TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 00 (2026) e70061

ISSN: 1756-8765 online
DOI: 10.1111/tops.70061

This article is part of the topic "Diverse Minds: Incorporating Diversity into Cognitive Science Research," Annika Tjuka, Yoolim Kim, Yağmur Deniz Kısa, Daniel B. M. Haun, and Russell D. Gray (Topic Editors).

Abstract Cognition Across Cultures

Amritpal Singh, Jiangxue Ning, Shikun Su, Daniel Casasanto

Department of Psychology, Cornell University

Received 8 September 2025; received in revised form 25 March 2026; accepted 29 April 2026

Abstract

How do minds differ across cultures? According to one long-standing perspective in the social sciences, people from Chinese cultures think less abstractly than people from Western cultures. Here, we argue that previous studies of attention, perception, and memory suggest that East Asians may think *more* abstractly than Westerners due to their greater sensitivity to social and physical contexts. We tested these competing hypotheses in studies focused on two operationalizations of abstract thought: event construal and analogical reasoning. In the first three studies ($N = 1797$), participants chose whether events were best described abstractly or concretely. Across six independent national samples, Chinese participants were more likely than U.S. participants to construe events abstractly. In the next study ($n = 965$), participants from a more context-sensitive Chinese subculture were more likely to construe events abstractly than those from a less context-sensitive subculture. In the final two studies, participants attempted to solve analogical reasoning problems. Across four independent national samples ($N = 677$), Chinese participants were more likely than U.S. participants to solve the analogical reasoning problems correctly. Together, these results challenge the claim that Westerners

A.S. is now at the Department of Psychology and Rutgers Center for Cognitive Science, Rutgers University, New Brunswick.

Preliminary portions of this work were presented in A.S.'s doctoral dissertation, S.S.'s master's thesis, at the 3rd NDSEG Fellows Conference, the 59th and 61st Annual Meetings of the Psychonomic Society, the 41st and 43rd Annual Conferences of the Cognitive Science Society, and the 2nd Annual Culture and Cognition Preconference.

Correspondence should be sent to Amritpal Singh, Rutgers Center for Cognitive Science, 152 Frelinghuysen Rd, Piscataway, NJ, 08854, USA. E-mail: as3695@cornell.edu

think more abstractly than Chinese people and suggest that context-sensitivity may be an engine of abstract cognition.

Keywords: Abstract cognition; Context-sensitivity; Culture; Event cognition; Analogy

1. Introduction

People from Western and East Asian cultures differ in how they perceive and think about the world (Markus & Conner, 2014; Nisbett, 2003). Whereas Western people (typically sampled from Caucasian North American or Western European populations) think of the self as distinct from others (an independent self), East Asian people (typically sampled from Chinese, Japanese, or South Korean populations) view the self primarily as related to others (an interdependent self; Markus & Kitayama, 1991). Westerners are characterized as having an analytic style of perception and cognition, meaning they tend to focus on objects, categories, and the rules that govern them. By contrast, people from East Asian cultures are characterized as having a holistic style of perception and cognition, meaning they tend to focus on the relationships between objects and their environment (Nisbett, Peng, Choi, & Norenzayan, 2001). These social, cognitive, and perceptual differences reflect a broader difference: Western cultures promote greater focus on decontextualized people, objects, and events, whereas East Asian cultures promote greater sensitivity to the social and physical contexts of those people, objects, and events (Hofstede, 2001; Markus & Conner, 2014; Masuda, Gonzalez, Kwan, & Nisbett, 2008; Nisbett, 2003; Nisbett & Masuda, 2003; Nisbett et al., 2001; Wang, 2013).

How might more or less attention to context affect cognition? According to Nisbett and colleagues, non-Western people, and specifically Chinese people, may have difficulty reasoning abstractly because they attend more to context than Westerners. In a landmark paper, Nisbett et al. (2001) argued that Westerners, inheriting a tradition of formal logic from the Ancient Greeks, "should be more capable of…reasoning based on logical rules" (p. 296) than East Asians who, according to the authors, make "relatively little use of categories and formal logic" (p. 291). They claim that whereas Westerners tend toward a "search for abstractions," East Asian people tend toward "a concern with concrete objects and events" (p. 306).

Norenzayan, Smith, Kim, and Nisbett (2002, Studies 3–4) tested this hypothesis in a task where East Asian and Western participants attempted to reason about objects and events based on formal rules while ignoring the real-world contexts in which these objects and events are usually embedded. Western participants were more likely than East Asian participants to use abstract, rule-based strategies to complete the tasks successfully (Norenzayan et al., 2002, but see Klein et al., 2009 and Murphy, Bosch, & Kim, 2017). The authors concluded that the East Asian participants made more errors because they were less likely to "decontextualize the content of an argument from its logical structure" (Norenzayan et al., 2002, p. 678).

Chinese language and culture have been characterized as impediments to abstract thinking. According to one of the first Western sociologists to study Chinese society, Marcel Granet, "the Chinese have no taste for abstract symbols" (Granet, 1934, pp. 478–479 in Moser, 1996, p. 5). According to the philosopher and linguist David Moser, Chinese thinking was viewed by

Western scholars as “indifferent to abstraction” (Moser, 1996, p. 7), and the Chinese language was thought to be “an unsuitable vehicle for abstract ideas” (Moser, 1996, p. 6).

Within experimental psychology, claims about Westerners’ superiority in abstract thinking have focused on formal reasoning and categorization (Nisbett & Masuda, 2003; Nisbett et al., 2001; Norenzayan et al., 2002). However, rule-based reasoning is only one type of abstract thought. In the next section, we review several ways in which researchers have distinguished between abstract and concrete thought. For each notion of abstractness, we will argue that the empirical results suggest that East Asians, and specifically people from Chinese culture, may think more abstractly than Westerners due to their greater context-sensitivity.

2. Abstractness and concreteness

Abstractness[1] is defined in different ways by different researchers, but there are key overlapping aspects across these definitions. First, whereas concrete entities more often afford perceptuo-motor experience (e.g., cats, cars, or cups), abstract entities usually exist only in the mind (e.g., time, justice, or cogitation; Paivio, Yuille, & Madigan, 1968). Additionally, abstractness and concreteness can also be thought of as describing hierarchies where the superordinate (higher) levels of the hierarchy (e.g., a corporation) are more abstract, and the subordinate (lower) levels are more concrete (e.g., an employee of a corporation; Trope & Liberman, 2010), though superordinates at one level may be the subordinates of yet higher levels (Burgoon, Henderson, & Markman, 2013; Freitas, Gollwitzer, & Trope, 2004). Some scholars have compiled large lists of definitions and measures of abstract thought (e.g., Burgoon et al., 2013; Hale, 1988), but we will review only those notions of abstractness, which illuminate potential relationships between context, culture, and abstract thinking.[2] Below we describe three ways in which people can perceive, attend to, or construe their experiences more abstractly or concretely: (i) Attending to wholes versus parts; (ii) attending to relationships versus relata; (iii) construing actions in terms of their goals versus their implementations.

2.1. Abstract wholes and concrete parts

Colloquially, English speakers talk about abstract thinking as “seeing the forest for the trees.” This idiom refers to focusing on broader, more abstract wholes rather than the parts which make them up (Burgoon et al., 2013). Wholes correspond to superordinate levels of a hierarchy and parts to subordinate levels. Wholes can be the parts of yet greater wholes; each superordinate level provides a social, perceptual, or physical context for its subordinate levels (e.g., branches are parts of trees, trees are parts of forests, and forests can be parts of landscapes).

This distinction between wholes and parts aligns with a distinction in psychology between perceiving holistic, “global” gestalts as opposed to their analytic, “local” constituents (Burgoon et al., 2013; Navon, 1977). In Navon’s (1977) influential study of hierarchical perception, participants saw a series of large letters composed of small letters (e.g., a large

H made of small S's). Participants pressed a button when a target letter appeared as either a large, global letter or a small, local letter. Perceiving a global target requires ignoring the local parts and attending to the higher-level whole (i.e., the gestalt). U.S. participants perceived targets more quickly at the global level than the local level, a difference called "global precedence." Navon (1977) likened the global precedence effect to a tendency to see the "forest before the trees." In a cross-cultural study, East Asian participants, 84% of whom were from China, demonstrated greater global precedence than Caucasian Western participants (McKone et al., 2010), consistent with other results showing that East Asians attend more to contextual information in the perceptual field than Westerners do (Masuda & Nisbett, 2001; Nisbett & Masuda, 2003; Nisbett et al., 2001). In the case of wholes versus parts, Chinese people seem to be more inclined than Westerners to see the forest before the trees.

2.2. *Abstract relations and concrete relata*

Thinking about relations or relational categories is more abstract than thinking about relata (i.e., the things being related; Gentner & Asmuth, 2019; Hafri & Firestone, 2021; Hale, 1988).[3] Whereas relata are typically more available to perceptuo-motor experience, a relationship between those relata must be inferred and is less available to our senses (Gentner & Asmuth, 2019; Paivio et al., 1968). For example, we can perceive adults and children but we cannot literally see the kinship relationship between them. Hierarchically, superordinate relationships comprise their subordinate relata and provide social or physical contexts for their relata (e.g., solar systems provide physical contexts in which planets are embedded and related to one another). Thus, thinking about relationships means thinking about the contexts in which objects and ideas are embedded.

Chinese participants show more relational, context-sensitive patterns of attention, perception, and memory relative to participants from Western cultures. For example, in one cross-cultural study of attention (Ji, Peng, & Nisbett, 2000, Study 2), participants viewed a square frame with a rod sitting inside the frame. The rod and frame rotated independently of each other. Participants were asked to judge when the rod was vertical, regardless of the frame's position. Relative to U.S. Caucasian Americans, East Asians' (i.e., Chinese, Japanese, and Korean participants) verticality judgments were more influenced by the visual relationship between the rod and frame. In a cross-cultural study of autobiographical memory, East Asians and Asian Americans (the majority from Chinese cultures) tended to recall more memories based on general routines and more memories that dealt with interpersonal interactions. By contrast, U.S. Caucasian Americans remembered more memories about specific, detailed moments, and their memories typically focused on themselves as main characters (Wang & Ross, 2005). East Asian and Asian American participants' memories were more relational because they were thematically related by common actions and focused more on the interactions that relate people, rather than on individuals alone as main characters. These studies, alongside others (e.g., Chua, Boland, & Nisbett, 2005; Masuda & Nisbett, 2001; Masuda & Nisbett, 2006; Unsworth, Sears, & Pexman, 2005; Wang, 2013), suggest that people from Chinese cultures seem to perceive and remember the world in a more relational, and thus abstract, manner.

We have interpreted the above studies of attention, perception, and memory to suggest that people from Chinese cultures may think more abstractly than Westerners due to their greater sensitivity to context. However, the tasks described above were not designed to test how abstractly or concretely individuals think (i.e., they were not meant to test for differences in abstract cognition). In the following section, we turn to a third notion of abstractness and concreteness and a validated cognitive measure of that notion.

2.3. Abstract goals and concrete implementations

Actions are structured hierarchically and can be conceptualized at different levels of abstractness (Burgoon et al., 2013; Reilly et al., 2025; Trope & Liberman, 2010). One can construe actions in terms of either the concrete details of how the action is implemented or in terms of the action's abstract goals. The goals of actions are more abstract than their implementations along several interrelated dimensions of abstractness: in terms of sensorimotor availability (Paivio et al., 1968) because goals are typically not as available to direct experience as the implementations by which we arrive at those goals; in terms of spatiotemporal psychological distance (Trope & Liberman, 2010) because goals are typically more temporally and spatially distant than their implementations; and in terms of relationality and information specificity (Gentner & Asmuth, 2019; Gentner & Hoyos, 2017) because goals are less specific than their implementations and causally associate dissimilar actions and events. Further, goals correspond to the superordinate levels of action hierarchies that are composed of the subordinate, implementational levels (Burgoon et al., 2013; Trope & Liberman, 2010, Vallacher & Wegner, 1987). Finally, goals provide a context that connects actions and their implementational details through functional relationships.

Vallacher and Wegner (1989) validated a questionnaire called the Behavioral Identification Form (BIF) to measure at what level people construe events or actions: the higher, more abstract level of the goal or the lower, more concrete level of the goal's implementation. For 25 different common events or actions, participants choose whether that item (e.g., *voting*) is best described in terms of its abstract goals (*influencing the election*) or in terms of its concrete implementational details (*marking a ballot*). Two previous studies have suggested that Chinese participants construe events more abstractly than U.S. participants (Hong & Lee, 2010; Miyamoto, Knoepfler, Ishii, & Ji, 2013)

2.4. The present studies

Here, in Studies 1a–1c, we conduct further tests of differences in abstract versus concrete action construal between Caucasian American and Chinese participants. If Chinese people tend to think more concretely than Westerners do (Moser, 1996; Nisbett et al., 2001; Norenzayan et al., 2002), then Chinese participants should choose concrete descriptions more often than American participants. However, if people from Chinese culture tend to think more abstractly than Westerners, consistent with their greater context sensitivity and with the results of the attention, perception, and memory studies reviewed above, then Chinese participants should choose abstract descriptions more often than U.S. participants. In Study 2, we test for differences in concrete versus abstract event construal within China, comparing two subcultures shown previously to differ in context sensitivity. In Study 3, we test whether cross-cultural differences between Chinese and U.S. participants extend to a different measure of abstract thinking, analogical reasoning.

3. Studies 1a–1c: Event cognition across the United States and China

3.1. Study 1a

Two previous studies have administered versions of the BIF to Chinese and American participants (Hong & Lee, 2010, Experiment 5; Miyamoto et al., 2013, Study 3). However, one study used only a subset of the BIF items (Miyamoto et al., 2013), and neither study controlled for relevant demographic factors that could potentially explain away differences in abstract thinking between groups, such as age, education, and sex. Thus, we conducted Studies 1a–c for four reasons: (i) to use the entire validated measure of event construal (i.e., the BIF); (ii) to conduct multiple exact replications before conducting conceptual replications, following best practices (e.g., Pashler, 2018); (iii) to ensure we would have a sufficient number of participants in a combined dataset to conduct analyses that statistically controlled for the demographic factors of age, education, and sex; (iv) to ensure that we had enough Chinese participants for a subsequent study investigating abstract thought within China (see Study 2).

3.1.1. Method

3.1.1.1. Transparency and openness: Methods were preregistered at osf.io (https://doi.org/10.17605/OSF.IO/HWD85). All data, analysis code, and research materials are available on OSF (https://osf.io/m6rbv/?view_only=6c96471d67cd42eba20f467d22df6ea1). We chose a target sample size of 400 for the first study because we were not sure how the effect size might potentially be affected by online participation in uncontrolled and potentially distracting environments. We report all data exclusions performed and all measures in this study. Data were analyzed using the statistical package lme4 (version 1.1-31; Bates, Mächler, Bolker, & Walker, 2015) using R (version 4.2.2; R Core Team, 2022) in RStudio (version 2023.03.1+446; Posit Team, 2023) to run mixed effects linear models (Baayen, Davidson, & Bates, 2008).

3.1.1.2. Participants: Our target sample size was 400 participants before exclusions (200 Chinese participants, 200 U.S. participants). We collected data from a total of 418 participants (18 participants more than we had planned due to difficulty collecting exactly the intended number with the online platforms we used; Chinese: $n = 215$; U.S.: $n = 203$). We excluded participants who: participated twice ($n = 28$); did not complete the survey ($n = 4$); were not monolingual native U.S. Caucasian Americans ($n = 54$) or native Chinese individuals ($n = 17$); gave incomplete or nonsensical answers to demographic questions ($n = 1$); and one Chinese participant who took longer than 1 h to complete the task (the amount of time U.S. participants had to finish). After exclusions, 313 participants remained whose data were analyzed (Chinese: $n = 168$ [107 females, 61 males, mean age = 35.52 years old]; Caucasian American: $n = 145$ [62 females, 83 males, mean age = 36.10 years old]).

3.1.1.3. Materials: We used all 25 items of the Behavior Identification Form (BIF; Vallacher & Wegner, 1989) and created a Mandarin translation for this study. We produced the Mandarin version using a translation and back-translation process with two native Mandarin speakers. One native speaker translated the English BIF into Mandarin, and the other translated the Mandarin version back into English without having seen the original text. Any discrepancies between the back-translation and the original English instructions were discussed by the two Mandarin speakers and A.S. to finalize the Mandarin version of the questionnaire.

We created four versions of the BIF, each with a different fixed randomized order of the items. In each version, the abstract description appeared as the top option for 13 of the items, and as the bottom option for the remaining 12 items.

3.1.1.4. Procedure: We administered the BIF to participants over online survey platforms in the United States (Amazon Mechanical Turk) and in China (wjx.cn, a Chinese survey website). For each item, participants chose either the concrete or the abstract description of the action, following these instructions: "Your task is to choose the description that seems best to you. Choose one description for each pair. There are no right or wrong answers; people simply differ in their preferences for the different descriptions, and we are interested in your personal preferences."

Participants then answered 11 demographics questions: sex, first language, height, age, political orientation, birth country, race, occupation, income, education, and handedness. First language, birth country, and race were used for exclusions; age, sex, and education were collected for analyses controlling for those demographic factors. The remaining measures were collected for an unrelated study.

3.1.2. Results and discussion

We analyzed choice of construal (abstract vs. concrete) as a function of culture (Chinese or Caucasian American) in a binomial logistic regression with random intercepts for item and subject and a random slope for culture over the items intercept. We tested for the statistical significance of a factor of interest by comparing a full model including the factor of interest (e.g., culture) to a reduced model which included all variables except for the factor of interest. If a model failed to converge, we simplified the random effects structure of the model to try to facilitate convergence (Barr, Levy, Scheepers, & Tily, 2013).

Chinese participants were more likely to select the abstract description than U.S. participants, Wald $\chi^2(1) = 6.08$, $p = .01$, OR = 1.96, 95% CI [1.16, 3.33] (see Figs. 1a and 2). These results suggest that Chinese culture fosters more abstract thinking relative to U.S. culture. These results are also consistent with the findings of two previous studies that administered the BIF to Chinese and American participants (Hong & Lee, 2010, experiment 5; Miyamoto et al., 2013, study 3).

3.2. Study 1b

Study 1b attempts to exactly replicate the result of Study 1a with greater power.

3.2.1. Method

3.2.1.1. Transparency and openness: Methods were preregistered at osf.io (https://doi.org/10.17605/OSF.IO/72E39). We report how we determined the sample size for Study 1b, all data exclusions performed, and all measures in this study. The materials, procedure, exclusion criteria, analysis software, and analyses were identical to Study 1a.

3.2.1.2. Participants: To determine sample size, we conducted a power analysis for generalized linear models using the pwr package (version 1.2-2; Champely, 2020), extracting R2 using the piecewiseSEM package (version 2.1.0; Lefcheck, 2016). According to this analysis, Study 1a had 44.92% power, and a sample of 664 participants would be needed to reach 80% power in Study 1b. Based on attrition in the first study, we targeted a total of 800 participants (400 from each culture). We recruited 830 participants (Chinese $n = 422$; U.S. $n = 408$). We excluded participants who: participated twice ($n = 2$); did not complete the survey ($n = 39$); were not monolingual native U.S. Caucasian Americans ($n = 94$) or native Chinese individuals ($n = 12$); gave incomplete or nonsensical answers to demographic questions ($n = 2$); participated in Study 1a ($n = 3$); and two Chinese participants who took longer than 1 h (the amount of time U.S. participants had to finish). After exclusions, Study 1b $N = 676$ (Chinese: $n = 372$ [178 females, 194 males, mean age = 33.32 years old]; U.S. American: $n = 304$ [158 females, 146 males, mean age = 36.77 years old]).

3.2.2. Results and discussion

Study 1b successfully replicated the result of Study 1a: Chinese participants were more likely to construe events abstractly than U.S. participants, Wald $\chi^2(1) = 15.93$, $p < .001$, OR = 2.60, 95% CI [1.69, 4.00] (see Figs. 1b and 2).

3.3. Study 1c

Study 1c attempts to exactly replicate Studies 1a and 1b with greater power.

3.3.1. Method

3.3.1.1. Transparency and openness: Methods were preregistered at osf.io (https://doi.org/10.17605/OSF.IO/AEU8W). We report how we determined the sample size for Study 1b and all data exclusions performed. The materials, procedure, exclusion principles, analysis software, and analyses were identical to those of Studies 1a–b except that the sample size was increased. We also collected data on participants' self-construal tendencies for a separate set of analyses, which we do not report here.

3.3.1.2 Participants: The target sample size for this study was based on a power analysis that showed that Study 1b had 63.33% power and that a sample of 822 participants would reach 80% power. We wanted to reach at least 80% power to detect an effect of the same size as Study 1b. Based on attrition in the first two studies, we aimed to recruit a total of 1000 participants (500 Chinese participants, 500 U.S. participants) to increase the chances that we reach our target sample size after exclusions. We collected data from a total of 1049 participants (Chinese $n = 537$, U.S. $n = 512$). We excluded participants who: participated twice ($n = 2$); did not complete the survey ($n = 12$); were not monolingual native U.S. Caucasian Americans ($n = 139$) or monolingual native Chinese individuals ($n = 37$); gave nonsensical answers to self-construal questions ($n = 35$); participated in Studies 1a or 1b ($n = 13$); and three Chinese participants who took longer than 1 h (the amount of time U.S. participants had to finish). After exclusions, Study 1c $N = 808$ (Chinese: $n = 456$ [255 females, 201 males, mean age = 31.11 years old]; Caucasian American: $n = 352$ [190 females, 162 males, mean age = 37.48 years old]).

3.3.2. Results and discussion

Study 1c successfully replicated Studies 1a and 1b: Again, Chinese participants construed events abstractly more often than U.S. participants did, Wald $\chi^2(1) = 5.83$, $p = .02$, OR = 1.68, 95% CI [1.12, 2.53] (see Figs. 1c and 2).

3.4. Combined results

Across Studies 1a–1c, Chinese participants chose abstract descriptions of events more often than U.S. participants. Next, we sought to examine whether relevant demographic variables

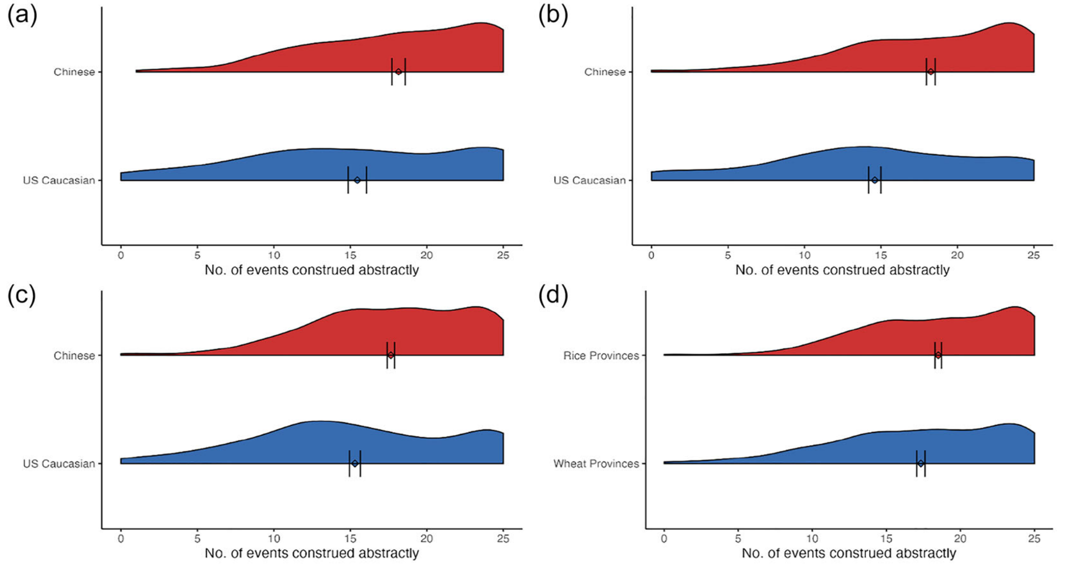

Fig. 1. Distributions of Participants' Behavioral Identification Form (BIF) scores.
Note. Panels 1a–1c: Studies 1a–1c. Panel 1d: Study 2. Diamonds are means. Error bars represent the standard error of the mean (SEM).

(age, education, sex) could explain away the difference in BIF performance in the combined data from Studies 1a–1c ($N = 1797$: Chinese $n = 996$ [540 females, 456 males, mean age = 32.17 years old]; U.S. Caucasians $n = 801$ [410 females, 391 males, mean age = 36.96 years old]). The data from these studies were combined after establishing that there were no statistically significant differences in event construal between the three U.S. samples, Wald $\chi^2(2) = 3.44$, $p = .179$, nor the three Chinese samples, Wald $\chi^2(2) = 4.74$, $p = .093$. In the combined data, Chinese participants construed events abstractly more often than U.S. participants, Wald $\chi^2(1) = 11.57$, $p < .001$, OR = 2.02, 95% CI = [1.40, 2.93], and this difference remained significant when controlling separately for education, Wald $\chi^2(1) = 44.90$, $p < .001$, sex, Wald $\chi^2(1) = 11.75$, $p < .001$, and age, Wald $\chi^2(1) = 12.72$, $p < .001$. Models attempting to control for all these variables simultaneously failed to converge even after model simplification. We did not find a significant main effect of education, Wald $\chi^2(1) = 4.76$, $p = .45$. We found significant main effects of age, Wald $\chi^2(1) = 5.93$, $p = .02$, and sex, Wald $\chi^2(1) = 6.27$, $p = .01$, with older adults and males construing events abstractly more often in the combined dataset. However, the main effects of these demographic variables do not bear directly on our hypotheses regarding Chinese and American culture. Altogether, these results firmly establish a cultural difference in event construal via multiple exact replications and by showing that the difference between cultures remains significant even when statistically controlling for relevant demographic variables.

However, could the differences in event construal have been due to participants' use of difference languages to complete the BIF rather than culture-linked cognitive differences in context-sensitivity? American participants completed the BIF in English, and Chinese participants completed it in Mandarin. We used a translation-backtranslation process to minimize differences between the English and Mandarin versions of the BIF, but it is possible that the online use of different languages to complete the questionnaire may have led to the differences found between American and Chinese groups. One aim of Study 2 was to address this concern by testing whether two (sub)cultures that differ in relational thought, and thus in context-sensitivity, would differ in event construal even when all participants completed the task in one language (Mandarin).

4. Study 2: Event cognition within China

Study 2 had two aims. The first aim was to triangulate on the relationship between context-sensitivity and differences in abstract cognition. If differences in context-sensitivity drive cultural differences in abstract and concrete thinking, then (sub)cultural groups that differ in context-sensitivity should differ in how abstractly or concretely they tend to think, even when they are found within the same country. But how do we identify relevant subcultures? Subsistence style (historical or present) influences cultural mindsets, even beyond the time during which the culture practices that subsistence style (Berry, 1967; Uskul, Kitayama, & Nisbett, 2008). Specifically, rice-paddy agricultures are thought to foster more interdependent and relational (and thus more context-sensitive) patterns of thinking than agricultures based in wheat farming due to the greater large-group coordination and long-term planning needed for rice paddy farming, compared to wheat farming (Dong, Talhelm, & Ren, 2019; Talhelm et al., 2014). Therefore, Study 2 tested for a difference in abstract versus concrete thinking between participants who lived in (historically) rice or wheat provinces from Study 1's Chinese sample. Based on the results of Studies 1a–1c and previous characterizations of rice and wheat provinces (Talhelm et al., 2014), we predicted that participants from rice provinces would construe events abstractly more often than participants from wheat provinces.

Study 2's second aim was to address the question of whether cross-cultural differences in event construal are due to cognitive differences (i.e., differences in sensitivity to context) or due to the online use of different languages to complete the BIF (English for U.S. participants and Mandarin for Chinese participants). Study 2 addresses this question by testing groups of people that all completed the BIF in Mandarin.

4.1. Method

4.1.1. Transparency and openness

Sample size was determined by combining the Chinese data from Studies 1a–1b and excluding participants for whom we could not determine the province they were from, or who were not from a rice-based or wheat-based province. We report all data exclusions performed and all measures in the study. This study's design and analysis were not preregistered.

From our total sample of Study 1 Chinese participants ($n = 996$), we excluded participants for whom we could not determine the province where they lived ($n = 20$) or who were not from a historically rice-based or wheat-based province ($n = 11$). After exclusions, Study 2 $n = 965$ (rice provinces $n = 538$ [304 females, 234 males, mean age = 31.23 years old]; wheat provinces $n = 427$ [220 females, 207 males, mean age = 33.17 years old]). We based our categorization of participants' provinces on participant information provided by wjx.cn and on previous research comparing rice and wheat provinces (Talhelm et al., 2014, Fig. 2). We categorized the following provinces as "wheat provinces" in our analysis: Anhui, Beijing, Gansu, Guizhou, Hebei, Heilongjiang, Henan, Jilin, Liaoning, Ningxia, Qinghai, Shaanxi, Shanxi, Shandong, Tianjin, and Yunnan. We categorized the following provinces as "rice provinces" in our analysis: Chongqing, Fujian, Guangdong, Guangxi, Hainan, Hubei, Hunan, Jiangsu, Jiangxi, Shanghai, Sichuan, and Zhejiang.

4.2. Results and discussion

Within China, participants from historically rice-growing provinces construed events abstractly more often than those from historically wheat-growing provinces, according to a binomial logistic regression with subsistence style as a fixed effect and random intercepts for subject, item, and province, with a random slope of subsistence style over the intercept for items, Wald $\chi^2(1) = 5.21$, $p = .02$, OR = 1.34, 95% CI [1.05, 1.71] (see Figs. 1d and 2). The effect of subsistence style remained significant when controlling separately for the participants' education, Wald $\chi^2(1) = 5.40$, $p = .02$, sex, Wald $\chi^2(1) = 5.46$, $p = .019$, and age, Wald $\chi^2(1) = 5.87$, $p = .02$. We did not find significant main effects of education, Wald $\chi^2(1) = 4.28$, $p = .51$, sex, Wald $\chi^2(1) = 3.41$, $p = .06$, or age, Wald $\chi^2(1) = 2.23$, $p = .14$. Partici-

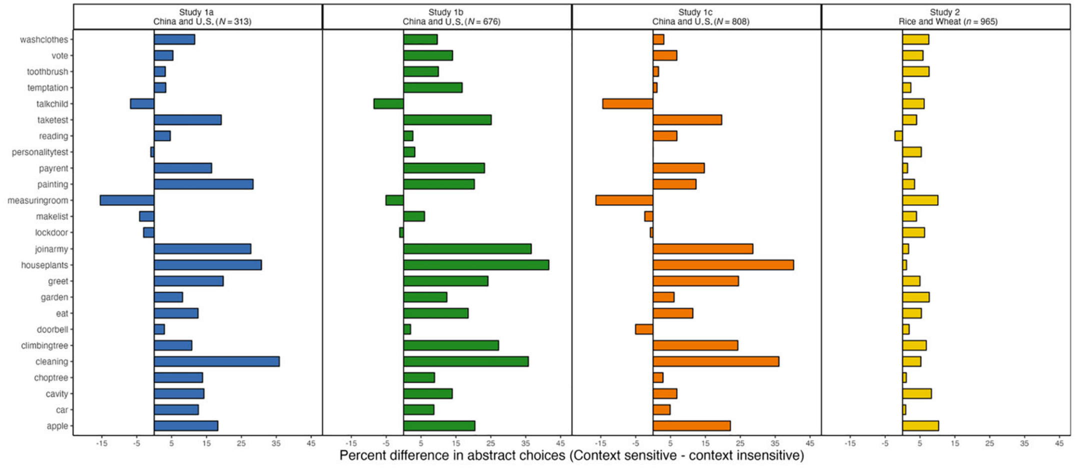

Fig. 2. Construal differences for each BIF item in Studies 1a–1c and 2.
Note. The x-axis represents whether an item was construed abstractly more often by the more or less context-sensitive group. The vertical black lines represent 0% difference. Bars extending to the right indicate items construed abstractly more often by the more context-sensitive group. Across studies, most items were construed abstractly more often by the more context-sensitive group.

pants from both historically rice-growing, Wald $\chi^2(1) = 14.92$, p $< .001$, and wheat-growing provinces, $\chi^2(1) = 6.63$, $p = .01$, were statistically more likely to construe events abstractly than Study 1 U.S. participants.

This (sub-)cultural difference between participants who all completed our Mandarin translation of the BIF addresses concerns that the differences in abstract thought in Study 1 were due only to the language participants used during the task. Further, the difference between rice and wheat provinces parallels the difference between Chinese and American participants: In both comparisons, participants from the more context-sensitive group were more likely to construe events abstractly, compared to those from the less context-sensitive group.

5. Studies 3a-3b: Analogical reasoning across the United States and Chinese cultures

Do the differences in Studies 1–2 reflect variation in abstract cognition more broadly, across multiple domains of abstractness, or are they limited to event cognition, specifically? To investigate this question, Study 3 tested for cultural differences using another paradigm case of abstract thinking: Analogical reasoning (Gentner, 1983; Gentner & Asmuth, 2019). Analogical reasoning is abstract because analogy requires one to discover invisible, intangible relationships between objects or ideas that may appear dissimilar at first. Discovering analogical relationships requires thinking about objects or ideas relationally, in the context of other objects or ideas.

On the basis of differences in context sensitivity across cultures, we hypothesized that Chinese people would be more likely to successfully engage in analogical reasoning relative to U.S. Caucasian Americans. To test this hypothesis, we administered a validated measure of analogical reasoning, Raven's (2003) Standard Progressive Matrices (SPM),[4] to American and Chinese participants. With stimuli composed of line drawings, the SPM is entirely pictorial and, thus, avoids concerns about the online use of different languages.

Raven's (2003) SPM consists of five sets of "puzzles." For each item, participants view visual patterns distributed across a matrix, with one portion of the matrix missing. Participants must choose which of the eight possible responses below the matrix correctly completes the incomplete matrix. Whereas easier puzzles in Sets A–D can be solved using non-analogical strategies, the most difficult puzzles in Set E typically require discerning analogical relationships (Raven & Court, 1996). Thus, we administered only Set E because our goal was specifically to test analogical reasoning.

5.1. Study 3a

5.1.1. Method

Methods were preregistered at aspredicted.com (https://aspredicted.org/YON_KUT). We targeted a sample size of 400 for the first study because we did not have a priori predictions about what effect size to expect for this novel comparison. We report all data exclusions and all measures in this study. Data were analyzed using the same software as in Studies 1–2.

Our target sample size was 400 participants (200 Chinese participants, 200 U.S. participants). 834 participants accessed the study through the study link (Chinese: $n = 432$; U.S.: $n = 402$), using wjx.cn (in China) and Amazon Mechanical Turk (in the United States). Following our pre-registered criteria, we excluded participants who: were not monolingual speakers of Mandarin ($n = 15$) or English ($n = 7$); reported that their country of birth was not China or the U.S. ($n = 1$); were not Caucasians ($n = 61$) in the U.S. sample and not Asian in the Chinese sample ($n = 3$); took less than 3 s (consistent with automated "bot" responses) or more than 3 min to respond on one or more SPM items (China: $n = 130$; U.S.: $n = 73$); answered an attention check incorrectly ($n = 49$); reported having done the SPM previously (China: $n = 31$; U.S.: $n = 50$); gave incomplete or nonsensical answers to demographic questions ($n = 9$); or did not meet our a priori participation criteria for Amazon Mechanical Turk workers ($n = 5$). After exclusions, $N = 400$ (Chinese: $n = 224$ [115 females, 109 males, mean age = 30.25 years old]; U.S. participants: $n = 176$ [77 females, 99 males, mean age = 39.44 years old]).

We used the 12 SPM Set E items. The language used for the consent process, instructions, and demographic questions was English for U.S. participants and Mandarin for Chinese participants. The English instructions were adapted from the standard instructions in the Manual for Raven's SPM (Raven & Court, 1996). We updated the wording of the questions to reflect contemporary U.S. English and a computerized context. We constructed a Mandarin version of the written materials following the same translation procedure described in Study 1a.

Participants completed the study online, using Qualtrics software. Participants read detailed instructions and then answered an example question from an easier SPM set (Set A, question

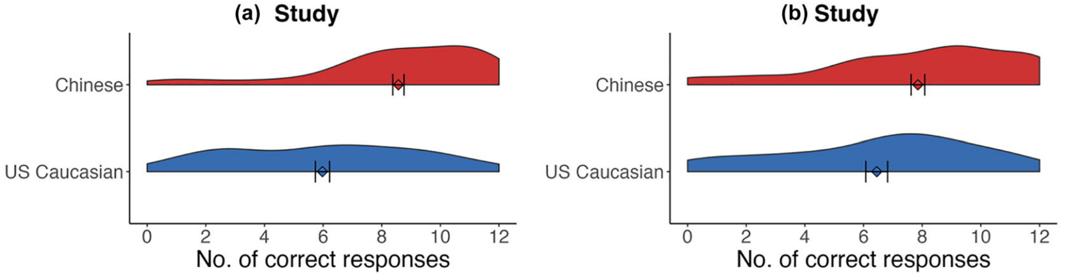

Fig. 3. Distributions of participants' Standard Progressive Matrices (SPM) scores.
Note. Diamonds show the mean number of correct responses on the SPM. Error bars represent the SEM.

A1). Participants were given the correct answer to the example question in the instructions; their job was to select this answer from the options provided. We used the example question as both a tutorial and a check: Participants who answered the example question incorrectly were excluded from analyses.

Next, participants completed SPM Set E. For each item, participants were instructed to select the option that completed each matrix and to respond as quickly and accurately as possible by clicking a radio button next to the response. Reaction times (RTs) were recorded for each response. Finally, participants answered demographic questions.

5.1.2. Results and discussion

We analyzed accuracy (correct or incorrect) as a function of culture (Chinese or Caucasian American group) in a binomial logistic regression with random intercepts for item and subject, and a random slope for culture over the intercept for item. We conducted similar model comparisons and, when necessary, model simplifications as in Studies 1–2. We ran further analyses to test whether the effect of culture would remain significant when including terms for potentially relevant demographic variables.

Chinese participants were more likely than U.S. participants to correctly solve the SPM matrix problems, $\chi^2(1) = 22.58$, $p < .001$, OR = 4.91, 95% CI [3, 8.04] (see Fig. 3a), indicative of more successful analogical reasoning. There were no significant main effects of age, $\chi^2(1) = 2.57$, $p = .11$, education, $\chi^2(2) = 0.59$, $p = .74$, income, $\chi^2(4) = 1.67$, $p = .80$, or sex $\chi^2(1) = 0.04$, $p = .84$, and the effect of culture remained significant even when controlling the fixed effects of education, $\chi^2(1) = 22.54$, $p < .001$, income, $\chi^2(1) = 63.33$, $p < .001$, age, $\chi^2(1) = 19.83$, $p < .001$, and sex, $\chi^2(1) = 22.61$, $p < .001$.

Could the difference in analogical reasoning reflect greater diligence in Chinese participants, who may have used more time than American participants to respond correctly? To investigate this alternative explanation for the cross-cultural difference in SPM performance, we tested for a speed-accuracy tradeoff. We determined that the residuals for RT data were not normally distributed and that it would be appropriate to transform the RTs when testing if RTs differed across cultures. We determined that the appropriate lambda value for a Box-Cox transformation was -0.106, using the CAR package (Fox & Weisberg, 2019). The residuals from a model using the transformed RTs appeared to be normally distributed. We found that transformed RTs did not differ significantly across cultures, $\chi^2(1) = 2.73$, $p = .09$. Additionally, the effect of culture on SPM accuracy remained significant even when including a random intercept of culture by untransformed[5] RT, $\chi^2(1) = 279.01, p < .001$, suggesting that the cultural difference was not due to a speed-accuracy tradeoff.

Therefore, Study 3a provides evidence that Chinese individuals may have a relative advantage in thinking abstractly relative to U.S. Caucasian Americans as indexed by analogical reasoning.

5.2. Study 3b

To the best of our knowledge, a cross-cultural difference in analogical reasoning had not been shown previously between U.S. and Chinese adult populations. Therefore, we conducted a direct replication of Study 3a.[6]

5.2.1. Method

3.8.1.1 Transparency and openness: Methods were preregistered at aspredicted.org (https://aspredicted.org/CMT_VBN). The materials, procedure, exclusion criteria, analysis software, and analyses were identical to Study 3a. Based on attrition in the previous study and one of the original planned analyses for this study, we targeted a sample size of 1000 participants before exclusions.

1014 participants accessed the study through the study link (Chinese: $n = 509$; U.S.: $n = 505$), using wjx.cn and Amazon Mechanical Turk. Following our pre-registered criteria, we excluded participants who: did not meet our a priori participation criteria on Amazon Mechanical Turk ($n = 11$); answered an attention check incorrectly ($n = 88$); responded inappropriately to digit span catch trials, consistent with either cheating or automated 'bot' responses ($n = 168$); reported that their country of birth was not China ($n = 1$) or the U.S. ($n = 13$); were not monolingual speakers of Mandarin ($n = 14$) or English ($n = 4$); reported having done the SPM previously (China: $n = 211$; U.S.: $n = 99$); were not Caucasians in the U.S. sample ($n = 37$); gave incomplete or nonsensical answers for the digit span task or demographic questions ($n = 6$); and took less than 3 s or more than 3 min to respond on one or more SPM items (China: $n = 66$; U.S.: $n = 19$). After these exclusions, 277 participants remained (Chinese: $n = 197$ [80 females, 117 males, mean age = 33 years old]; U.S. Caucasian Americans: $n = 80$ [32.99 females, 43 males, mean age = 38.86 years old]).

5.3. Results and discussion

Again, Chinese participants outperformed U.S. participants in solving the SPM problems $\chi^2(1) = 8.99, p = .0027$, OR = 2.59, 95% CI [1.41, 4.76] (Fig. 3b). We found significant main effects of age, $\chi^2(1) = 10.20$, $p = .001$, and income, $\chi^2(4) = 9.60$, $p = .047$, with younger and higher-income participants being more likely to accurately solve SPM problems. We did not find significant main effects of education, $\chi^2(2) = 1.08$, $p = .58$, or sex, $\chi^2(1) = .0007$, $p = .98$. Of primary interest, the effect of culture on SPM performance remained significant when controlling the fixed effects of age, $\chi^2(1) = 4.89$, $p = .03$, education, $\chi^2(1) = 8.02$, $p = .0046$, income, $\chi^2(1) = 7.13$, $p = .0075$, and sex, $\chi^2(1) = 8.95$, $p = .0027$.

We then tested for a speed-accuracy tradeoff using the same methods as in Study 3a. We determined it was appropriate to transform the RT data when testing whether RTs differed across cultures. We used a Box-Cox transformation with the suggested lambda (−0.071). RTs did not differ significantly across cultures, $\chi^2(1) = 2.28$, $p = .13$. A model testing for the effect of culture with a random intercept of culture by untransformed RTs produced singular fits even after model simplification. We then ran a model that controlled for RTs by including the interaction of untransformed RTs by culture as a fixed effect while testing for the independent effect of culture. The effect of culture on SPM accuracy remained significant above and beyond the interaction of culture by untransformed RTs, $\chi^2(1) = 4.81$, $p = .03$, suggesting there was no speed-accuracy tradeoff. Study 3b successfully replicates Study 3a: Chinese participants outperformed their American counterparts in analogical reasoning[7]. Together, Studies 1 and 3 suggest that the differences we observed are cultural differences in abstract cognition generally.

6. General discussion

Across 10 national samples ($N = 2474$), we provide evidence that Chinese people may have a stronger propensity for thinking abstractly than U.S. Americans. Across Studies 1a–1c, Chinese participants were consistently more likely than U.S. participants to construe events abstractly in terms of their high-level goals vs. their concrete, mechanical details. In Study 2, within China, members of a more interdependent and relational cultural group (rice-province residents) construed actions more abstractly than a relatively independent and less relational group (wheat-province residents), paralleling the difference between Chinese and American participants. Across Studies 3a–3b, Chinese participants outperformed U.S. Americans in a standard measure of abstract, analogical reasoning. These studies demonstrate the generalizability of the cultural differences in abstract thought by providing convergent evidence across two types of abstract cognition. These results support the hypothesis that Chinese people may have an advantage over their Western counterparts in abstract thought, consistent with Chinese people's greater sensitivity to physical and social contexts (e.g., Markus & Kitayama, 1991; Masuda et al., 2008).

The results of Study 1 are consistent with two previous experiments that reported differences in event cognition between American and Chinese samples (Hong & Lee, 2010, experiment 5; Miyamoto et al., 2013, study 3). These earlier studies did not control for demographic variables that could potentially explain cross-cultural differences in event construal. Here, we collected a large enough sample to conduct control analyses, which demonstrated that differences in event construal cannot be explained away by demographic variables that can influence abstract thought (i.e., age, education, and sex). Thus, Studies 1a–1c establish a consistent cultural difference in abstract thought and eliminate several alternative explanatory variables left open in the previous studies. Study 2 addressed an open question concerning all cross-cultural BIF studies (including ours): Are the differences in event construal due to cognitive differences or due to participants from the two groups using different languages? Study 2 addressed this question by finding a difference in event construal between Chinese participants from groups that differ cognitively but who all completed the BIF in Mandarin. Further, Study 2's results demonstrate that context-sensitivity differences can predict cognitive variation both across *and* within cultures. Thus, Study 2's results suggest that differences in context-sensitivity may be a mechanism by which cultural differences in abstract thought arise. Finally, we extend the differences in event cognition to a different paradigm case of abstract thinking, showing parallel differences across cultures in analogical reasoning.

To the best of our knowledge, these data provide the first direct comparison of analogical reasoning between East Asian and Western adults (see Footnote 5), but multiple studies have compared East Asian and Western children on versions of Raven's Progressive Matrices (Huang & Hanley, 1995; Lynn, 1991; Lynn, Chan, & Eysenck, 1991; Lynn, Pagliari, & Chan, 1988; Ma & Leung, 1991; Raven, 2000). Typically, East Asian children score higher than Western children, consistent with the present findings. In principle, cross-cultural differences in SPM scores measured in children could be due to differences in parental or educational practices (Chan, 1994; Chen & Chen, 1988) and could disappear over development. The present data do not rule out the possibility that parental or educational practices could contribute to the observed cross-cultural difference, but they do show that this difference persists in adulthood.

Although the SPM puzzles are pictorial, it is possible that the visual features of the SPM stimuli may be more familiar to Western participants than to Chinese participants because the SPM was designed by a Western scientist. Could differential familiarity with the stimulus features account for the cultural difference we observed in analogical reasoning? No, on the contrary, greater familiarity with the figures on the part of Westerners should benefit American participants, rather than Chinese participants. Thus, potential familiarity differences with stimulus features would predict the opposite pattern of results from what we observed.

In principle, a difference in analogical reasoning could be explained by different amounts of experience with the SPM, per se, in one culture or the other. To address this potential issue, we excluded participants from both cultures who reported having done the task before. The number of participants excluded for having done the SPM previously was greater in the U.S. sample for Study 1 and in the Chinese sample for Study 2; thus, these data do not provide any clear indication that one culture had more exposure to the SPM than the other. Together, these facts militate against an explanation of our findings that appeals to different amounts of familiarity or experience with the SPM across cultures.

Does this work inform (re)interpretations of culture-specific preferences for taxonomic and thematic relations? For example, in the triad task, participants view three words or pictures in each trial (e.g., a monkey, a panda, a banana) and are asked which two of the items go together. Two items share a taxonomic relationship (e.g., the taxonomic category *animal* relates monkeys to pandas), and two items share a functional/thematic relationship (e.g., the thematic category of *eating* relates monkeys to bananas). People from Chinese cultures (and historically rice-growing subcultures in China) tend to choose the items that share thematic relationships more than Westerners (and historically wheat-growing subcultures in China; Chiu, 1972; Ji, Zhang, & Nisbett, 2004; Talhelm et al., 2014). However, it is not clear whether taxonomic or thematic pairings are more abstract since both response types require participants to generate a more general relation that connects two relata at a more superordinate level of a hierarchy (see section *Abstract relations vs. relata* above).

Our experimental hypothesis was that environments which foster greater sensitivity to context may foster more abstract thinking. Thus, we recruited participants from groups that prior work had shown to differ in their sensitivity to information in physical and social contexts, namely, U.S. Americans and Chinese people. If we are correct about our hypothesized mechanism, then future work should observe variation in abstract cognition in other cultural groups that differ in context-sensitivity, beyond those studied here. Differences in abstract thought are likely influenced by many different factors across different types of abstractness. However, future work could reveal the degree to which diversity in abstract cognition is driven by the mechanism we highlight here: differential sensitivity to context. Such work is also motivated by the goal to expand the populations studied in cultural psychology beyond the North American, European, and East Asian samples that currently dominate much of the field (Krys et al., 2025).

We report convergent evidence that cultures may differ in abstract thought, to the extent that different notions of abstractness are successfully operationalized by event construal and analogical reasoning (e.g., information specificity, discerning relations, goal-directedness, sensorimotor availability). However, there are a variety of somewhat-overlapping, somewhat-distinct notions of abstract thought. Thus, we believe that our findings motivate future work to test for potential cultural differences when operationalizing other notions of abstractness (see Burgoon et al., 2013, for a review). Such work would further test the generality of the differences reported here. Additionally, by investigating which types of abstract thought differ (or do not differ) across cultures, such work would shed further light on what mechanisms may drive or constrain abstract cognition across cultures.

The finding that Chinese participants think more abstractly than their American counterparts, as indexed by validated measures of abstract versus concrete event construal (Trope & Liberman, 2010; Vallacher & Wegner, 1989) and analogical reasoning (Gentner & Asmuth, 2019; Raven, 2003), challenges a century-long tradition in the social sciences that suggests the opposite pattern should be found (e.g., Moser, 1996; Nisbett et al., 2001; Norenzayan et al., 2002; Wallace, 1871). The results of Studies 1–3 are also inconsistent with Norenzayan et al. (2002), who reported that Western participants performed better than East Asian participants in abstract, formal reasoning tasks. Furthermore, Norenzayan et al.'s (2002) findings failed to replicate in multiple independent attempted replications (Klein et al., 2009; Murphy et al., 2017). Thus, to our knowledge, no experimental evidence supports the notion that East Asian, and specifically Chinese, people differ from Western people in *formal* reasoning.

The arguments and data presented here are also consistent with recent work on social metacognition and abstract thought (Borghi, Mazzuca, & Tummolini, 2025). Social metacognition refers to the process of monitoring the adequacy of one's own conceptual knowledge and then interacting with others (or oneself through simulated dialogue) to supplement and align that knowledge with the knowledge of others (Borghi et al., 2025). Borghi et al. (2025) argue that social metacognition plays a central role in the acquisition and use of abstract concepts because such concepts are difficult to acquire from concrete experience without input from others. Such social metacognition is argued to be particularly necessary for abstract concepts that are imprecise or vague, because their meaning remains open to social negotiation. If it is true that social metacognition is necessary for the acquisition of abstract concepts, then cultures (such as those in China) that foster greater attention to social relations (i.e., to the social context) may facilitate the acquisition and *usage* of abstract concepts by promoting more social metacognition, relative to cultures such as those in the United States. Thus, future research is motivated to investigate the relationship between a culture's sensitivity to social context and social metacognition, and its propensity toward abstract thought.

Together, the results reported here show that people from cultures that foster more interdependent, holistic, and context-sensitive thinking are more likely to construe events abstractly and to engage in successful analogical reasoning compared to people who come from cultures that foster more independent, analytic, and less context-sensitive thinking. These results challenge the belief that East Asian people, and specifically Chinese people, are "more concrete in their conceptualizations" than Westerners (Moser, 1996, p. 134; Nisbett et al., 2001, p. 306). Rather than being an impediment to thinking abstractly, sensitivity to context may be an engine of abstract thought. To the extent that abstract thinking is characterized by perceiving global patterns (rather than their local constituents), attending to relationships (rather than the objects being related), construing actions in terms of the high-level goals and outcomes (rather than their mechanical details), and reasoning analogically, it appears that Chinese people are more likely than Westerners to think abstractly.

Notes

1 We distinguish the *process of abstraction* (i.e., of extracting general tendencies from experience, language, or concepts) from the categories of *concrete vs. abstract thought* (e.g., Reilly et al., 2025). This paper focuses on several notions of abstract thought (described in this section) but does not focus on the process of abstraction.
2 All human beings, and other animals, engage in *the process of abstraction* (e.g., in the formation of generalizable categories). The differences we hypothesize concern relative cross-cultural differences in people's tendency to engage in concrete vs. abstract thinking, as defined in the introduction and in Studies 1–3.
3 Arguably, all categories are relational, and are thus abstract, to some degree because every category relates the members of that category to each other (e.g., the category "dog" is more abstract than a specific dog that you can see, hear, or touch). However, different categories are abstract to varying degrees. Relational categories are more abstract than entity categories because they involve relationships beyond the relationship that exists between categories and their members. The difference between relational and entity categories is therefore relative.
4 Raven's SPM is often used as a measure of general intelligence (Brouwers, Van de Vijver, & Van Hemert, 2009; Carpenter, Just, & Shell, 1990; Laidra, Pullmann, & Allik, 2007; Lynn, Chan, & Eysenck, 1991; Van den Broek & Bradshaw, 1994) and as a non-verbal test to measure Spearman's g (Mackintosh & Bennett, 2005; Spearman, 1946, p.127).

Our goal is not to compare intelligence across cultures but to use the SPM as an operationalization of analogical reasoning, per se.

5 Culture significantly predicted SPM accuracy whether we used transformed or untransformed RTs in this model.

6 We preregistered both the SPM and a working memory task (forward digit span task; DST) as a control task. The DST was intended to test whether Chinese participants outperformed U.S. participants for abstract cognitive tasks only, or also for more concrete tasks. Thus, our exclusion criteria for Study 3b included correctly answering DST catch trials. The DST results are not reported here because, after data collection, we learned that DST performance differs between English and Mandarin speakers because Mandarin digit names have a shorter duration than English digit names on average – even for monosyllabic number words (see e.g., Chincotta & Underwood, 1997; Elliott, 1992; Mattys, Baddeley, & Trenkic, 2018). Thus, more digits can be maintained in the phonological loop in Mandarin than English. This difference produces a Mandarin speaker advantage. Thus, the DST is an unsuitable control task due to the documented influence of hypothesis-irrelevant linguistic factors.

7 These results have since been additionally replicated by an independent lab and reported in Cao et al. (2024). That work cites "Su (2020)", the Master's thesis of this paper's author S.S. supervised by A.S. and D.C.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Borghi, A. M., Mazzuca, C., & Tummolini, L. (2025). The role of social interaction in the formation and use of abstract concepts. *Nature Reviews Psychology*, *4*, 70–483.

Berry, J.W. (1967). Independence and conformity in subsistence-level societies. *Journal of Personality and Social Psychology*, *7*(4), 415–418.

Brouwers, S. A., Van de Vijver, F. J., & Van Hemert, D. A. (2009). Variation in Raven's Progressive Matrices scores across time and place. *Learning and Individual Differences*, *19*(3), 330–338.

Burgoon, E. M., Henderson, M. D., & Markman, A. B. (2013). There are many ways to see the forest for the trees: A tour guide for abstraction. *Perspectives on Psychological Science*, *8*(5), 501–520.

Cao, A., Carstensen, A., Gao, S., & Frank, M. C. (2024). United States–China differences in cognition and perception across 12 tasks: Replicability, robustness, and within-culture variation. *Journal of Experimental Psychology: General*, *153*(11), 2657–2685. https://doi.org/10.1037/xge0001559

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404.

Champely, S. (2020). pwr: Basic Functions for Power Analysis. Available at: https://cran.r-project.org/web/packages/pwr/pwr.pdf

Chan, M. J. (1994). Chinese and Australian concepts of intelligence. *Psychology and Developing Societies*, *6*(2), 103–117.

Chen, M. J., & Chen, H. C. (1988). Concepts of intelligence: A comparison of Chinese graduates from Chinese and English schools in Hong Kong. *International Journal of Psychology*, *23*(1-6), 471–487.

Chiu, L. H. (1972). A cross-cultural comparison of cognitive styles in Chinese and American children. *International Journal of Psychology*, *7*(4), 235–242.

Chincotta, D., & Underwood, G. (1997). Digit span and articulatory suppression: A cross-linguistic comparison. *European Journal of Cognitive Psychology*, *9*(1), 89–96.

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences*, *102*(35), 12629–12633.

Dong, X., Talhelm, T., & Ren, X. (2019). Teens in rice county are more interdependent and think more holistically than nearby wheat county. *Social Psychological and Personality Science*, *10*(7), 966–976.

Elliott, J. M. (1992). Forward digit span and articulation speed for Malay, English, and two Chinese dialects. *Perceptual and Motor Skills*, *74*(1), 291–295.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Thousand Oaks, CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Freitas, A. L., Gollwitzer, P., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, *40*(6), 739–752.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170.

Gentner, D., & Asmuth, J. (2019). Metaphoric extension, relational categories, and abstraction. *Language*, *Cognition and Neuroscience*, *34*(10), 1298–1307.

Gentner, D., & Hoyos, C. (2017). Analogy and abstraction. *Topics in Cognitive Science*, *9*(3), 672–693.

Granet, M. (1934). *La Pensée Chinoise*. Paris: Albin Michel.

Hafri, A., & Firestone, C. (2021). The perception of relations. *Trends in Cognitive Sciences*, *25*(6), 475–492.

Hale, S. C. (1988). Spacetime and the abstract/concrete distinction. *Philosophical Studies*, *53*(1), 85–102.

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

Hong, J., & Lee, A. Y. (2010). Feeling mixed but not torn: The moderating role of construal level in mixed emotions appeals. *Journal of Consumer Research*, *37*(3), 456–472.

Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition*, *54*(1), 73–98.

Ji, L. J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, *78*(5), 943–955.

Ji, L. J., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*, *87*(1), 57–65.

Klein, H. A., Lin, M. H., Radford, M., Masuda, T., Choi, I., Lien, Y., Yeh, Y., & Boff, K. R. (2009). Cultural differences in cognition: Rosetta Phase I. *Psychological Reports*, *105*(2), 659–674.

Krys, K., de Almeida, I., Wasiel, A., & Vignoles, V. L. (2025). WEIRD–Confucian comparisons: Ongoing cultural biases in psychology's evidence base and some recommendations for improving global representation. *American Psychologist*, *80*(2), 247–263. https://doi.org/10.1037/amp0001298

Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, *42*(3), 441–451.

Lefcheck, J. S. (2016). PIECEWISESEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579. https://doi.org/10.1111/2041-210X.12512

Lynn, R. (1991). Intelligence in China. *Social Behavior and Personality: An International Journal*, *19*(1), 1–4.

Lynn, R., Chan, J. W. C., & Eysenck, H. J. (1991). Reaction times and intelligence in Chinese and British children. *Perceptual and Motor Skills*, *72*(2), 443–452.

Lynn, R., Pagliari, C., & Chan, J. (1988). Intelligence in Hong Kong measured for Spearman's g and the visuospatial and verbal primaries. *Intelligence*, *12*(4), 423–433.

Ma, H. K., & Leung, M. C. (1991). Altruistic orientation in children: Construction and validation of the Child Altruism Inventory. *International Journal of Psychology*, *26*(6), 745–759.

Mackintosh, N. J., & Bennett, E. S. (2005). What do Raven's Matrices measure? An analysis in terms of sex differences. *Intelligence*, *33*(6), 663–674.

Markus, H. R., & Conner, A. (2014). *Clash!: How to thrive in a multicultural world*. New York: Penguin.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253.

Masuda, T., Gonzalez, R., Kwan, L., & Nisbett, R. E. (2008). Culture and aesthetic preference: Comparing the attention to context of East Asians and Americans. *Personality and Social Psychology Bulletin*, *34*(9), 1260–1275.

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, *81*(5), 922–934.

Masuda, T., & Nisbett, R. E. (2006). Culture and change blindness. *Cognitive Science*, *30*(2), 381–399.

Mattys, S. L., Baddeley, A., & Trenkic, D. (2018). Is the superior verbal memory span of Mandarin speakers due to faster rehearsal? *Memory & Cognition*, *46*, 361–369.

McKone, E., Davies, A. A., Fernando, D., Aalders, R., Leung, H., Wickramariyaratne, T., & Platow, M. J. (2010). Asia has the global advantage: Race and visual attention. *Vision Research*, *50*(16), 1540–1549.

Miyamoto, Y., Knoepfler, C. A., Ishii, K., & Ji, L. J. (2013). Cultural variation in the focus on goals versus processes of actions. *Personality and Social Psychology Bulletin*, *39*(6), 707–719.

Moser, D. J. (1996). *Abstract thinking and thought in ancient Chinese and early Greek* (Doctoral dissertation, University of Michigan).

Murphy, G. L., Bosch, D. A., & Kim, S. (2017). Do Americans have a preference for rule-based classification? *Cognitive Science*, *41*(8), 2026–2052.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383.

Nisbett, R. E. (2003). *The geography of thought*. New York: The Free Press.

Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences*, *100*(19), 11163–11170.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*(2), 291–310.

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, *26*(5), 653–684.

Paivio, A., Yuille, J. C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1, Pt. 2), 1–25.

Pashler, H. (2018). Building a more replicable experimental psychology: Key challenges. *The 59th Annual Meeting of the Psychonomic Society*, New Orleans, LA.

Posit team (2023). RStudio: Integrated Development environment for R. Posit Software, PBC, Boston, MA. Available at: http://www.posit.co/

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48.

Raven, J., & Raven, J. (2003). Raven progressive matrices. In R. Steve McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 223–237). Boston, MA: Springer.

Raven, J. C., & Court, J. (1996). *Manual for Raven's standard progressive matrices and vocabulary scales* (1996 edition). Oxford, England: Oxford Psychologists Press.

Reilly, J., Shain, C., Borghesani, V., Kuhnke, P., Vigliocco, G., Peelle, J. E., Mahon, B. Z., Buxbaum, L. J., Majid, A., Brysbaert, M., Borghi, A. M., Deyne, S. D.e, Dove, G., Papeo, L., PexmRRean, P. M., Poeppel, D., Lupyan, G., Boggio, P., Hickok, G., Gwilliams, L., Fernandino, L., Mirman, D., Chrysikou, E. G., Sandberg, C. W., Crutch, S. J., Pylkkänen, L., Yee, E., Jackson, R. L., Rodd, J. M., Bedny, M., Connell, L., Kiefer, M., Kemmerer, D., Zubicaray, G., Jefferies, E., Lynott, D., Siew, C. S. Q., Desai, R. H., McRae, K., Diaz, M. T., Bolognesi, M., Fedorenko, E., Kiran, S., Montefinese, M., Binder, J. R., Yap, M. J., Hartwigsen, G., Cantlon, J., Bi, Y., Hoffman, P., Garcea, F. E., & Vinson, D. (2025). What we mean when we say semantic: Toward a multidisciplinary semantic glossary. *Psychonomic Bulletin & Review*, *32*(1), 243–280.

Spearman, C. (1946). Theory of general factor. *British Journal of Psychology*, *36*(3), 117–131.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*, *344*(6184), 603–608.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463.

Unsworth, S. J., Sears, C. R., & Pexman, P. M. (2005). Cultural influences on categorization processes. *Journal of Cross-Cultural Psychology*, *36*(6), 662–688.

Uskul, A. K., Kitayama, S., & Nisbett, R. E. (2008). Ecocultural basis of cognition: Farmers and fishermen are more holistic than herders. *Proceedings of the National Academy of Sciences*, *105*(25), 8552–8556.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*(1), 3–15.

Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, *57*(4), 660–671.

Van den Broek, M. D., & Bradshaw, C. M. (1994). Detection of acquired deficits in general intelligence using the National Adult Reading Test and Raven's Standard Progressive Matrices. *British Journal of Clinical Psychology*, *33*(4), 509–515.

Wallace, A. R. (1871). *Contributions to the theory of natural selection: A series of essays*. New York: Macmillan and Co.

Wang, Q. (2013). *The autobiographical self in time and culture*. Oxford, England: Oxford University Press.

Wang, Q., & Ross, M. (2005). What we remember and what we tell: The effects of culture and self-priming on memory representations and narratives. *Memory*, *13*(6), 594–606.