US relationship list were adopted. In the second round, a new list of 258 relationships was generated by Chinese NLP algorithms (see the details in Supplementary Method 1), which included numerous Chinese-unique relationships (that is, some cannot be translated linguistically, and others are culture-specific; see the full list in Supplementary Table 4). Our analysis revealed no significant differences between the datasets of directly translated relationships and those generated via Chinese NLP algorithms (all $r > 0.622$, all $P < 0.001$; Supplementary Fig. 8), confirming that our results were not influenced by language or translation. There were more intriguing findings in the direct comparisons between the USA and China. We found, when understanding closeness in human relationships, Americans seemed to focus more on physical distance, whereas Chinese focused on psychological distance (Fig. 4c). For example, ancestor–descendant was considered by Americans to be a distant relationship because two sides have infinitely far physical distance. Chinese evaluated this relationship as being less distant due to ancestor veneration in the foundational philosophy of Confucianism (for example, high spiritual intimacy with ancestors). When understanding power in human relationships, individuals in China hold stronger stereotypes of inequality among family members (for example, uncle–nephew; Fig. 4d), which is consistent with the Confucian ideal of filial piety. When evaluating social exchange in private relationships, Americans seemed to experience more concrete resource exchanges than Chinese, which could be associated with their higher modernization level or foundational values linked to capitalism (Fig. 4e). For example, long-distance lovers in the USA often buy gifts such as flowers for each other, whereas symbolic exchanges, such as long telephone calls, were typically observed in Chinese long-distance partners. Together, these subtle cultural differences in relationship conceptualization seemed to be highly interdependent with USA–China differences in religion and modernization level.

Finally, as all 19 global regions were industrial societies, we validated the FAVEE-HPP model in a non-industrial society—the Chinese Mosuo tribe, a small-scale matrilineal society living near Lugu Lake in the Tibetan Himalayas. As a traditional agrarian society, the Mosuo society is distinct from industrialized societies in social organization, economy system, language, beliefs and lifestyle (see key features of the Mosuo society in Extended Data Fig. 5). Field research data from 229 native Mosuo people indicated that Mosuo culture still conforms to FAVEE-HPP structures when understanding relationships. Highly similar representational geometries can be observed between the Mosuo, Chinese Han and other industrial societies in the world (all $r > 0.600$, all $P < 0.001$; Extended Data Figs. 5b,d). This confirmed that people from non-industrial and industrial societies share the same set of conceptual structures for relationships.

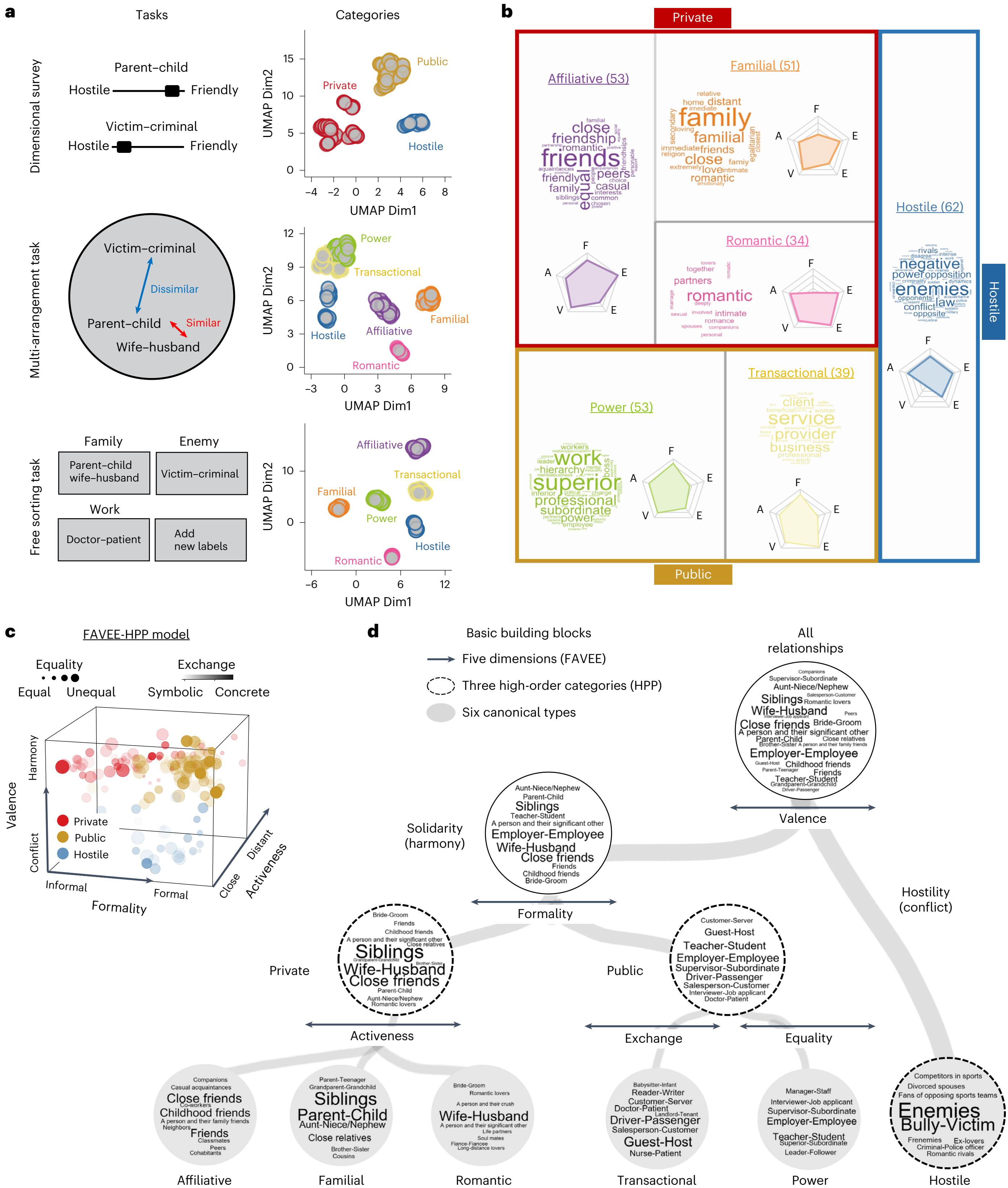
Study 2 demonstrated that relationship concepts reside in a universal, low-dimensional space shared by people around the world. While many concepts are similarly positioned in FAVEE-HPP space regardless of culture, other concepts exhibit significant cultural variability (see Extended Data Fig. 6 for conceptual differences of 'neighbours' in the USA, China and Israel as a vivid example). This variation was found to be tied to religion and modernization differences between regions.

**Fig. 2 | Categorical and dimensional representations of human relationships.** **a**, Three behavioural tasks and their corresponding categories (via UMAP and $k$-means clustering). The dimensional survey indicated three high-order categories (hostile, private and public, abbreviated as HPP), whereas both cognitive tasks identified six canonical relationship types. **b**, Text analysis on categorical labels in the free sorting task revealed the label hierarchy: the hierarchical clustering algorithm first derived three HPP categories; next, the 'public' cluster was further split into 'transactional' and 'power', and the 'private' cluster was subdivided into 'family', 'romantic' and 'affiliative'. The word clouds in the treemap display label names in each cluster. The radar plots illustrate the average scores on the FAVEE dimensions across relationships in each category. **c**, A dimension–category hybrid model was evaluated by applying $k$-means clustering to the 5D FAVEE embedding space. Three HPP clusters were found, suggesting that HPP could emerge from the FAVEE space. **d**, The three building blocks of the relationship concept system and their internal associations. The flow chart illustrates how people represent relationships via the joint framework of dimensions and categories. Five FAVEE dimensions can be understood as filters (for example, all relationships can be filtered into hostile and solidarity relationships via the valence dimension), and relationship categories are formed by uneven distributions projected on dimensions (see the details in Supplementary Fig. 14). Our data suggest that six canonical relationship types (in the grey circles) originate from three HPP categories (in the dashed circles), which inherently emerge from the 5D FAVEE space (arrows).

## Study 3: relationship representations in ancient cultures

Study 1 investigated how human relationships are mentally represented and discovered the FAVEE-HPP structures. Study 2 examined where in the world the FAVEE-HPP model applies and showed its generalizability to diverse global regions. In Study 3, we explored when in history this model can apply. In Studies 1 and 2, we only examined contemporary societies, which are far from representative of all cultures. An investigation on ancient cultures will help verify the persistence of the FAVEE-HPP model through time.

We employed state-of-the-art NLP techniques to capture ancient people's perception and comprehension of human relationships. This involved analysing large-scale text corpora sourced from historical archives, enabling us to gain insights into their conceptualization of relationships. Analysing texts can offer a unique window into human



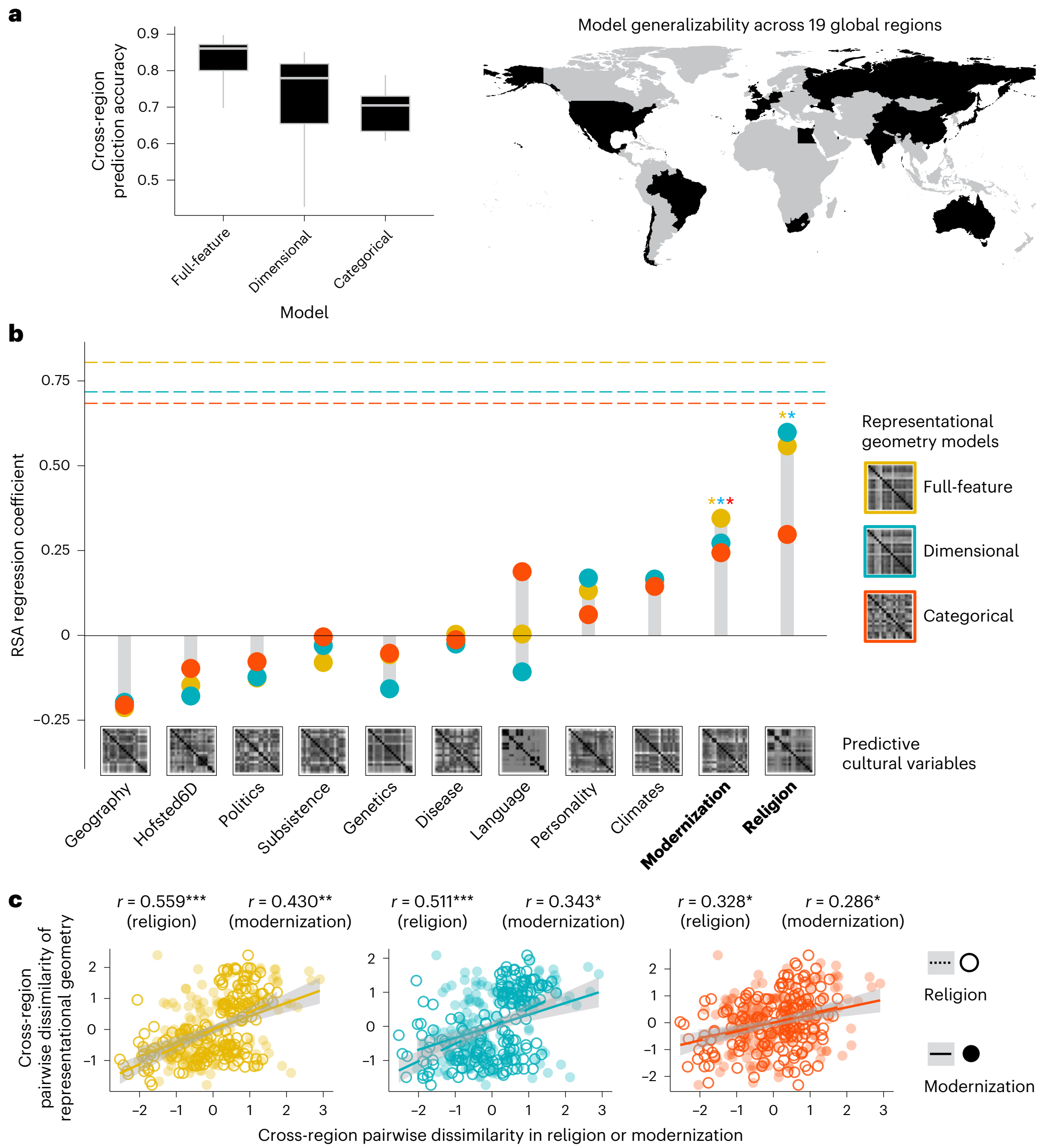



**Fig. 3 | Universality and cultural variability of relationship representational geometries. a**, Three representational geometries were derived from the full-feature model (all evaluative features), the dimensional model (FAVEE) and the categorical model (HPP). Using leave-one-region-out cross-validation, high similarities of representational geometries (all $r > 0.687$) were observed across 19 global regions (black areas on the world map), suggesting that FAVEE dimensions and HPP categories are commonly shared across the world. The box plots show the median (horizontal line) and the interquartile range (IQR) ± 1.5 × IQR. Here the full-feature model represents the noise ceiling of cross-region representational similarity. **b**, Computational modelling of representational geometries using RSA multiple regression. A set of ecological, biological and sociocultural RDMs were fit to predict the cross-regional relationship RDM, and religion and modernization level were the only two factors that could significantly predict the cross-regional variability of representational geometries. The asterisks indicate significant regression coefficients according to the Mantel test, and the dashed lines indicate noise ceilings (Methods). **c**, Significant correlations between cross-region pairwise dissimilarity of representational geometry and cross-region pairwise dissimilarity in religion and modernization level. The shaded area represents the 95% confidence interval. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; one-sided permutation tests.

psychology[17]. Prior research has suggested that word embeddings (representations) reflect the ways people understand concepts such as object knowledge[18], personality traits[19] and mental states[20]. The advent of pretrained language models (PLMs) and large language models (LLMs) revolutionizes tools for analysing texts on a massive scale[21]. Probing language models pretrained on Chinese historical text corpora thus allows us to query relationship understanding from people in ancient China (for example, Qin Dynasty, 221 BCE)—populations otherwise inaccessible to modern researchers.

We conducted an initial investigation to examine whether language models can generate human-like relationship understanding (Fig. 5a). This was achieved by employing an approach that combines PLMs, as proposed by Cutler and Condon[19], with the use of LLMs such as GPT-4 (see the details in Methods). Specifically, we designed the following query (in Chinese) as the input for the pretrained model:

> '[DESC] The most salient feature of the relationship [TERM] is [MASK].'

where the [TERM] token was substituted with one of the 258 Chinese relationship terms, while the [MASK] token represented the conceptualization of the target relationship. During the pretraining, the

 https://doi.org/10.1038/s41562-025-02122-8

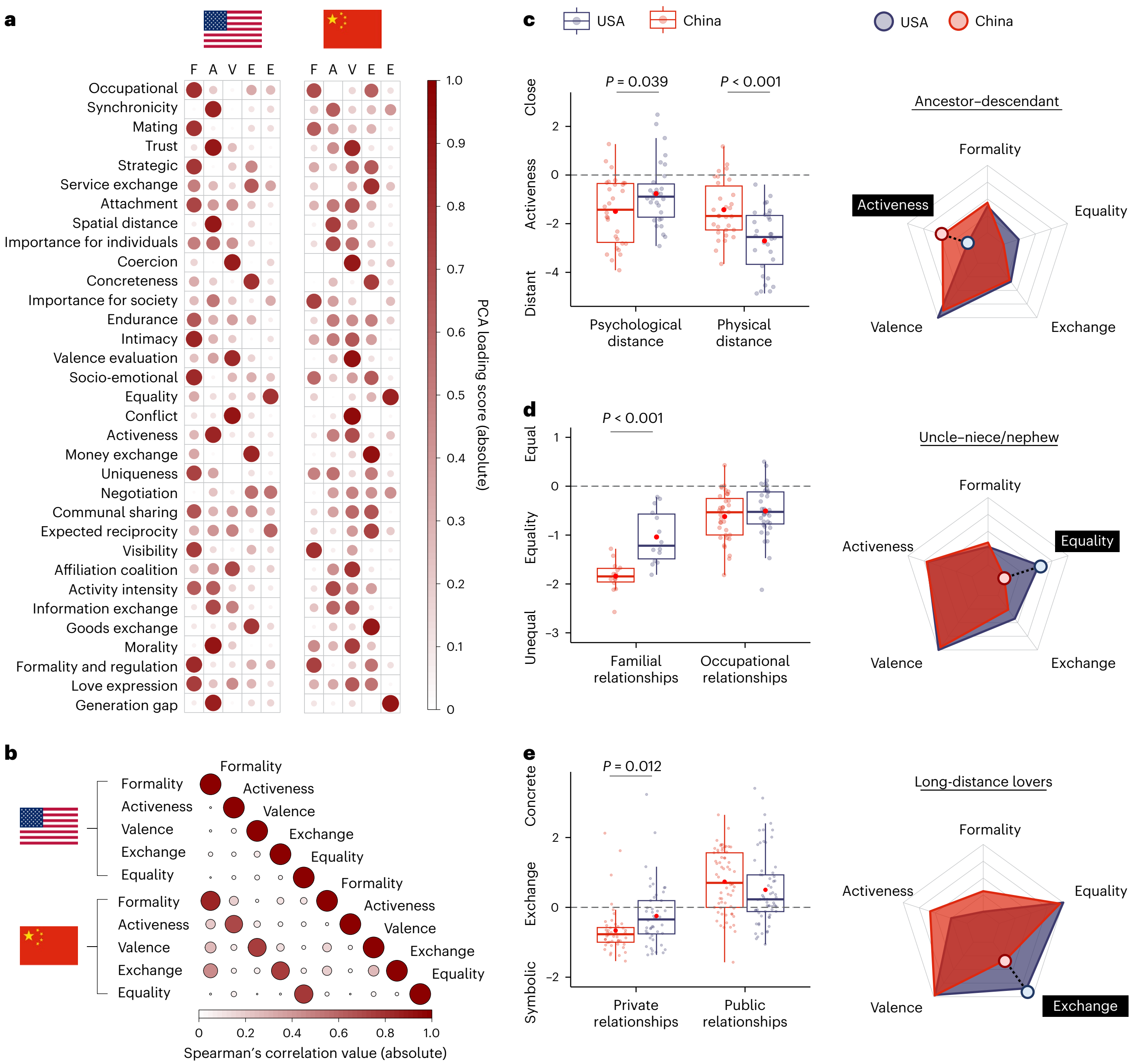



**Fig. 4 | Comparisons between the USA and China. a**, Similar PCA loadings for the two countries. **b**, Correlation of relationship scores between the two countries. The high concordance of PCA loadings (all $r > 0.707$) and relationship scores (all $r > 0.704$) suggests a common FAVEE space in the two countries. **c**, USA versus China on the conceptualization of close relationships. PCA scores on psychological distance dimensions (for example, attachment, love expression and intimacy) and physical distance dimensions (for example, spatial distance and synchronicity) were computed for the 30 most distant relationships in each country. Americans weighted physical distance more when judging close relationships, whereas the Chinese considered both psychological and physical distance. For example, the Chinese rated ancestor–descendant less distant than Americans, possibly due to ancestor worship (that is, close psychological distance). **d**, USA versus China on the conceptualization of power relationships. While both groups shared similar views on equality in occupational relationships, the Chinese judged familial relationships as more unequal than Americans (for example, uncle–nephew). **e**, USA versus China on the conceptualization of social change in private and public relationships. While both countries believed that public relationships mainly exchange concrete resources (for example, money, goods and services) and private relationships exchange symbolic resources (for example, love, advice and information), Americans experienced more concrete resource exchanges in private relationships than the Chinese. For example, long-distance lovers in the USA send gifts to each other regularly, whereas long-distance lovers in China spend long hours on telephone chats instead. All box plots show the median (horizontal line) and the 25th and 75th quantiles (box edges), and the whiskers extend to the most extreme data points within 1.5 times the IQR. Analysis of variance and $t$-test were used to compare cultural differences with a two-tailed test.

language model used contextualized embeddings of the [MASK] token to predict the most probable words to occur in that position, given the contexts. To enrich the contextual information, we incorporated [DESC], which denotes relationship-specific descriptions generated by a state-of-the-art LLM, GPT-4. These descriptions played a pivotal role in establishing a contextual framework for the subsequent representations of relationships by the language model. After systematic testing with different query types, token positions and

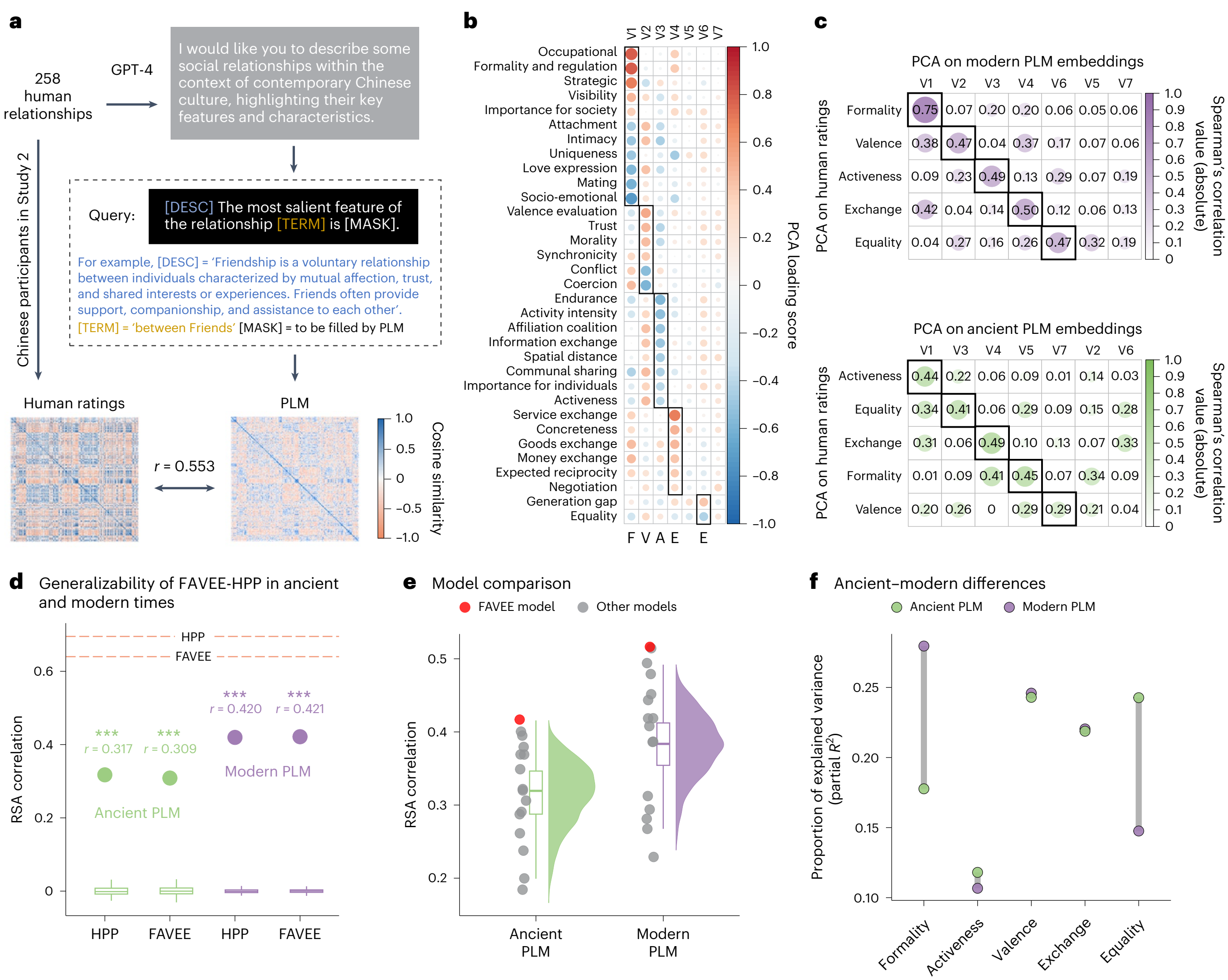



**Fig. 5 | Validation of the FAVEE-HPP model in modern and ancient China using language models. a**, Pipeline for generating PLM embeddings. The query (dashed box) was formulated as '[DESC] The most salient feature of the relationship [TERM] is [MASK]', where the [DESC] component was generated by GPT-4 (with the prompt in the grey box) and the [TERM] component was replaced with one of the 258 relationships. The [MASK] component was filled by the PLM with words that are most likely to occur there, given the context. The last layer vector (768 dimensions) was extracted on the position of [MASK] for each of 258 relationships, resulting in a 258 × 768 matrix. The resemblance of human representations (that is, the similarity matrix of 258 relationships from Chinese populations in Study 2) and PLM representations (pretrained on a modern Chinese corpus) was 0.553. **b**, PCA on modern PLM embeddings. By correlating human ratings on 33 dimensional features with the first seven principal components of PLM embeddings, we found that PLM representations (V1–V4 and V6) had captured corresponding features of FAVEE structures. **c**, Correspondence between PCA on human ratings and PCA on modern PLM embeddings (in purple) and ancient PLM embeddings (in green). **d**, Generalizability of the FAVEE-HPP model in ancient and modern times. RSA correlation on RDMs suggested that the FAVEE dimensions and HPP categories can significantly predict relationship representations in historical (as reflected by ancient PLM embeddings) and modern times (as reflected by modern PLM embeddings). The box plots indicate RSA correlation distributions at chance level by permutation, and the dashed lines indicate noise ceilings. **e**, Model comparison analysis suggested that the FAVEE model outperformed 15 other existing theories in predicting ancient and modern PLM embeddings. Broadly, FAVEE was significantly better than any random combination of theoretical features (the null distribution is indicated by the box plots and density plots; modern: $P < 0.001$, ancient: $P = 0.006$). **f**, Ancient–modern differences. Using linear combinations of the five FAVEE components as regressors to predict ancient and modern PLM embeddings, we found substantial changes on 'formality' and 'equality' for their explained variance in ancient and modern PLM representations, suggesting that these two dimensions may contribute differently to relationship conceptualization across time. ***$P < 0.001$; one-sided permutation tests. All box plots show the median (horizontal line) and the interquartile range (IQR) ± 1.5 × IQR.

embedding layers of the language model (Supplementary Fig. 9), we were able to identify optimal PLM representations highly resembling human relationship representations ($r = 0.553$, $P < 0.001$; Fig. 5a). Critically, PCA on PLM representations generated components (Fig. 5b) corresponding well with the FAVEE structures (all $r > 0.470$, all $P < 0.001$; Fig. 5c in purple).

Since we had confirmed that PLM embeddings could reflect human-like relationship understanding, we harnessed the ancient PLM as a proxy of the ancient human mind and sought evidence of FAVEE-HPP structures in an ancient language model pretrained on a comprehensive compilation of historical Chinese texts ranging from the Zhou Dynasty (~1046 BCE) to the Qing Dynasty (1912 CE)[22]. For more accurate historical context, we first prompted GPT-4 to describe the relationships within the context of ancient China. We then recruited human experts in ancient Chinese language, literature and history to manually refine the descriptions and express them in Classical Chinese. This ensures that the DESC effectively matches the linguistic features and relationship characteristics of the ancient era (Supplementary Method 2). These experts also carefully selected 120 relationships that existed in ancient China (Supplementary Table 6). As expected, FAVEE structures can be identified after applying PCA on ancient PLM embeddings (all $r > 0.287$, all $P < 0.001$; Fig. 5c in green). Next, if the FAVEE-HPP model can capture relationship representations in history, then the relationships that are closer to each other within FAVEE-HPP space should be represented by vectors that are closer to each other in ancient PLM embeddings. Indeed, for both FAVEE dimensions and HPP categories, we found significant correlations between RDMs in human ratings and RDMs in ancient PLM embeddings (Fig. 5d). Model comparison analysis suggested that the FAVEE model outperformed other theoretical models in predicting ancient and modern PLM embeddings (Fig. 5e). To further reveal the difference between ancient and modern China, we evaluated the relative contribution of each FAVEE dimension when predicting relationship representations in ancient and modern PLMs (Fig. 5f). We found that 'formality' explained more variance in modern than in ancient times (modern, 0.279; ancient, 0.178), whereas 'equality' accounted for more variance in ancient than in modern times (modern, 0.148; ancient, 0.243). This suggests that, compared with modern Chinese, ancient Chinese might put more weight on equality features (for example, social hierarchy) but less on formality features (for example, occupations) when understanding relationships.

We also performed expert validation on the ancient PLM to check whether it had expert-like knowledge on ancient relationships. A group of university scholars ($n = 44$) were asked to rate all 120 relationships in the context of ancient Chinese culture, and FAVEE-HPP structures can be reliably identified from their ratings (Supplementary Fig. 10). Critically, ancient PLM embeddings showed higher agreement with expert ratings than with non-expert ratings, suggesting that our PLM embeddings did capture scholarly knowledge and insights on how ancient Chinese conceptualized relationships.

Study 3 demonstrated that language models can generate human-like relationship understanding and have expertise in historical contexts. By decomposing PLM embeddings, we can identify FAVEE-HPP structures in ancient and modern representations of relationships. Furthermore, the FAVEE-HPP model outperformed other models across different time periods. These findings highlight the broad and effective generalization of the FAVEE-HPP model from contemporary society to the ancient world, spanning a history of 3,000 years.

## Discussion

In the past 50 years, social scientists have sought to understand the nature of human sociality, but there is still no consensus on the elemental forms and overarching organization of human relationships. To help address this long-standing question, the present study examined how conceptual knowledge of relationships is mentally represented and organized. We created a generalized framework that unifies existing theories across multiple disciplines and discovered a set of five dimensions (FAVEE) and three categories (HPP) that scaffold the conceptual space of human relationships (Study 1). Converging evidence suggests that the FAVEE-HPP framework is commonly shared across modern societies (Study 2) and historical cultures (Study 3), and it exceeds existing theories in model performance (Supplementary Fig. 6), consistency across global regions (Extended Data Fig. 3) and endurance over time (Fig. 5e). We also extended the FAVEE framework to non-dyadic relationships (Extended Data Fig. 7) and confirmed its generalizability to triadic relations (for example, love triangle) and group relations (for example, rich–poor, Democrats–Republicans).

As a parsimonious model of human relationships, the FAVEE-HPP framework will inspire theory, experimental design, hypothesis testing and reinterpretations of empirical data in the social sciences. For example, since socialization is hypothesized to be one of the major drivers behind the evolution of cognitive abilities, the FAVEE-HPP framework can be applied to study the link between sociality and cognition[4]. Specifically, as the human mind has culturally informed, motivationally powered, emotionally imbued and morally guided models of how people think, feel and behave in relationships, the framework provides implications for how human cognitive processes (for example, affects, motives and decisions) are adaptively configured and operated for different interpersonal contexts (see our discussions on the functionality of each FAVEE dimension and HPP category in Extended Data Fig. 8). This could help us understand why humans are able to form and maintain relationships that extend beyond immediate family members and how humans evolved from 'animal' (that is, no cooperation, but hostility towards others) to 'social animal' (that is, small-scale cooperation based on private relationships) and finally to 'cultural animal' (that is, large-scale cooperation based on public relationships)[23]. From a practical point of view, the FAVEE-HPP model builds a computational framework that can objectively and quantitatively measure human relationships at a high level of granularity. It provides a standard frame of reference that can be used to optimally design, manipulate, control and model interpersonal factors in relationship science, similar to the role of the 'Big Five' framework for personality science.

Our research has provided concrete evidence that relationship understanding is both universal and culturally variable. We demonstrated that the global architecture (backbone) of relationship representations (that is, FAVEE and HPP) is universally shared across cultures, but local fine-grained representational geometries (for example, the concept of 'neighbour' in Extended Data Fig. 6b), which are malleable by culture, could be quantitatively different. Computational modelling further suggests that the variations among modern societies are associated with religion and modernization, and the major difference between ancient and modern relationship concepts might come from changes in formality (for example, public/private boundaries) and equality (for example, social stratification). The universality and cultural variability of relationship conceptualization have wide-ranging implications for science and society. For example, a detailed delineation and elaboration of cross-cultural similarities and differences in relationships could inform whether human relations in languages, laws, social policies, moral codes and ideologies are equivalent in different countries and eras, whether relationships are expected to have the same impact on health and happiness across cultural groups and generations[3], or whether the same artificial intelligence algorithms can be applied to decode interpersonal relations via daily conversations and videos around the globe[24]. Understanding the role of culture in relationships could also contribute to cross-cultural adaptations of communications, literature, film, art, social networking, dating and marriage[25] and could facilitate efforts at cross-cultural diplomacy and commerce in a rapidly globalizing society[26].

Since the FAVEE-HPP model situates relationships as a fluid construct that is permitted to freely vary within and between people, future research could investigate how relationship representations are constructed during human development and how we form idiosyncratic impressions on relationships. It has been suggested that humans begin to accrue cognitive heuristics and stereotypes about interpersonal relations at birth[27] (for example, stranger danger and respect for elders), and thus relationship dimensions and categories could be gradually built via a combination of explicit instruction, indirect observation and personal experience. For example, infants' caregivers may introduce and transmit information about human relationships through bedtime stories, and preverbal infants learn basic dimensions such as valence (friends versus foes) and activeness (family versus outsiders). Later, they have social learning opportunities through indirect observation and direct experiences with others and understand new dimensions such as activeness (old versus new friends), equality (for example, teachers–peers) and exchange (for example, seller–buyer). In adulthood, acculturation to a new society involves learning the host culture's social norms and rules when interacting with local people. In addition, the present work investigated relationship conceptualization at only the cultural and population levels. It is apparent that cognition about relationships is subjective, varied and dynamic at the individual level, and how people think about relationships might vary depending on salient features in the contexts. The FAVEE dimensions and HPP categories could function as cognitive maps to help individuals navigate social environments (such as a 'relationship compass') and set standards to determine the satisfaction and stability of a relationship[28,29]. For example, individuals who grew up in a family with challenges and had chronic peer rejection might form negative impressions about familial and affiliative relationships (for example, with negative scores in valence and activeness). Likewise, individuals who had harmonious experiences with employers, clients or co-workers might adopt more positive views on public relationships (for example, with positive scores in valence and equality). The FAVEE-HPP framework establishes relatively objective and quantitative measures of relationships that can be compared across contexts, individuals and groups. Future research could use the framework to develop psychometric tests to measure where an individual lies on the spectrum of each of the five dimensions (like the Big Five personality test) and quantitatively examine how individual differences in relationship representations are linked to interpersonal difficulties in daily life[30] and whether relationship representations are abnormal in clinical populations (for example, those with autism or sociopathy).

The present work features replication and generalization. We attempted to extend and improve on prior work by being more comprehensive in several aspects, including preregistering our studies, using high-powered samples, including diverse types of relationships, analysing data with different tools and algorithms, and replicating representational models across different cultures (contemporary industrial societies, ancient societies and matrilineal tribes) and interpersonal contexts (dyadic, triadic and group relations). We also quantified the robustness of all results and showed that a subset of 40 relationships was good enough to replicate all findings based on 159 relationships (Supplementary Fig. 11).

However, our work also has several limitations. First, the mental representations of relationships are an organized body of information that reflects values, rules, concepts, scripts, affects, motives, expectations and memories associated with a relationship. The present work only taps the lay theory (that is, vernacular beliefs), which may differ from the actual organization of relationships in human society[31]. Future work needs to examine the social acts and interactions across relationships. Second, FAVEE-HPP as the universal structure of relationships is far from conclusive. The present work primarily used online populations and data-driven approaches, which was a double-edged sword. More data and investigations are needed to explore factors or boundary conditions that could influence the stability, validity, representativeness and generalizability of the FAVEE-HPP model. For simplicity and convenience, we chose the acronym FAVEE as the name for our model, but the global data showed that formality is not always the most important dimension. The different ordering of dimensions in different regions requires further investigation as it could reveal interesting cultural differences. Third, the FAVEE-HPP model was decomposed from many theoretical features originating from layperson languages. A more scientifically rigorous approach is needed to create a valid and reliable taxonomy of human relationships. Fourth, due to limited resources of ancient culture experts and high-quality PLMs, Study 3 examined relationship representations only in ancient China. Future research is encouraged to validate the FAVEE-HPP model in other historical contexts (for example, in Hebrew, Greek, Tamil and Old English).

# Methods

## Participants

All studies in this report were approved by the Institutional Review Board of Beijing Normal University (IRB_A_0024_2021002), and informed consent was obtained from all participants. Study 1 recruited 1,065 online US participants via MTurk and 60 offline US participants. Study 2 was preregistered (https://osf.io/swr2c) and recruited 17,686 online participants across 19 global regions via MTurk, CloudResearch, Credamo and the NaoDao platform[32,33]. In addition, 229 native Mosuo people were recruited from Yongning Township (Yunnan Province, China), using a field research data collection style (that is, through face-to-face interviews and door-to-door paper surveys). Study 3 recruited 44 scholars specialized in ancient Chinese culture for expert evaluation of the NLP method. Moreover, to test the FAVEE-HPP model in non-dyadic relationships, we recruited 380 online US participants (via MTurk) and 242 online Chinese participants (via the NaoDao platform). Participants across all studies were native speakers who grew up or lived for the longest period of their life in the targeted regions, with diverse demographics (Supplementary Fig. 13). The survey was translated into the local written language, and detailed guidelines for translation can be found at the Open Science Framework website. All participants received monetary compensation after completing the tasks.

Power analysis was performed to predetermine the sample size. To establish a design with adequate statistical power, we conducted a pilot study ($n = 721$, recruited from MTurk) using the dimensional survey from Wish et al.[10]. We collected at least 80 participant responses for each relationship on each evaluative feature, and the results of Wish et al.[10] were completely replicated (Supplementary Fig. 12). We ran a Monte Carlo simulation test to derive the minimally required responses in each condition to maintain a stable and consistent PCA result. PCA was performed on each subsample (from 2 to 40, with 1,000 iterations for each subsample), and loading scores and relationship scores were compared with the overall dataset using Pearson's correlation. The simulation results (Supplementary Fig. 12c) indicated that subsamples with ten responses were almost identical to the entire dataset (rating correlation $r > 0.95$) and thus should be adequate to ensure highly similar derived PCA components (loading score correlation $r > 0.90$; relationship score correlation $r > 0.95$).

## Sampling of human relationships

A data-driven approach based on NLP was used to generate a comprehensive list of human relationships (see Supplementary Method 1 for the details). Seed words were created via brainstorming and social media searches by a set of participants ($n = 15$ for the USA and $n = 27$ for China). Text embedding was used to find high-co-occurrence words relating to seed words by calculating the cosine distance between word vectors. The list of words was filtered to leave only nouns. Next, the list was filtered for frequency and was manually checked to keep only words related to human relationships. Finally, we paired the words on the basis of the meaning of relationships and added relationships that were pulled from the literature, resulting in the final relationships word list (159 for the USA and 258 for China). See further methodological details in Supplementary Figs. 1 and 8 and the full list of 159 English relationships and 258 Chinese relationships in Supplementary Tables 2 and 4.

## Evaluative features

A comprehensive literature search was performed to find all relevant theories and models that were proposed to explore the basic forms of human relationships. Thirty conceptual features were summarized and extracted from 15 prominent theories in Study 1. Redundant features were combined across theories (see Extended Data Fig. 1 and Supplementary Table 1 for the details). Note that many of these theoretical features were originally derived from dimensionality reduction or clustering techniques, but here they were prepared to be further reduced into higher-order components. Study 2 added three extra theoretical features (morality, trust and generation gap) from the cross-cultural literature[34–36], so a total of 33 features were evaluated.

### Dimensional survey

The participants completed an online survey where they rated human relationships on bipolar Likert scales. At the top of each page, the participants were cued to rate relationships on a given evaluative feature (for example, activeness), along with two phrases on opposite ends of a presented slider bar (for example, passive versus active). These two phrases represented the opposite ends of the bipolar features. Participants moved the slider towards the phrase that they felt best related to the presented relationships. Since certain features were quite obscure (for example, communality and reciprocity), we presented each feature with a detailed definition plus an exemplary relationship in the survey (Supplementary Table 1). Once the participants confirmed their understanding of each feature, they moved to the rating part. The participants were asked to consider all aspects of the relationships, including the way the individuals in each relationship typically think and feel about each other, how they act and react towards each other, how they talk and listen to each other, and any other characteristics of the relationships that occurred to them. The participants were instructed to focus not on their personal experiences with a specific relationship but rather on their general knowledge (that is, common sense or stereotypical understanding) about such relationships. Attention-check questions were used to ensure that the online participants were actively engaged in the survey and not answering questions in specific patterns or answering randomly. To avoid potential fatigue and inattentiveness, a between-subject design was used for all online participants to keep the survey short and effective (~20 min). Each participant was randomly assigned to a subset of relationships (for example, five to eight relationships) and had to rate them on a subset of evaluative features (for example, 10–11 features). To replicate the results from the between-subject design, a within-subject design was adopted for offline participants in Study 1, where each participant was asked to rate all relationships on all features in the laboratory (which took them three hours to complete). To rule out the effects of cross-cultural variations in online data quality and general semantic knowledge, the participants were asked additional questions on the size and colour of common objects (for example, animals, fruits, vehicles, tools and outdoor scenes). We found very low cross-regional variations in this object knowledge (pairwise correlations were >0.991), and there is no evidence that it can predict the cross-regional variation in relationship understanding (all $P > 0.352$; Extended Data Table 1). The cultural variability reported in Study 2 thus seems to be unique to relationship concepts, not merely arising from the variability of general semantic knowledge or data quality differences across global regions.

### Cognitive tasks

Along with the dimensional survey, two laboratory cognitive tasks were implemented to measure the categorization of relationship concepts. The multi-arrangement task is a behavioural paradigm to collect intuitive similarity judgements on semantic concepts[12]. The participants were asked to 'arrange the 159 relationships according to their similarity' in a 2D circle on a computer screen via mouse drag-and-drop so that similar relationships were placed close together, and dissimilar ones were placed further apart. The free sorting task asks participants to deliberately classify the 159 relationships into labelled categories[13]. They were allowed to make as many groupings as they liked (up to eight). Both tasks were conducted via the Meadows platform and the Naodao platform.

Text analysis was performed on the categorical labels assigned by participants in the free sorting task. Initially, 444 labels were obtained, and they were coded by assigning 159 relationships (444 × 159 matrix). For example, the 'family' label was assigned to 'wife–husband' but not 'doctor–patient', so the former was coded as 1 and the latter was coded as 0. Hierarchical clustering (the Ward method) was performed on the label × relationship matrix. After a noisy cluster containing miscellaneous labels was excluded, three and six clusters were observed on the remaining 292 labels.

### Dimensionality reduction and clustering

Python (v.3.9.1) was used to clean and organize all data. Any participants who did not pass the attention check were excluded from the analysis (the data exclusion criteria can be found in the preregistration at https://osf.io/swr2c). On the basis of this criterion, 129 participants (out of 721; 17.89%) were excluded in the pilot study, 248 participants (out of 1,065; 23.29%) were excluded in Study 1 and 2,441 participants (out of 18,537; 13.17%) were excluded in Study 2. Before applying any dimensionality reduction or clustering, we created a matrix from the average ratings of each relationship on each evaluative feature across participants. This matrix was normalized by using the preprocessing command from the scikit-learn package (v.1.4.2).

PCA was adopted as the primary dimensionality reduction technique to derive all dimensional models (using the prcomp function from R (v.4.3.3)). A varimax rotation was used for individual evaluative features to load maximally onto the components. Since the PCA does not provide labels for the components, we named the components by considering both the top five highest loadings (absolute value) and the distribution of relationship scores. To determine the optimal number of PCA components, we checked four data-driven metrics (that is, parallel analysis, the Kaiser–Guttman rule, Cattell's scree test and optimal coordinates) and examined the interpretability of each component (Extended Data Fig. 2a). Solutions with cross-metrics agreement and high interpretability were chosen. We also implemented other dimensionality reduction techniques to validate the PCA results (see Supplementary Fig. 3 for the details), and five identical components were observed using independent component analysis, exploratory factor analysis, multidimensional scaling and network analysis.

We adopted $k$-means clustering as the primary clustering technique to derive categorical models, although other clustering techniques (such as hierarchical clustering and HDBSCAN) were also conducted to validate the $k$-means results. A dissimilarity matrix of 159 relationships was prepared as input. For the dimensional survey, the Euclidean distance matrix was calculated using relationships' ratings on all evaluative features. For the multi-arrangement task, the distance matrix was retrieved from the original data, in which the value indicated two relationships' distances in the 2D circle. For the free sorting task, the distance matrix was calculated on the basis of the probability that two relationships were classified in the same category. Uniform manifold approximation and projection (UMAP) was used as a preprocessing step to boost the performance of $k$-means clustering, given that this method is flexible and powerful in finding and balancing the local and global structure of the data. Two UMAP parameters were manually set: the nearest neighbour parameter (which determines how much of the local versus global structure to consider) and the minimum distance value (which determines how closely together the data points should be in the final solution). A low-to-medium value (15) for the nearest neighbour and a low value for the minimum distance (0.01) were selected, as they can effectively produce tighter clusters that are easier to process for the subsequent clustering algorithms[37]. To determine the optimal number of clusters, the silhouette score was considered (Supplementary Fig. 15), and the stability and interpretability of the output clusters were also examined. Solutions that were insensitive to algorithm/parameter choice, were consistent across different clustering algorithms, and had high interpretability and high silhouette scores were chosen.

### Language models and embeddings

We used PLMs and LLMs to probe ancient people's perception and comprehension of human relationships. For the modern Chinese PLM, we employed the word-based Chinese-RoBERTa-Base model from UER-py Modelzoo[38]. We selected this model due to its focus on the mask language modelling task during the pretraining phase. Moreover, it takes into account the characteristics of the Chinese language by using words rather than characters as units, and it has been trained on a large-scale, publicly available corpus of modern Chinese text. For the ancient Chinese PLM, we used BERT-ancient-Chinese[22], which was trained on a large-scale ancient Chinese corpus including historical texts from 1046 BCE to 1912 CE.

We adopted an approach to generate human-like PLM embeddings (Fig. 5a), which was previously proposed by Cutler and Condon[19] to identify Big Five personality structures in language models. We compared different queries and layers of embeddings (Supplementary Fig. 9). The [DESC] component in the query was generated by GPT-4 in October 2023 with the temperature parameter set to zero to ensure reproducibility (see exemplar prompts in Supplementary Method 2). Details of the labels and descriptions for ancient and modern Chinese relationships can be accessed via the Open Science Framework website.

### RSA and model comparison

To uncover which cultural variables account for the cross-cultural variance in relationship representations, we performed RSA multiple regression[39] (Fig. 3b). For each global region, cultural variables of language, personality, socio-ecology (that is, subsistence style, historical disease prevalence and climates), modernization, genetics, religion, politics and the Hofstede 6D culture model were collected from multiple open databases, such as the World Values Survey, Timeanddate and Worldbank (see Supplementary Table 3 for the details). For each cultural variable (for example, modernization), an RDM was computed where each cell represents the dissimilarity of two regions on this variable (for example, the dissimilarity of China and Portugal according to their modernization level). For each representational geometry (that is, full-feature, dimensional or categorical), we also created an RDM to represent the dissimilarity of relationship representations across regions. We then performed a linear regression model in which cultural variable RDMs were predictors, and relationship representational geometry RDM was the outcome variable. The noise ceiling was estimated using the mean relationship RDMs of $n-1$ regions to predict the relationship RDM of the remaining region, which reflected the inherent heterogeneity of the relationship RDMs. The Mantel test was used to assess the statistical significance of each RSA[40,41]. We permuted the order of RDMs of cultural variables while holding the representational geometries constant, recalculated the regression and repeated the process 10,000 times. This test allowed us to compute a $P$ value for the representational geometries based on the $F$ statistic of the multiple regression. We performed a one-sided test since a negative value is not meaningful and only positive similarities are expected[20,42].

Study 3 implemented RSA correlations between language models and the human-rating FAVEE-HPP model. Specifically, we transformed PLM embeddings ($258 \times 768$ or $120 \times 768$ matrix) into a cosine similarity matrix ($258 \times 258$ or $120 \times 120$). This matrix was then correlated with the lower triangle of the RDMs derived from the FAVEE dimensions (which represents the distances between pairs of relationships in 5D FAVEE space) or RDMs from the HPP categories using Spearman correlation. The noise ceiling was estimated by correlating human-rating RDMs derived from the FAVEE-HPP model with human-rating RDMs from 33 dimensional features (Fig. 5d).

### Robustness test

The robustness test across different numbers of relationships was quantified using the same method as Lin et al.[43]. We removed human relationships one by one and reperformed all analyses (for example, PCA, clustering and cross-cultural RSA). The sequence to remove relationships was implemented as follows: all pairs of relationships were ranked from the most to the least similarity in the multi-arrangement task, and the relationship with the lower familiarity rating was removed first from each pair. Pearson correlations were calculated between metrics from the full set and from the subsets to determine the robustness of the results (see Supplementary Fig. 11 for the details).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All data used in this project are accessible via GitHub (https://github.com/BNU-Wang-MSN-Lab/FAVEE-HPP) and deposited in the Open Science Framework (https://osf.io/nfkmj) and can be interactively viewed and freely downloaded at a dedicated website (https://bnu-wang-msn-lab.github.io/FAVEE-HPP). A supplementary video is also provided to elaborate the FAVEE-HPP framework (https://insula.oss-cn-chengdu.aliyuncs.com/favee/FAVEE-HPP.mp4).

## Code availability

All data analysis code is available via GitHub (https://github.com/BNU-Wang-MSN-Lab/FAVEE-HPP).

## References

1. Fiske, A. P. The four elementary forms of sociality: framework for a unified theory of social relations. *Psychol. Rev.* **99**, 689 (1992).
2. Reis, H. T., Collins, W. A. & Berscheid, E. The relationship context of human behavior and development. *Psychol. Bull.* **126**, 844 (2000).
3. Holt-Lunstad, J. Why social relationships are important for physical health: a systems approach to understanding and modifying risk and protection. *Annu. Rev. Psychol.* **69**, 437–458 (2018).
4. Fiske, A. P. & Haslam, N. Social cognition is thinking about relationships. *Curr. Dir. Psychol. Sci.* **5**, 143–148 (1996).
5. Cacioppo, J. T. & Hawkley, L. C. Perceived social isolation and cognition. *Trends Cogn. Sci.* **13**, 447–454 (2009).
6. Sapolsky, R. M. The influence of social hierarchy on primate health. *Science* **308**, 648–652 (2005).
7. Rai, T. S. & Fiske, A. P. Moral psychology is relationship regulation: moral motives for unity, hierarchy, equality, and proportionality. *Psychol. Rev.* **118**, 57–75 (2011).
8. Argyle, M., Henderson, M. & Furnham, A. The rules of social relationships. *Br. J. Soc. Psychol.* **24**, 125–139 (1985).
9. Marwell, G. & Hage, J. The organization of role-relationships: a systematic description. *Am. Sociol. Rev.* **35**, 884–900 (1970).
10. Wish, M., Deutsch, M. & Kaplan, S. J. Perceived dimensions of interpersonal relations. *J. Pers. Soc. Psychol.* **33**, 409–420 (1976).
11. Montgomery, B. M. in *Handbook of Personal Relationships: Theory, Research and Interventions* (ed. Duck, S.) 343–359 (Wiley, 1988).
12. Kriegeskorte, N. & Mur, M. Inverse MDS: inferring dissimilarity structure from multiple item arrangements. *Front. Psychol.* **3**, 245 (2012).
13. Burton, M. L. & Romney, A. K. A multidimensional representation of role terms. *Am. Ethnol.* **2**, 397–407 (1975).
14. Freeman, J. B., Stolier, R. M., Brooks, J. A. & Stillerman, B. S. The neural representational geometry of social perception. *Curr. Opin. Psychol.* **24**, 83–91 (2018).
15. Kriegeskorte, N. & Kievit, R. A. Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn. Sci.* **17**, 401–412 (2013).
16. Greenfield, P. M. Linking social change and developmental change: shifting pathways of human development. *Dev. Psychol.* **45**, 401–418 (2009).

17. Jackson, J. C. et al. From text to thought: how analyzing language can advance psychological science. *Perspect. Psychol. Sci.* **17**, 805–826 (2022).
18. Grand, G., Blank, I. A., Pereira, F. & Fedorenko, E. Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nat. Hum. Behav.* **6**, 975–987 (2022).
19. Cutler, A. & Condon, D. M. Deep lexical hypothesis: identifying personality structure in natural language. *J. Pers. Soc. Psychol.* **125**, 173–197 (2023).
20. Thornton, M. A., Wolf, S., Reilly, B. J., Slingerland, E. G. & Tamir, D. I. The 3D mind model characterizes how people understand mental states across modern and historical cultures. *Affect. Sci.* **3**, 93–104 (2022).
21. Demszky, D. et al. Using large language models in psychology. *Nat. Rev. Psychol.* **2**, 688–701 (2023).
22. Wang, P. & Ren, Z. The uncertainty-based retrieval framework for ancient Chinese CWS and POS. Preprint at https://doi.org/10.48550/arXiv.2310.08496 (2023).
23. Baumeister, R. F. *The Cultural Animal: Human Nature, Meaning, and Social Life* (Oxford Univ. Press, 2005).
24. Cowen, A. S. et al. Sixteen facial expressions occur in similar contexts worldwide. *Nature* **589**, 251–257 (2021).
25. Wood, A., Kleinbaum, A. M. & Wheatley, T. Cultural diversity broadens social networks. *J. Pers. Soc. Psychol.* **124**, 109–122 (2022).
26. Lindquist, K. A., Jackson, J. C., Leshin, J., Satpute, A. B. & Gendron, M. The cultural evolution of emotion. *Nat. Rev. Psychol.* **1**, 669–681 (2022).
27. Hamlin, J. K., Wynn, K. & Bloom, P. Social evaluation by preverbal infants. *Nature* **450**, 557–559 (2007).
28. Clark, M. S. & Reis, H. T. Interpersonal processes in close relationships. *Annu. Rev. Psychol.* **39**, 609–672 (1988).
29. Fiske, A. P. & Haslam, N. in *Making Friends: The Influences of Culture and Development* (eds Meyer, L. H. et al.) 385–392 (Paul H. Brookes, 1998).
30. Robins, G. & Boldero, J. Relational discrepancy theory: the implications of self-discrepancy theory for dyadic relationships and for the emergence of social structure. *Pers. Soc. Psychol. Rev.* **7**, 56–74 (2003).
31. Fiske, A. P. The lexical fallacy in emotion research: mistaking vernacular words for psychological entities. *Psychol. Rev.* **127**, 95 (2020).
32. Litman, L., Robinson, J. & Abberbock, T. TurkPrime.com: a versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behav. Res. Methods* **49**, 433–442 (2017).
33. Chen, G. et al. Online research in psychology and its future in China. *J. Psychol. Sci.* **46**, 1262–1271 (2023).
34. Zhang, Z. X. Perceiving relationships with others: the Chinese dimensions. *Indig. Psychol. Res. Chin. Soc.* **261**, 288 (1999).
35. Lau, K. H., Au, W. T., Lv, X. & Jiang, Y. The conceptualization of Chinese guanxi of Hong Kong University students by using multi-dimensional scaling: an empirical approach. *Acta Psychol. Sin.* **37**, 122–125 (2005).
36. Wang, R. & Sang, B. A study on the characteristics and development of children's conceptual structure of social relations. *J. Psychol. Sci.* **33**, 1344–1351 (2010).
37. McInnes, L., Healy, J. & Melville, J. Umap: uniform manifold approximation and projection for dimension reduction. Preprint at https://doi.org/10.48550/arXiv.1802.03426 (2018).
38. Zhao, Z. et al. UER: an open-source toolkit for pre-training models. Preprint at https://doi.org/10.48550/arXiv.1909.05658 (2019).
39. Stolier, R. M., Hehman, E. & Freeman, J. B. A dynamic structure of social trait space. *Trends Cogn. Sci.* **22**, 197–200 (2018).
40. Popal, H., Wang, Y. & Olson, I. R. A guide to representational similarity analysis for social neuroscience. *Soc. Cogn. Affect. Neurosci.* **14**, 1243–1253 (2019).
41. Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis—connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
42. Finn, E. S. et al. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. *NeuroImage* **215**, 116828 (2020).
43. Lin, C., Keles, U. & Adolphs, R. Four dimensions characterize attributions from faces using a representative set of English trait words. *Nat. Commun.* **12**, 5168 (2021).
44. Harari, Y. N. *Sapiens: A Brief History of Humankind* (Random House, 2014).

## Acknowledgements

We thank J. Zhou, J. Yang, J. Li, Y. Wu, C. Xu, C. Mou and W. Lin for assistance in the entire project; K. Yao for help with global survey data collection; X. Zhang for help with ancient Chinese literature; 阿七尼玛次尔 (杨成龙) (NimaCier Aqi), 阿七独支玛 (DuzhiMa Aqi), 杨剑明 (J. Yang) and 和智利 (Z. He) for help with Chinese Mosuo data collection; T. Talhelm and M. Muthukrishna for sharing global cultural data; and A. P. Fiske, N. Haslem, D. M. Condon, S. Han, C. Lin, J. Zhang and X. Wang for valuable discussions and comments. This work was supported by the National Natural Science Foundation of China (grant nos 32422033, 32171032, 32430041, 32371139, 32000776 and 31925020), the National Science and Technology Innovation 2030 Major Program (grant nos 2022ZD0211000, 2021ZD0200500 and 2021ZD0204104), Fundamental Research Funds for the Central Universities (grant no. 2233300002), the Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning (grant no. CNLZD2103), the National Social Science Foundation of China (grant no. 21BZX005), National Institute of Health grants (nos R01MH091113, R01HD099165, R21HD098509, 2R56MH091113-11 and R01DC013063) and the start-up funding from the State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, Beijing Normal University (to Y.W.).

## Author contributions

X.C., H.P., R.H. and I.R.O. designed and planned the research. X.C., H.W., M.Z. and Y.Z. performed the global data collection. X.C., H.P., H.W., R.H., M.Z. and Y.Z. analysed the data. X.C., H.P., H.W. and Y.Z. wrote the initial draft of the manuscript. R.H., M.A.T., Y.M., C.H., Y.B., J.R. and I.R.O. edited and reviewed the final manuscript. Y.W. conceived and supervised the project and was involved in all aspects of the research.

## Competing interests

The authors declare no competing interests.

## Additional information

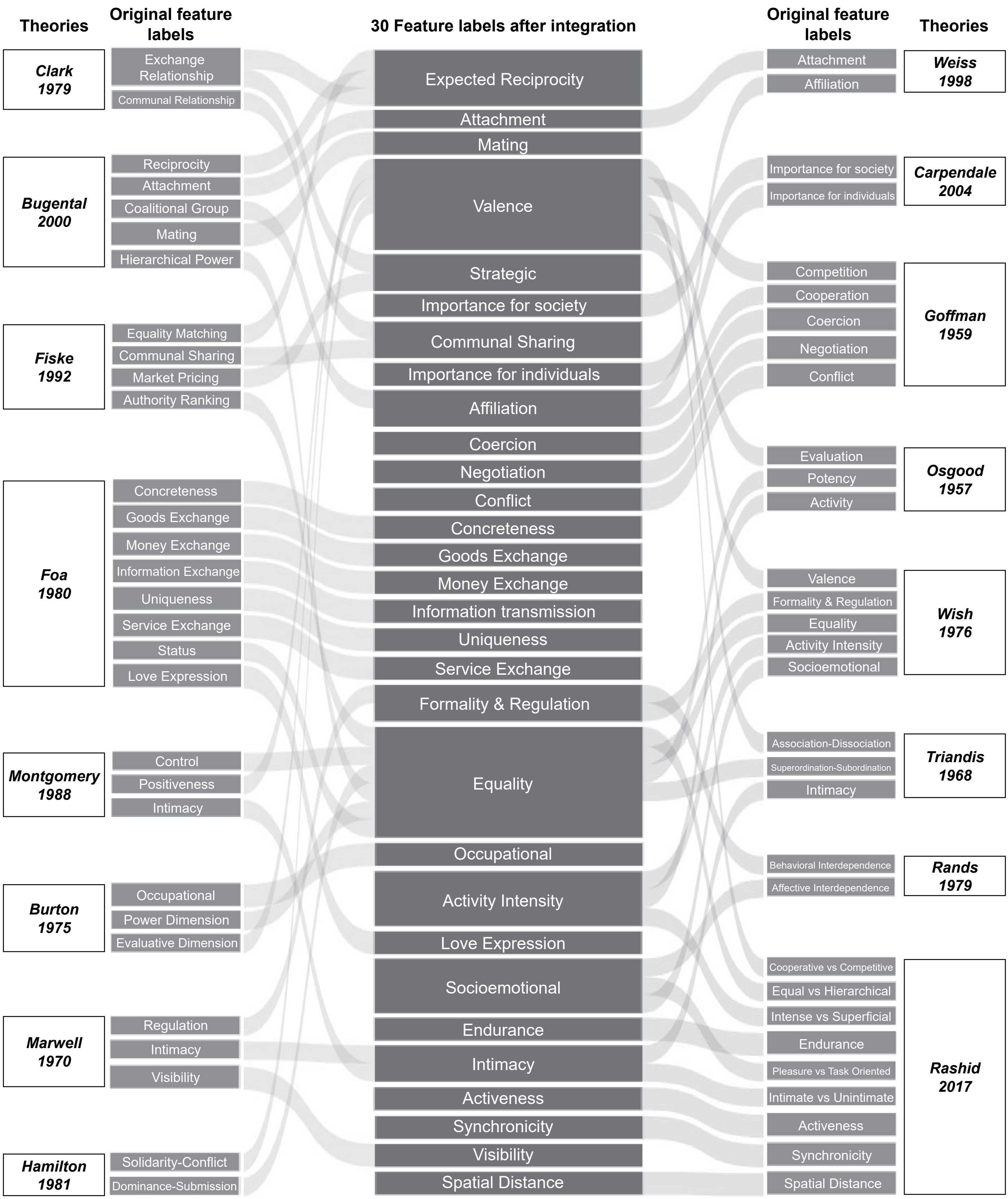
**Extended data** is available for this paper at https://doi.org/10.1038/s41562-025-02122-8.

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41562-025-02122-8.

**Correspondence and requests for materials** should be addressed to Yin Wang.

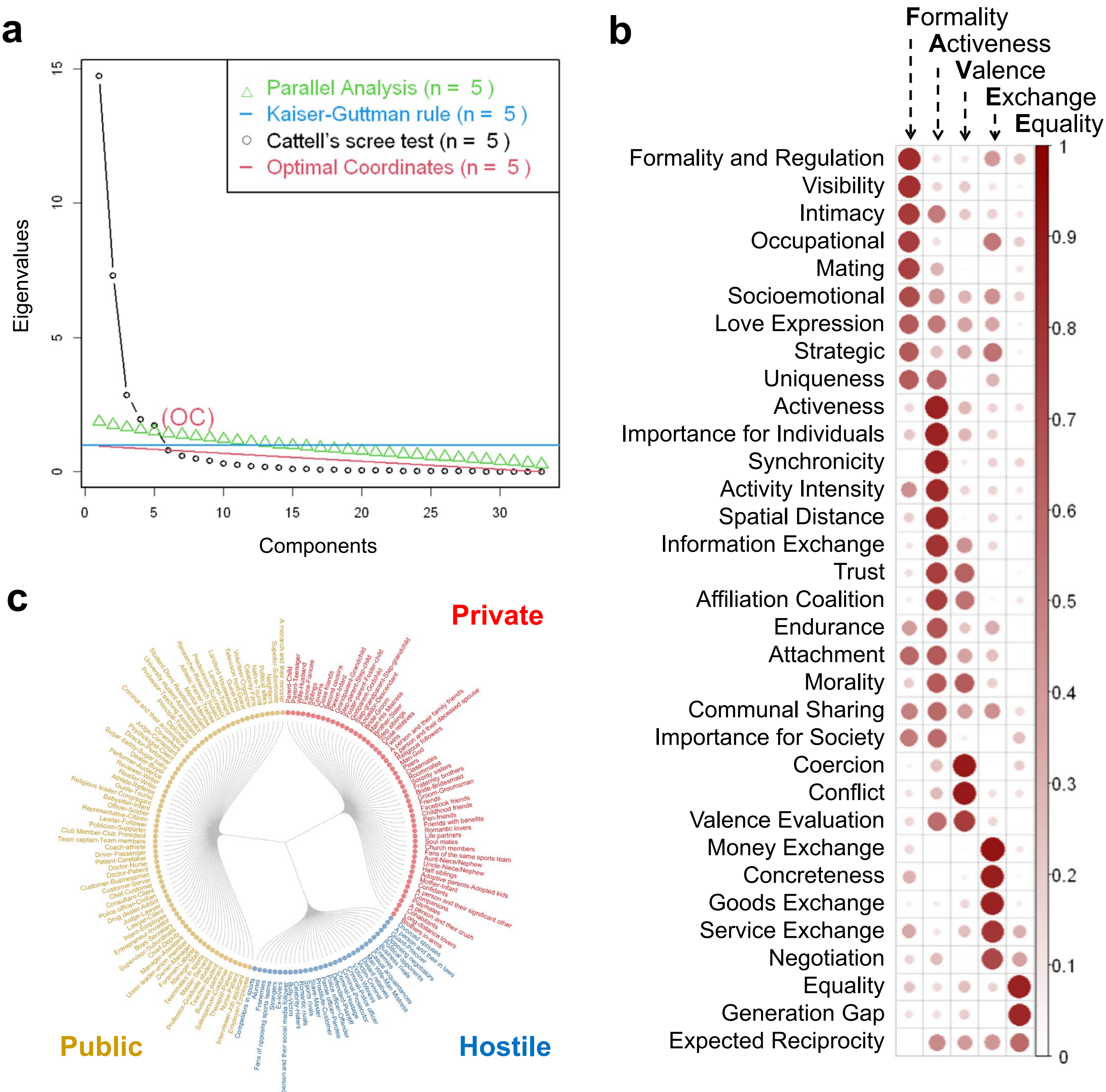
**Peer review information** *Nature Human Behaviour* thanks Alan Fiske, Christopher Cox, Daniel Hruschka and Joshua Jackson for their contribution to the peer review of this work.

https://doi.org/10.1038/s41562-025-02122-8

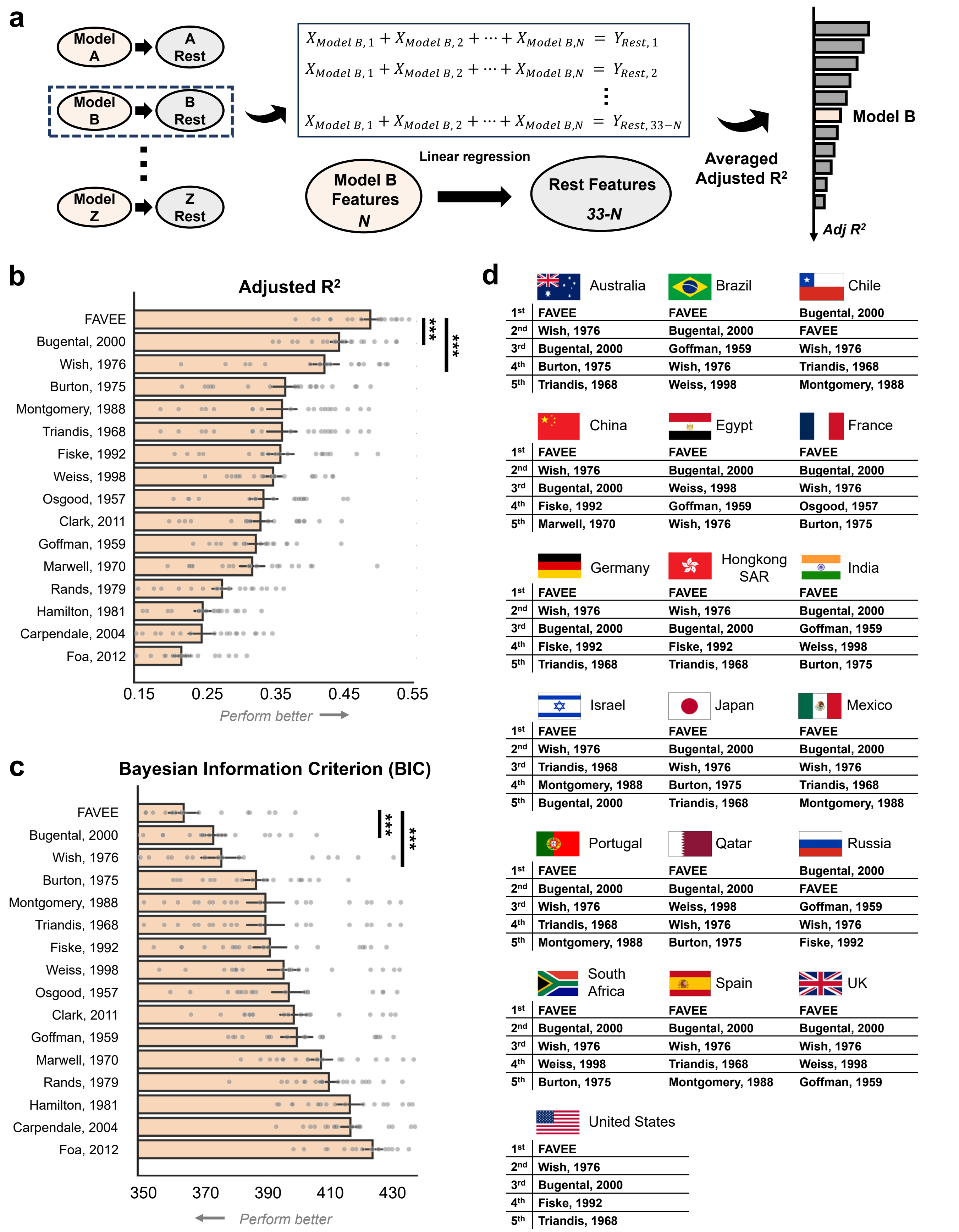



**Extended Data Fig. 1 | Existing theories in relationship science.** 30 conceptual features were summarized from 15 prominent theories. Redundant features were combined across theories (see central column for final 30 features). It seems that valence and equality were the most frequently mentioned features. Note that many of these theoretical features were originally derived from dimensionality reduction or clustering techniques, but here, they were prepared to be further reduced into higher-order components. Three extra theoretical features from cross-cultural literature (that is, morality, trust, generation gap) were added for Study 2, which were not listed here. Please see more detailed information about each feature in Supplementary Table 1.

https://doi.org/10.1038/s41562-025-02122-8



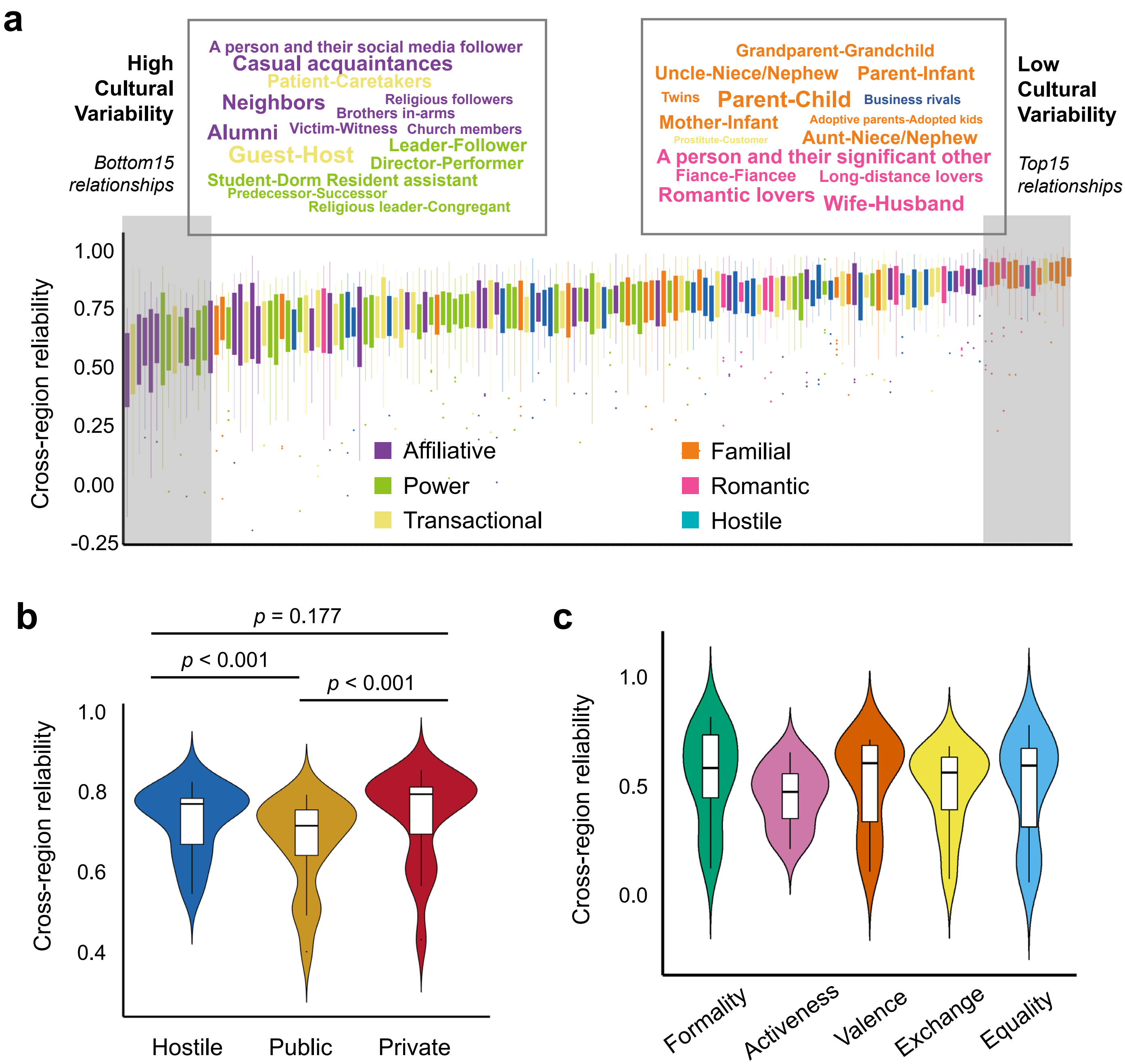

**Extended Data Fig. 2 | The dimensional and categorical models derived from global data (n = 17,686). a**, Four data-driven metrics consistently indicated that the optimal number for PCA was five. **b**, PCA loadings for five principal dimensions. **c**, K-means clustering solution identified three categories labelled as Hostile, Private, and Public, with the highest silhouette score of 0.295 at k = 3. These results suggested that the FAVEE-HPP model proposed in Study 1 can be well replicated by large-scale global data. In addition, for each global region, the same five dimensions and three clusters can also be identified (see Supplementary Fig. 4 and Supplementary Fig. 5).




**Extended Data Fig. 3 | See next page for caption.**

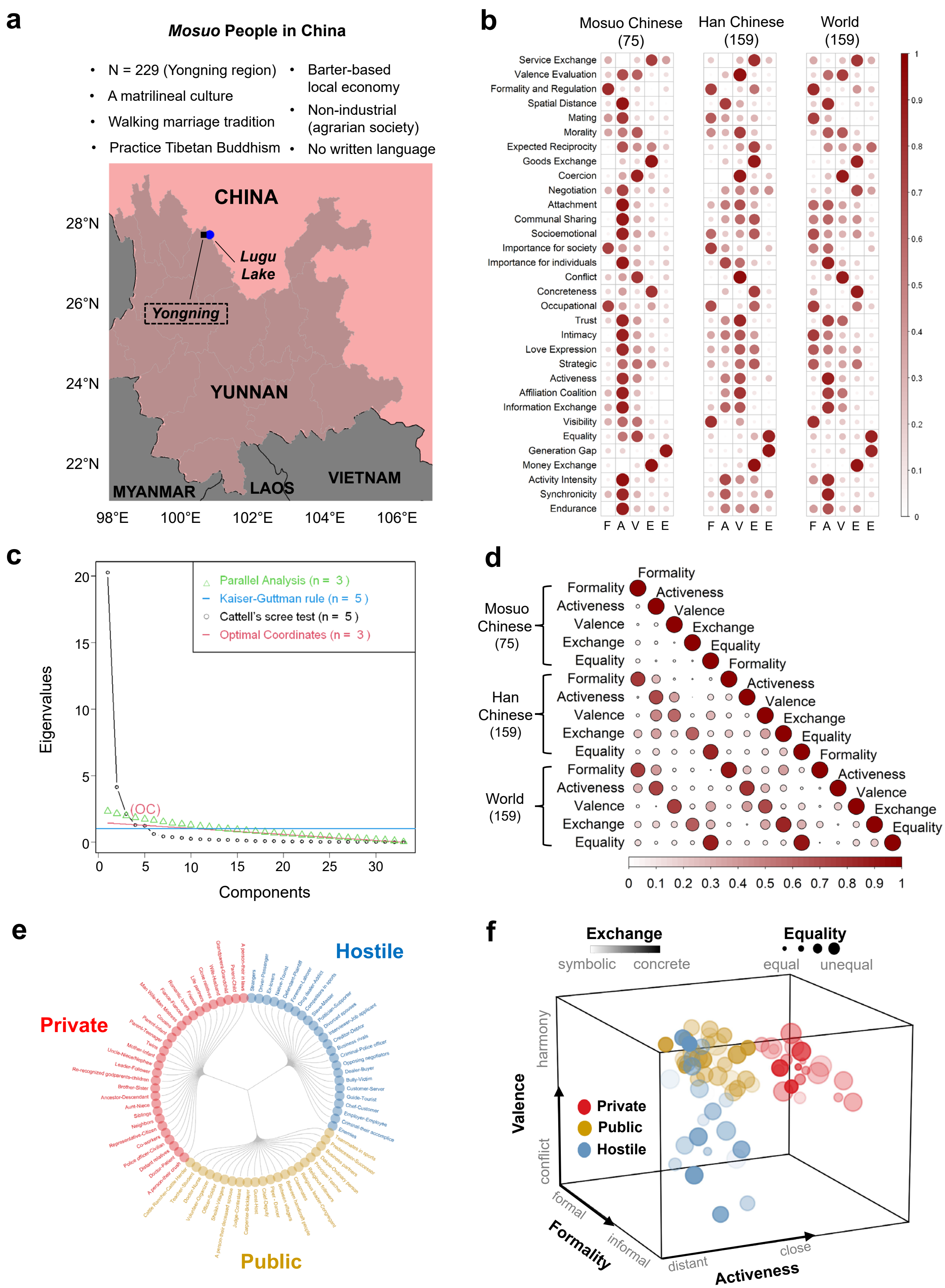
**Extended Data Fig. 3 | Model comparison in performance consistency across the globe.** **a**, To examine how well a model can represent all theoretical relationship features, we used linear combinations of features in each model as regressors to predict each of remaining theoretical features (that were not included in that model) and calculated adjusted $R^2$ for each region. **b** and **c**, Across global regions, FAVEE model (mean adjusted $R^2 = 0.489$, mean BIC = 364.794) outperformed other 15 existing theories in both explained variance and data fitting, with 100,000 bootstrap resamples used to estimate the mean differences (99.9% confidence interval). Error bar (standard error) represents performance variance across 19 regions. **d**, Top five models in each global region (FAVEE was the best in 17 out of 19). Note: A more stringent way of model comparison was attempted where the number of predicted features was controlled between two models, and similar results were found (see Supplementary Fig. 6). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

https://doi.org/10.1038/s41562-025-02122-8




**Extended Data Fig. 4 | Cultural variability across relationships, dimensions, and categories. a**, Cultural variability in 159 relationship concepts. The top 15 relationships with the lowest cross-region variability were primarily familial and romantic relationships (right word cloud box), whereas the bottom 15 relationships with the highest cross-region variability were mainly affiliative and power relationships (left word cloud box). Each bar represents a single relationship, and the order was arranged according to its cross-region reliability (that is, mean LOOCV across 19 global regions). The font size of word clouds proportionally reflects the familiarity of the relationships, which was uncorrelated with cultural variability (p > 0.994). **b**, For three HPP categories, public relationship concepts had more culturally variable meanings than hostile and private relationship concepts (ANOVA: $F(2, 36) = 41.113$, $P < 0.001$, $\eta_p^2 = 0.695$; post hoc, two-tailed paired-sample $t$-tests (Bonferroni's correction): 'Hostile:Private': $t(18) = -2.015$, $P = 0.177$; 'Hostile:Public': $t(18) = 7.329$, $P < 0.001$, Cohen's $d = 0.580$; 'Public:Private': $t(18) = -10.070$, $P < 0.001$, Cohen's $d = -0.698$). **c**, The cross-region reliability was comparable among the five principal dimensions, suggesting that no single dimension was selectively influenced by culture (ANOVA: $F(4, 72) = 2.053$, $P = 0.096$). All box plots show interquartile range (IQR) ± 1.5×IQR.

https://doi.org/10.1038/s41562-025-02122-8

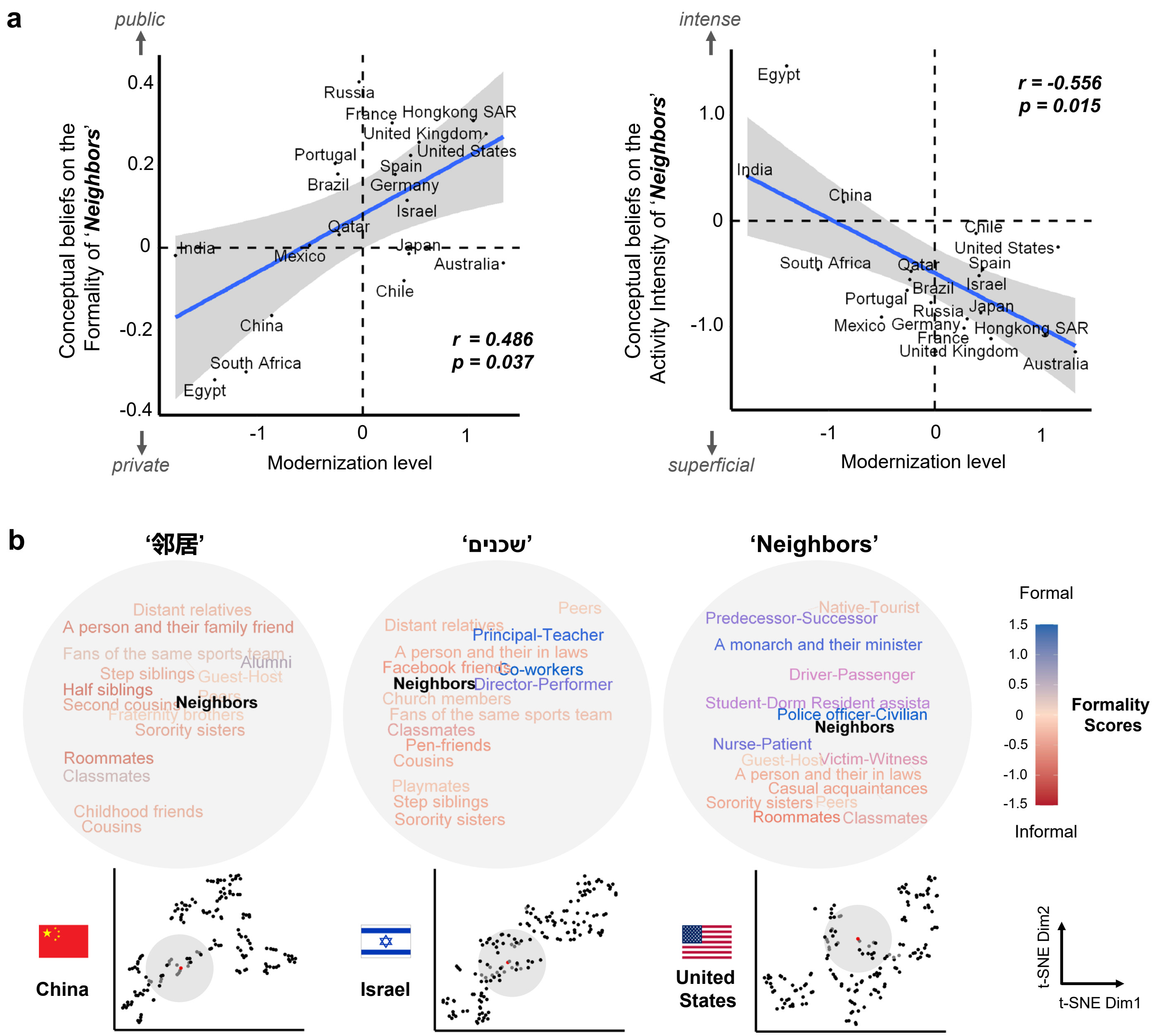



**Extended Data Fig. 5 | See next page for caption.**

https://doi.org/10.1038/s41562-025-02122-8

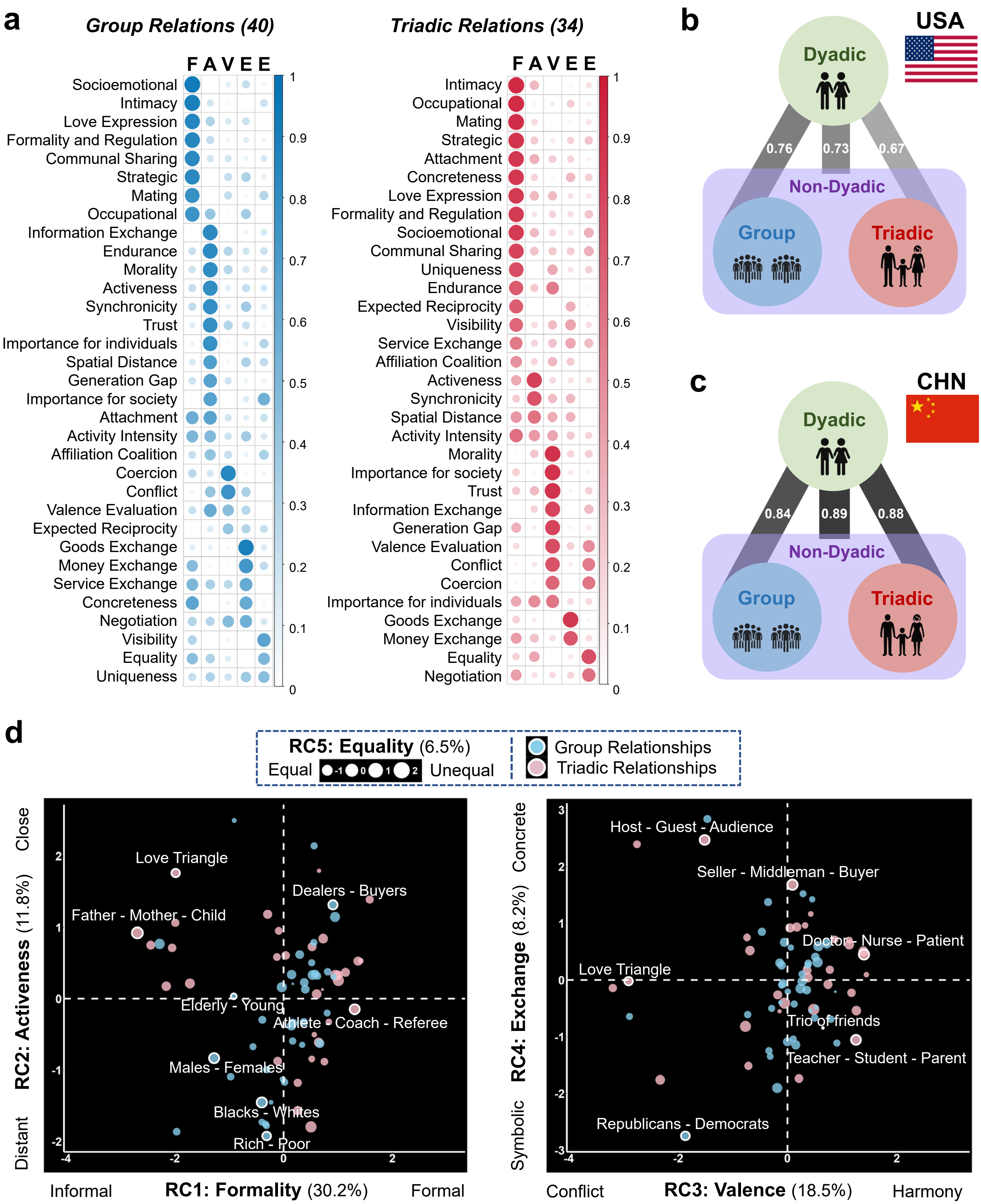
**Extended Data Fig. 5 | Model validation in a non-industrial society.** **a**, key features of the Mosuo society and its geographical location (dash line box). **b**, PCA showed identical FAVEE dimensions for Mosuo Chinese, Han Chinese, and world-averaged data. Through field work, we identified 75 typical relationships in Mosuo culture (see Supplementary Table 5). **c**, The optimal number of PCA components for Mosuo was five. **d**, Spearman's correlation of loading scores across three datasets. Their derived FAVEE dimensions were well corresponded. **e**, K-means clustering on Mosuo relationships identified the HPP model. **f**, A similar dimension-category hybrid model was observed in the Mosuo society, which replicated the findings in Study 1 (see Fig. 2c).

 https://doi.org/10.1038/s41562-025-02122-8




**Extended Data Fig. 6 | Conceptual differences in the word '*neighbours*' across the globe. a**, For each region, people's understanding (conceptual beliefs) of '*neighbours*' was estimated by interrogating its surrounding relationships in the semantic neighbourhood of representational space. Fifteen nearest relationship words were selected based on the smallest Euclidean distance with '*neighbours*' on all evaluative features. We found that a country's modernization level was positively correlated with the formality score of '*neighbours*' surrounding relationships but negatively correlated with the activity intensity score (Spearman correlation, two-tailed). This suggests that as a country's modernization level increases, '*neighbours*' become a more public, impersonal, and superficial relationship. The shaded area represents the 95% confidence interval. **b**, Taking China (middle level of modernization), Israel (high level modernization) and the US (highest level of modernization) as examples. All 159 relationships were plotted in the 2D t-SNE space, with the nearby 15 relationships zoomed in for better visualization. For China, only informal relationships (in red colour) surround the Chinese word '*neighbour*' ('邻居'), indicating that '*neighbours*' are considered private and personal relationships. However, for Israel and the US, an increasing number of public relationships (in blue colour) appear nearby, indicating that '*neighbours*' are conceptualized as being more formal and impersonal. Together, these results demonstrate that although translation dictionaries provide equivalent words of relationships in different languages, their conceptual meanings are not always the same. Their variations (at least for the concept '*neighbours*') were dependent on the country's level of modernization (for example, '*neighbours*' in large cities are often unknown to each other due to greater mobility led by urbanization). On the other side, these results illustrate how cultural factors such as modernization can deform the local representational geometries of relationship concepts.

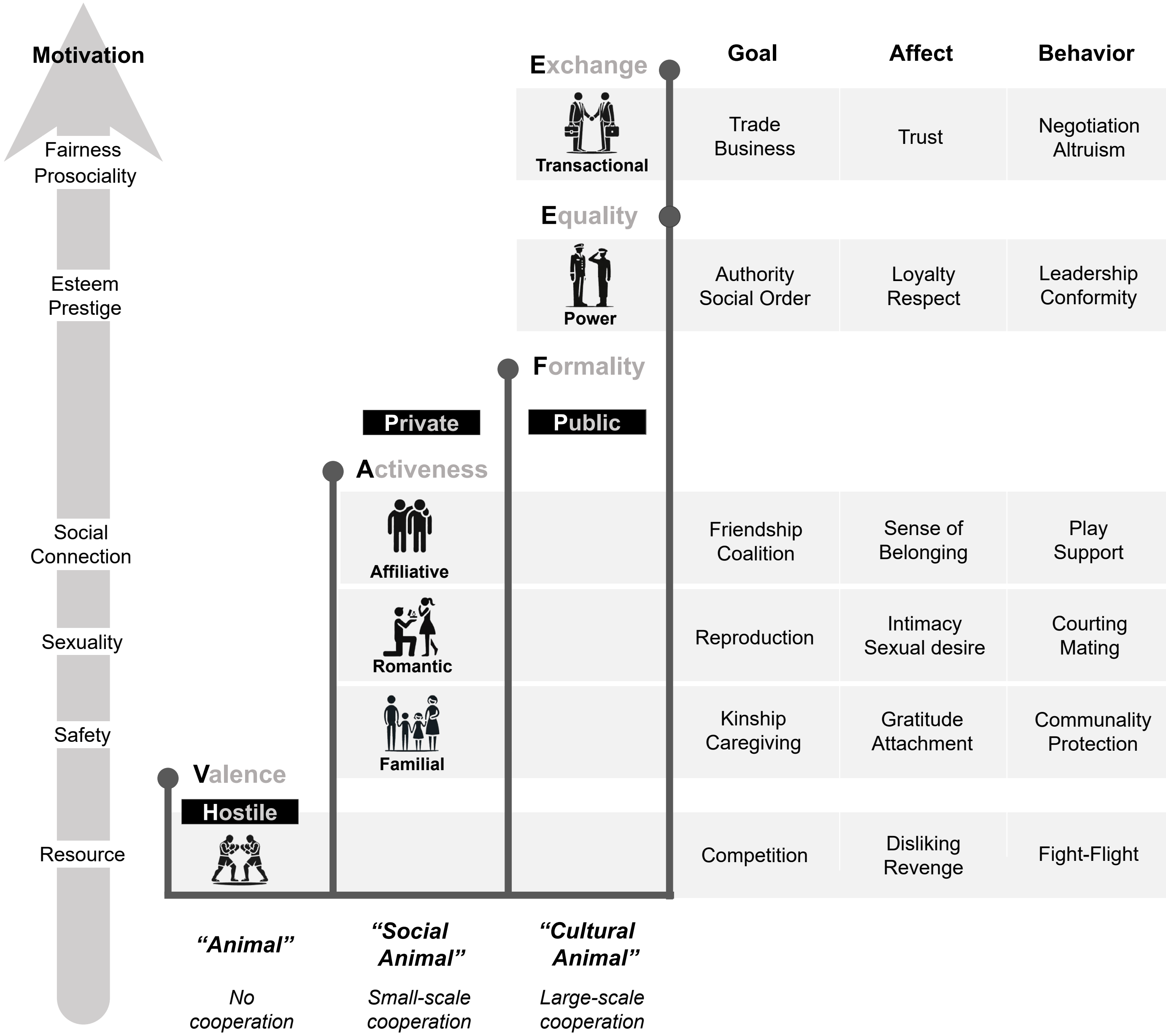
 https://doi.org/10.1038/s41562-025-02122-8




**Extended Data Fig. 7 | See next page for caption.**

https://doi.org/10.1038/s41562-025-02122-8

**Extended Data Fig. 7 | Good generalizability of the FAVEE model to non-dyadic relations in the US and China. a**, PCA loadings on 33 theoretical features for group relations and triadic relations in the US (n = 380). See Supplementary Table 7 for the full list of 40 group relations and 34 triadic relations. **b**, In general, there were high correlations of FAVEE structures between dyadic and non-dyadic relations in the US (r = 0.73). Within non-dyadic relations, dyadic relations also showed high correlations with group relations (r = 0.76) and triadic relations (r = 0.67). **c**, Similar results were observed in China (n = 242), with high correlations of FAVEE structures between dyadic and non-dyadic relations (r = 0.89). **d**, For illustration purpose, all group relations (blue) and triadic relations (red) in the US data were plotted in the 5D space based on their scores on each FAVEE dimension.




**Extended Data Fig. 8 | Functionality of FAVEE dimensions and HPP categories.** The human mind involves implicit cognitive models for forming and maintaining relationships ('relational schemas'), such as a shared understanding of the rules and norms governing interactions and the coordination of mental processes for social navigation and adaptation. The FAVEE-HPP framework posits that relationship concepts are primarily organized in a five-dimensional space with three default categories. These five dimensions might reflect different levels of motivations (for example, Maslow's hierarchy of needs, see left arrows), for example, valence for resource competition, activeness for emotional support and belongingness, formality for social order, equality for power, and exchange for fairness. Three categories might be configured for three levels of cooperation, which echoes Roy Baumeister's theory on how humans evolved from 'animals' (no cooperation, keeping hostile towards others), 'social animals' (small-scale cooperation based on private relationships) and to 'cultural animals' (large-scale cooperation based on public relationships)[23]. The three default HPP categories can be further classified into six canonical types of relationships, which are assumed to be associated with distinct goals, affects and behaviours. Circles and squares represent dimensions and categories, respectively. Please note that, although animals may have certain dimensions and categories, they are different from those of humans. For example, power in animals is typically defined by biological and behavioural characteristics (for example, body size, strength, vocalization), while high power in humans is often based on abstract symbols and cultural conventions (for example, reverence for elders and the divine[1]). Likewise, money creates a system of trust that enables exchange and cooperation between strangers in human society[44].

**Extended Data Table 1 | Multiple Regressions on Full-feature, Dimensional, and Categorical Models in Representational Similarity Analysis in Study 2 (significant results are in bold)**

| Main Analysis | Representational Geometry Models | | | | | |
|---|---|---|---|---|---|---|
| Predictors | Full-feature Model | | Dimensional Model | | Categorical Model | |
| | $\beta$ | $p$ | $\beta$ | $p$ | $\beta$ | $p$ |
| Climates | 0.163 | 0.180 | 0.167 | 0.200 | 0.146 | 0.154 |
| Demographics | -0.052 | 0.591 | -0.003 | 0.469 | -0.323 | 0.996 |
| Disease | 0.004 | 0.499 | -0.025 | 0.609 | -0.012 | 0.555 |
| Gene | -0.056 | 0.568 | -0.157 | 0.740 | -0.052 | 0.574 |
| Geography | -0.213 | 0.924 | -0.197 | 0.896 | -0.205 | 0.936 |
| Hofstede6D | -0.147 | 0.810 | -0.178 | 0.857 | -0.097 | 0.746 |
| Language | 0.005 | 0.498 | -0.107 | 0.667 | 0.189 | 0.208 |
| **Modernization** | **0.347** | **0.022** | **0.274** | **0.048** | **0.245** | **0.047** |
| Personality | 0.133 | 0.204 | 0.170 | 0.191 | 0.062 | 0.295 |
| Politics | -0.125 | 0.820 | -0.122 | 0.818 | -0.077 | 0.730 |
| **Religion** | **0.561** | **0.014** | **0.601** | **0.011** | 0.299 | 0.109 |
| Subsistence | -0.079 | 0.698 | -0.029 | 0.563 | -0.004 | 0.491 |
| Adjusted R-squared | 0.561 | | 0.532 | | 0.267 | |

| Follow-up Analysis | Representational Geometry Models | | | | | |
|---|---|---|---|---|---|---|
| Predictors | Full-feature Model | | Dimensional Model | | Categorical Model | |
| | $\beta$ | $p$ | $\beta$ | $p$ | $\beta$ | $p$ |
| Climates | 0.198 | 0.156 | 0.188 | 0.181 | 0.182 | 0.120 |
| Demographics | 0.035 | 0.395 | 0.063 | 0.332 | -0.278 | 0.986 |
| Disease | -0.014 | 0.575 | -0.027 | 0.632 | -0.052 | 0.729 |
| Education | 0.123 | 0.203 | 0.122 | 0.207 | 0.032 | 0.366 |
| Gene | -0.139 | 0.700 | -0.23 | 0.839 | -0.065 | 0.598 |
| Geography | -0.122 | 0.776 | -0.127 | 0.790 | -0.164 | 0.891 |
| Hofstede6D | -0.087 | 0.700 | -0.125 | 0.773 | -0.09 | 0.728 |
| Language | 0.091 | 0.369 | -0.047 | 0.573 | 0.238 | 0.154 |
| Personality | 0.146 | 0.188 | 0.188 | 0.177 | 0.05 | 0.321 |
| Politics | -0.078 | 0.716 | -0.083 | 0.730 | -0.063 | 0.691 |
| **Religion** | 0.317 | 0.109 | **0.419** | **0.048** | 0.177 | 0.223 |
| Subsistence | -0.045 | 0.614 | -0.005 | 0.502 | 0.023 | 0.413 |
| **Urbanization** | **0.445** | **0.018** | **0.326** | **0.050** | **0.307** | **0.044** |
| Wealth | -0.072 | 0.693 | -0.077 | 0.708 | 0.08 | 0.241 |
| Object knowledge | 0.017 | 0.439 | 0.056 | 0.359 | -0.098 | 0.749 |
| Adjusted R-squared | 0.618 | | 0.559 | | 0.299 | |

Note: As modernization was found to be significant in the main analysis, we decomposed it into subcomponents (that is, urbanization, wealth, and education) in the follow-up analysis. Full statistics can be found in Supplementary Table 8. β = standardized regression coefficients.

nature portfolio

Corresponding author(s): YIN WANG

Last updated by author(s): Oct 14, 2024

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size ($n$) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. $F$, $t$, $r$) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's $d$, Pearson's $r$), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Qualtrics, Meadows, MTurk, CloudResearch, Credamo, and NaoDao platform. |
| Data analysis | All data analysis codes were uploaded to GitHub (https://github.com/BNU-Wang-MSN-Lab/FAVEE-HPP). The Python version used for the analysis is 3.9.19, and the R version is 4.3.3. Detailed information on the libraries and packages used can be found in the requirements.txt file available on GitHub. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data in this project were accessible on Github (https://github.com/BNU-Wang-MSN-Lab/FAVEE-HPP) and deposited in the Open Science Framework (https://osf.io/nfkmj) and can be interactively viewed and freely downloaded at a dedicated website (https://bnu-wang-msn-lab.github.io/FAVEE-HPP).

## Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

| | |
|---|---|
| Reporting on sex and gender | See detailed demographics in Supplementary Fig. 13 |
| Reporting on race, ethnicity, or other socially relevant groupings | See detailed demographics in Supplementary Fig. 13 |
| Population characteristics | See detailed demographics in Supplementary Fig. 13 |
| Recruitment | In all studies, online participants were recruited via MTurk, CloudResearch, Credamo, and NaoDao platform. Ofline participants in study2 took part in lab experiments, which were conducted via Meadows. |
| Ethics oversight | All studies in this report were approved by the Institutional Review Board of Beijing Normal University (IRB_A_0024_2021002) |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Study description | All studies are quantitative. By using natural language processing, online surveys, laboratory cognitive tasks, and computational modelling on diverse modern cultures across the world (total n = 20,427) and ancient cultures across more than spanning 3,000 years of history, we examined universality and cultural variability in the ways that people conceptualize relationships. |
| Research sample | Study 1 recruited 1,065 online US participants via MTurk and 60 offline US participants. Study 2 was preregistered (https://osf.io/swr2c) and recruited 17,686 online participants across 19 global regions spanning five continents and covering 10 different languages. In addition, 229 native Mosuo people were recruited from Yongning Township (Yunnan Province, China). Study 3 recruited 44 scholars specialized in ancient Chinese culture for expert evaluation of the NLP method. Moreover, to test the FAVEE-HPP model in non-dyadic relationships, we recruited 380 online US participants (via MTurk) and 242 online Chinese participants (via NaoDao platform). Participants across all studies were native speakers who grew up or lived for the longest period of their life in the targeted regions. We achieved broadly representative samples for age, gender, education, and ethnicity (see details in in Supplementary Fig. 13). |
| Sampling strategy | Power analysis was done to predetermine the sample size. To establish a design with adequate statistical power, a pilot study was conducted (n=721, recruited from MTurk) using the dimensional survey from Wish et al., (1976). We collected at least 80 participant responses for each relationship on each evaluative feature, and the results of Wish et al., (1976) were completely replicated(Supplementary Fig. 12a & 12b). We ran a Monte Carlo simulation test to derive the minimally required responses in each condition to maintain a stable and consistent PCA result. The PCA was performed on each subsample (from 2 to 40, with 1000 iterations for each subsample), and loading scores and relationship scores were compared with the overall dataset using Pearson's correlation. The simulations results Supplementary Fig. 12c) indicated that subsamples with 10 responses were almost identical with the entire dataset (rating correlation r > 0.95) and thus should be adequate to ensure highly similar derived PCA components (loading score correlation r > 0.90; relationship score correlation r > 0.95). |
| Data collection | Online participants were recruited via MTurk, CloudResearch, Credamo, and NaoDao platform. In addition, native Mosuo people were recruited from Yongning Township (Yunnan Province, China), using a field research data collection style (i.e., through face-to-face interviews and door-to-door paper surveys). We remained blinded during data analysis, as no specific experimental conditions were involved and participants were randomly sampled. |
| Timing | Pilot Study and Study1 data were collected in 2019-2022. Study2 and 3 were collected in 2023. |
| Data exclusions | We excluded data from dimensional survey based on attention-check questions and familiarity ratings. A total of 129 participants (pilot study), 248 participants (study1), and 2,441 participants (study2) were excluded due to their responses being outliers, suggesting they either did not respond carefully or were unfamiliar with relationships they rated. In addition, in the cognitive tasks of Study 1, there was no significant correlation between one participant and the left average results of the participants. Consequently, we viewed the participant as an outlier and remove this participant from the following analysis. |

| | |
|---|---|
| Non-participation | In pilot study, 129 participants (out of 721; 17.89%) were excluded. In study 1, 248 participants (out of 1,065; 23.29%). In study 2, 2,441 participants (out of 18,537; 13.17%) were excluded. |
| Randomization | Each participant was randomly assigned to a subset of relationships (e.g., 5-8 relationships) and had to rate them on a subset of evaluative features (e.g., 10-11 features). |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |
| ☒ | ☐ Plants |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Plants

| | |
|---|---|
| Seed stocks | *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.* |
| Novel plant genotypes | *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.* |
| Authentication | *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.* |