

Cultural Variation in the Structure of Internal Representations: Testing the Orthographic Hypothesis

Ryutaro Uchiyama^{1,2} 

Abstract

Cultural differences comprise not only outwardly observable behaviors, but also internal psychological traits. One poorly understood domain of cross-cultural psychological variation is the organization of internal mental representations, and how this variation arises from experiential differences. Such an understanding could help reveal fundamental ways in which culture shapes the mind. Here we use the Internal Representations Questionnaire (IRQ) to investigate cross-cultural differences in modalities of thought such as visual imagery and internal speech. We compare respondents in China and Japan to the original US sample, testing the preregistered hypothesis that the structure of internal representations is associated with variation in writing systems. We found evidence of differences in factor structure between the US sample and the Japanese and Chinese respondents. An exploratory factor analysis for the Chinese data revealed that some aspects of inner speech are statistically

Cross-Cultural Research

2026, Vol. 0(0) 1–42

© 2026 SAGE Publications



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10693971261415636

journals.sagepub.com/home/CCR



¹Department of Psychological and Behavioural Science, London School of Economics and Political Science, UK

²Humanities, Arts and Social Sciences Cluster, Singapore University of Technology and Design, Singapore

Corresponding Author:

Ryutaro Uchiyama, Humanities, Arts and Social Sciences Cluster, Singapore University of Technology and Design, 8 Somapah Rd, 487372, Singapore.

Email: ryutaro_uchiyama@sutd.edu.sg

inseparable from orthographic imagery in this population—an outcome consistent with psycholinguistic and neuroimaging findings about the development of an orthography-to-semantics direct pathway in Chinese readers but not in alphabetic readers. These findings suggest the presence of meaningful cultural variation in the structure of internal representations, which may be a downstream consequence of variation in writing systems.

Keywords

cultural psychology, cultural variation, inner speech, mental imagery, orthographic imagery, writing systems, cultural evolution

Received: 6 August 2025; accepted: 30 December 2025

Introduction

In one common conception, cultural differences consist largely of differences in behavior. Across societies, differences in communication, cooperation, rituals, mating, eating, work, and leisure are readily observable. Even when we talk about differences of mentality or attitude between cultural groups, we often infer such psychological attributes through the lens of behavior. But there are many aspects of human psychology that are not expressed as overt behavior, being opaque to casual observation and even to controlled experimental study (Lupyan et al., 2023). Inferring the internal cognitive machinery that underpins overt behavior is a challenging “inverse problem”, and one that is intimately tied to our understanding of internal representations, defined here as the informational structures of the mind that model and track the structure of the world (Edelman, 2008). In order to examine true psychological variation, we must understand the structure of variation in internal representations.

The study of *cognitive style* is one research program that has contributed to the study of psychological variation (Kozhevnikov, 2007; Witkin & Moore, 1977). Cognitive style describes consistent differences in people’s mode or strategy of information processing, typically understood as a trade-off between different strategies rather than an aptitude measure like intelligence. A standard example of cognitive style is the visualizer–verbalizer continuum (Kirby et al., 1988; Kraemer et al., 2009, 2014; Mayer & Massa, 2003), which can be further decomposed for example into object-oriented visualizers, more common among artists, and spatially oriented visualizers, more common among scientists and engineers (Blazhenkova & Kozhevnikov, 2010; Kozhevnikov et al., 2005). Although this kind of individual-level variation is

frequently treated as noise in the design of cognitive experiments, a fine-grained analysis can offer a deeper understanding of underlying cognitive mechanisms (Ansari et al., 2003; Karmiloff-Smith et al., 2004; Kosslyn et al., 2002; Noppeney, Penny, et al., 2006; Noppeney, Price, et al., 2006).

In parallel with the study of individual differences in cognitive style within populations, there has also been a separate but substantial body of work on cognitive style across cultures. This research is nested within a larger body of work on cross-cultural differences in perception and cognition (Henrich et al., 2023), which has uncovered significant variation in domains ranging from memory (Wang, 2021) and spatial cognition (Majid et al., 2004) to moral judgments (Barrett et al., 2016), affect (Wei et al., 2025), motivation (Yanaoka et al., 2024; Zhang et al., 2024), and economic decision-making (Henrich et al., 2005; House et al., 2020). Cross-cultural psychological variation had been explored insufficiently for decades due to the WEIRD (Western, educated, industrialized, rich, democratic) people problem in the psychological and behavioral sciences (Apicella et al., 2020; Barrett, 2020; Henrich et al., 2010). A foundational discovery within this field was the finding that Western people tend to adopt an “analytic” cognitive style, which includes features such as attention to focal objects, ascription of causality to agents, and use of abstract categories, whereas East Asians tend to adopt a “holistic” cognitive style, with features such as attention to relationships among elements in the perceptual field, ascription of causality to situations, and use of relational categories (Choi et al., 1999; Kitayama et al., 2003; Masuda & Nisbett, 2001).

The divergence in cognitive styles occurs not only between East and West, but also among Eastern countries and among Western countries (Kitayama et al., 2009), as well as among regions within countries (Kitayama et al., 2006; Talhelm et al., 2014). These studies have linked the divergence in cognitive styles to historical factors such as subsistence method or exploration of geographical frontiers, each of which are thought to influence social structure (e.g., degree of interdependence), and are linked to socio-psychological variables such as the strength of social norms (Talhelm & English, 2020) or the perceived relationship between self and society (Markus & Kitayama, 1991). Some studies suggest that the key explanatory variable may instead be kinship intensity (i.e., how central kinship is to the formation of personal identity and social relationships) (Schulz et al., 2019). The causal processes linking these societal factors to cognition and perception—how we think about and see the world—are not well understood (Kitayama et al., 2009). At a proximal level, there is evidence of caretaker-to-child transmission of culturally typical cognitive styles via joint attention and shared discourse (Senzaki et al., 2016), but one missing piece in this overall picture is the population-level dynamics that induces systematic divergence across cultures.

Much of the work on cultural differences in internal representation has relied on this contrast between analytic and holistic processing, often at the expense of other dimensions of psychological variation—for instance, [Bruder and Zehra \(2025\)](#) find cross-cultural variation in the reported intensity of sensory imagery.

Analytic–holistic cognitive style has been repeatedly deployed as an explanation in the context of cultural variation, but importantly, it does not appear to be able to explain variation among individuals within a culture ([Kitayama et al., 2009](#); [Na et al., 2010](#)). Analytic–holistic cognitive style is thus likely to be a group-level trait generated by forces acting upon the cultural group itself as a unit. However, these results conflict with the sizeable literature on individual differences in cognitive style mentioned above ([Kozhevnikov, 2007](#); [Witkin & Moore, 1977](#)), which finds substantial variation within samples of individuals with the same cultural background, often for dimensions of variation that are highly similar to analytic–holistic processing ([Allinson & Hayes, 1996](#)). This discrepancy motivates a reassessment of cognitive style and indeed the structure of internal representations in general.

In the present study, we employ the Internal Representations Questionnaire (IRQ) ([Roebuck & Lupyan, 2020](#)) to investigate cross-cultural differences in the structure of internal representations. The IRQ is an instrument designed to probe individual differences in modalities of thought. In the original study, conducted with a US sample, a factor analysis reveals a 4-modality structure: *visual imagery*, *internal verbalization*, *orthographic imagery*, and *representational manipulation*.

Orthographic imagery refers to mental imagery specifically of text or written language. Representational manipulation refers to the ability to dynamically manipulate mental representations regardless of modality, as captured by items such as “I can easily imagine the sound of a trumpet getting louder”. Visual imagery and internal verbalization have been widely studied, but orthographic imagery and representational manipulation constitute novel modalities that are not typically discussed in the literature. Importantly, responses to the IRQ also predict performance on a cue-target matching task in a modality-selective manner, confirming predictive validity with respect to behavioral consequences.

Although the IRQ reveals notable findings about the population structure of internal representations, it is unclear how much of this structure can be attributed to the effect of genes, culture, or idiosyncratic experience. For example, if there turned out to be large variation in internal representations between cultures while controlling for genotype, and less variation between ancestry groups or genotypes while controlling for cultural upbringing, this

would suggest a larger effect of culture than of genes (although the two will often be correlated). It would also offer hints about the plasticity of internal representations, as cultural evolution occurs more rapidly than genetic evolution (Richerson et al., 2010). Such an understanding of the sources of variation in internal representations can help us assess, among other things, the extent to which internal representations are amenable to interventions in settings such as education, professional training, or psychotherapy. A cross-cultural analysis is thus an important first step in understanding the status of internal representations.

In particular, we collected data from Japan and China, two populations that differ from the original US sample across various social and cultural dimensions. Our hypothesis pertains to differences in writing systems (Handel, 2019). Whereas English is written in a phonetic alphabet, Chinese writing is logographic and thus attributes both a semantic meaning and a sound to each character. For example, the English word “river” is formed by concatenating 5 graphemes, each with their own phonetic representation, whereas the Simplified Chinese (Hanzi) character for river, 河, is a unitary grapheme comprising elements that cue meaning and sound. The vast majority of Chinese characters represent some meaning (often more than one meaning) on their own, rather than represent meaning only when joined with other characters. Japanese writing makes heavy use of Chinese-derived logograms. But whereas Chinese logograms are organized in a one-to-one mapping between character and sound, Japanese logograms are commonly associated with multiple sounds. Moreover, the Japanese writing complements its logographic system with two additional phonetic syllabaries, yielding a hybrid script that shares features of both the English and Chinese systems.

We test the hypothesis that differences in writing systems explain differences in the structure of mental representations—an idea contemplated by thinkers ranging from Leibniz (1697) to McLuhan (1962). The IRQ identifies orthographic imagery as a modality of thought, and this modality is quite plausibly impacted by orthographic input from the cultural environment. Neuroimaging studies find differences in the profile of neural activation between Chinese and English reading (Perfetti et al., 2013; Wu et al., 2012), and there are several behavioral and cognitive differences that emerge during acquisition of these writing systems as well (McBride, 2016). Because reading reorganizes the brain not only in areas that directly subserve reading but also in areas devoted to other functions through downstream effects, for example face perception (Dehaene et al., 2015), the effect of exposure to a given writing system may extend beyond orthographic imagery into other modalities of internal representation as well. Due to its relative recency in human history, literacy is generally acknowledged to be a cultural adaptation rather than a

genetic one. Therefore, an understanding of how writing systems shape internal representations can help us understand more generally how culture shapes the human mind, also moving beyond the sometimes simplistic identification of East and West with Holistic and Analytic thinking respectively.

Materials & Methods

Preregistration

We preregistered a number of details about the study and its analysis following the format of AsPredicted <https://aspredicted.org>. Preregistrations were submitted twice—once during collection of the Japanese data (before anyone was able to see the data), and once more prior to collection of the Chinese data. Both submissions are made public at the Open Science Framework at the URL <https://osf.io/nxmg2/registrations>.

The Japanese preregistration specified details such as exclusion criteria for participants, expected sample size, the procedure of questionnaire administration, and the translations of the items. With respect to the analysis, we stated that we would follow the same analysis as (Roebuck & Lupyan, 2020), and do so in the analysis below. After verifying the predicted statistical differences in observed scores across IRQ factors, between the new Japanese sample and the previous US data collected by Roebuck and Lupyan, we subsequently decided to extend the analysis by collecting additional data from a group that makes even more extensive use of logographic writing, namely Chinese Hanzi users in the People's Republic of China. The Chinese preregistration was thus written with knowledge of the findings from the Japanese sample, as indicated in the preregistration. In the Chinese preregistration, we included several predictions for how demographic variables would be associated with the IRQ factor scores.

In addition to the analysis following Roebuck and Lupyan, we also conducted a confirmatory factor analysis for both the Japanese and Chinese data, and an exploratory factor analysis for the Chinese data, each of these post-hoc. There was a mixed effects model analysis that we had declared in the Chinese preregistration but that we chose to omit here, as the analysis assumed the validity of the US factor structure for the Chinese and Japanese samples. The post-hoc confirmatory factor analyses suggested that the factor structure extracted by Roebuck and Lupyan for their US sample was compatible with neither the Japanese nor Chinese data. Similarly, although we had made predictions about how demographic variables of the Chinese participants would be associated with scores for the US-derived factors, we omitted this

analysis for the same reason. Instead, we performed the same analysis using the Chinese-derived factor structure.

Participants

We recruited participants in China and Japan through survey management companies in each country. The pre-exclusion Japanese sample consisted of 122 consenting participants, but 22 (18%) met the preregistered exclusion criteria by either failing one of the two attention check questions or by giving identical Likert responses to 90% or more of the main questionnaire items. When preregistering exclusion criteria for the Japanese sample, we had originally proposed to exclude participants who gave the same response on all items, but due to the discovery of a small number of participants who gave the same response on nearly all items, we subsequently shifted the criterion to 90% and applied it to analysis of the Japanese sample and preregistered it for the subsequent Chinese sample. This change in criterion did not impact the results in any meaningful way. The pre-exclusion Chinese sample consisted of 470 consenting participants, and only 1 participant failed any of the exclusion criteria, suggesting that the Chinese data might be of better quality than the Japanese data. There were an additional 2 participants who were excluded due to uninterpretable results, resulting in a final Chinese sample size of 467, all of whom are from regions where the standard dialect is Mandarin Chinese and the standard writing system is Simplified Hanzi. Sample sizes were determined by research budget.

For the Japanese sample, 57 (57%) participants reported their gender as male while 43 (43%) reported female, and the mean age was 53.7 with a range of 21 to 72, with 81% of the sample aged 45 or above ([Figure 1\(A\)](#)). For the Chinese sample, 234 (50%) participants reported their gender as male while 233 (50%) reported female, and the mean age was 31.5 years, resulting in a considerably younger group than the Japanese sample: the age range was 20 to 70 but 87% of the sample were in their 20 s or 30 s ([Figure 1\(B\)](#)). For the Chinese sample, we also obtained responses for several demographic and background variables beyond age and gender: years of education, hours per week spent on dense reading (e.g., books and newspapers but not social media), frequency of thinking in English, and frequency of using English in daily life ([Figure 1\(C\)–\(F\)](#)).

Ethical approval for the study was granted by the London School of Economics Research Ethics Committee (REC), which has US Department of Health and Human Services IORG (IRB organization) status. The REC case number for the present study is 19583. All aspects of the study were conducted in accordance with relevant guidelines and regulations endorsed by the REC.

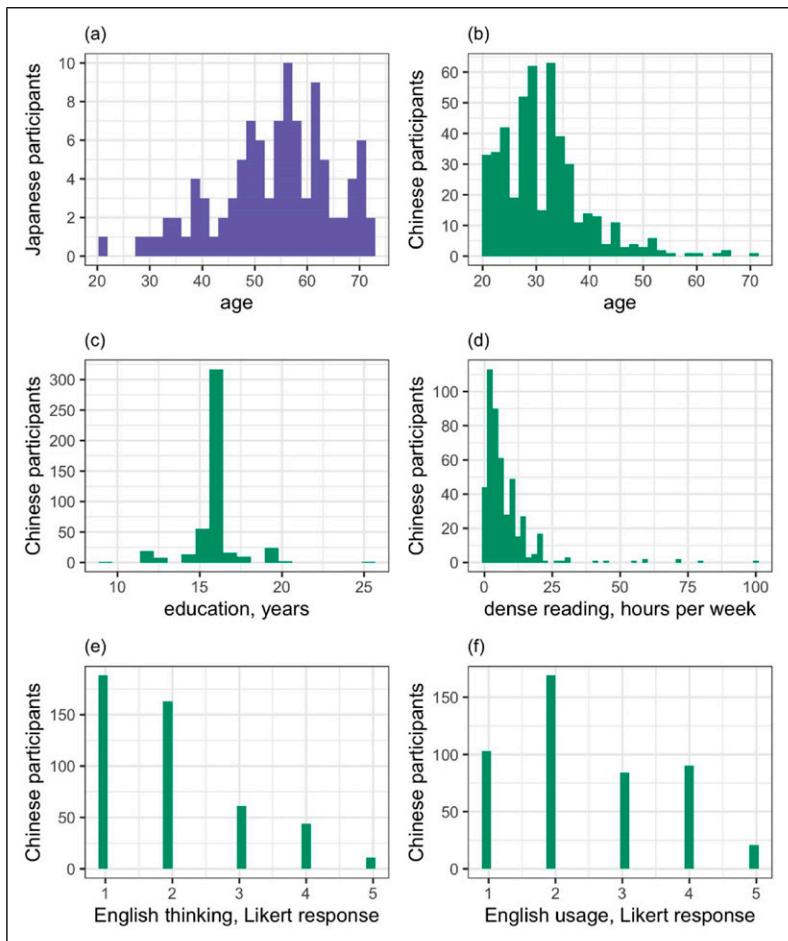


Figure 1. Self-reported demographic characteristics of the Japanese (A) and Chinese (B-F) samples. (E) English thinking indicates Likert responses to the Chinese translation of the statement “I frequently think in English” while (F) English usage indicates responses to the Chinese translation of the statement “I frequently use English in daily life (such as reading English texts, watching English films, engaging in English conversations, etc.)”. Likert responses range from “Strongly disagree” (1) to “Strongly agree” (5)

Informed consent was obtained from all respondents in their native languages, through an initial consent statement at the start of the questionnaire that described the purpose of the study, the possibility of their anonymized and aggregated data being published in academic venues, and the right to withdraw from the study at any point in time. Participants were transferred to the main questionnaire only if they had read through these terms and chose to accept them by selecting the acceptance option within the local survey management company's user interface.

Instrument

We administered the Internal Representations Questionnaire (IRQ) (Roebuck & Lupyan, 2020) after it was translated into Simplified Chinese (Hanzi) by a native Chinese speaker and into Japanese by a native Japanese speaker, each with professional-level competence in both their mother tongue and English. The IRQ consists of 36 items that probe the use of different forms of mental representation in everyday life. The questionnaire was originally constructed to investigate the role of internal verbalization in shaping perceptual and cognitive processing, but the researchers found 4 factors that each represent a different modality of internal representation: visual imagery, internal verbalization, orthographic imagery, and representational manipulation. The items in the IRQ were selected by an exploratory factor analysis on US samples consisting of university students and Amazon Mechanical Turk workers. The items of the IRQ and their factor assignment in the original study are listed in Table 9.

For both the Chinese and Japanese samples, the IRQ was administered through the smartphone interfaces employed by each survey management company. For each questionnaire item, participants were required to select a response from a 5-point Likert scale that consisted of the options “strongly disagree”, “disagree”, “neither agree nor disagree”, “agree”, and “strongly agree”. There were two reverse-coded items (items 13 and 33) whose response variables were inverted back for data analysis. The main questionnaire items were preceded by a consent question that allowed participants to opt-out of the study. The order of presentation of the main items was randomised, and the participant could only complete the study by providing responses for all questions. Two attention check questions were presented at randomized positions in the questionnaire.

Overview of Analysis

We first conducted simple comparisons of observed scores across the IRQ factors. The Chinese and Japanese scores were compared to the scores of the

US sample in [Roebuck and Lupyan \(2020\)](#); data published in online repository: <https://osf.io/8rdzh/>). Comparisons were made using both raw scores and within-culture standardized scores, the latter being a strategy to control for cross-cultural differences in response style ([Fischer, 2004](#)). As this simple comparison of observed scores was conducted without verification of the IRQ factor structure in the Chinese and Japanese samples, we performed a confirmatory factor analysis to evaluate the fit of the IRQ factors to the non-US samples and to test measurement invariance. The results were mixed, but taken in total suggested inadequate fit. To identify the difference in factor structure between the US and non-US data, we conducted an exploratory factor analysis only for the Chinese data, as the sample size of the Japanese data was insufficient. Finally, we obtained factor scores of the Chinese participants for the newly extracted factors, and used them in a regression analysis as outcome variables to be predicted by demographic variables.

Results

Cross-Cultural Comparison of Observed Scores

Comparison of Raw Scores. A comparison of raw observed scores between the 3 samples revealed salient cultural differences ([Table 1](#) and [Figure 2](#), top panel). The mean scores across all responses for the US, Chinese, and Japanese samples were 3.51, 3.52, and 3.15, respectively. The magnitude of US and Chinese responses were thus roughly similar on average, and both were about one-third of a Likert point higher than the Japanese responses. Despite their overall similarity in magnitude, US responses were considerably (more than half a Likert point) lower than Chinese responses on items associated with orthographic imagery. US responses were slightly higher than Chinese responses on items associated with the visual imagery and internal verbalization factors.

Japanese responses on average were lower than both the US and Chinese responses for visual imagery, internal verbalization, and representational manipulation. As the Japanese scores were closer to the mid-point of the 5-point Likert scale, this pattern may reflect either a middle response bias as previously reported in this population ([Chen et al., 1995](#); [Tasaki & Shin, 2017](#)), or a negative response bias relative to the US and Chinese samples. Despite their overall lower scores, Japanese participants were at roughly the same level as the US participants for items that load onto the orthographic imagery factor, although still lower than Chinese participants.

Table 1. Results of Welch's *t*-Tests for Simple Pairwise Comparison of Mean Scores Between US, Japanese (JP), and Chinese (CN) Samples, for the 4 Factors Extracted From the US Data in Roebuck and Lupyan (2020). The US Values are Computed From Data Published by Roebuck and Lupyan

Raw						Standardized			
Factor	Countries	t	df	p	Factor	Countries	t	df	p
Visual	US-JP	11.44	2181	<.0001	Visual	US-JP	-0.06	1832	0.950
	US-CN	4.27	4250	<.0001		US-CN	4.12	4532	<.0001
	JP-CN	-9.32	1573	<.0001		JP-CN	3.00	1445	0.003
	US-JP	16.52	2609	<.0001	Verbal	US-JP	5.67	2193	<.0001
Verbal	US-CN	5.55	5065	<.0001		US-CN	5.81	5400	<.0001
	JP-CN	-14.00	1871	<.0001		JP-CN	-1.99	1722	0.046
	US-JP	-1.67	1506	0.095	Orthographic	US-JP	-6.23	1263	<.0001
	US-CN	-13.59	2250	<.0001		US-CN	-12.72	2386	<.0001
Orthographic	JP-CN	-10.48	947	<.0001		JP-CN	-3.03	870	0.003
	US-JP	4.39	1435	<.0001	Manipulation	US-JP	-1.08	1202	0.280
	US-CN	-0.14	2459	0.886		US-CN	0.68	2619	0.494
	JP-CN	-5.18	988	<.0001		JP-CN	1.71	902	0.087

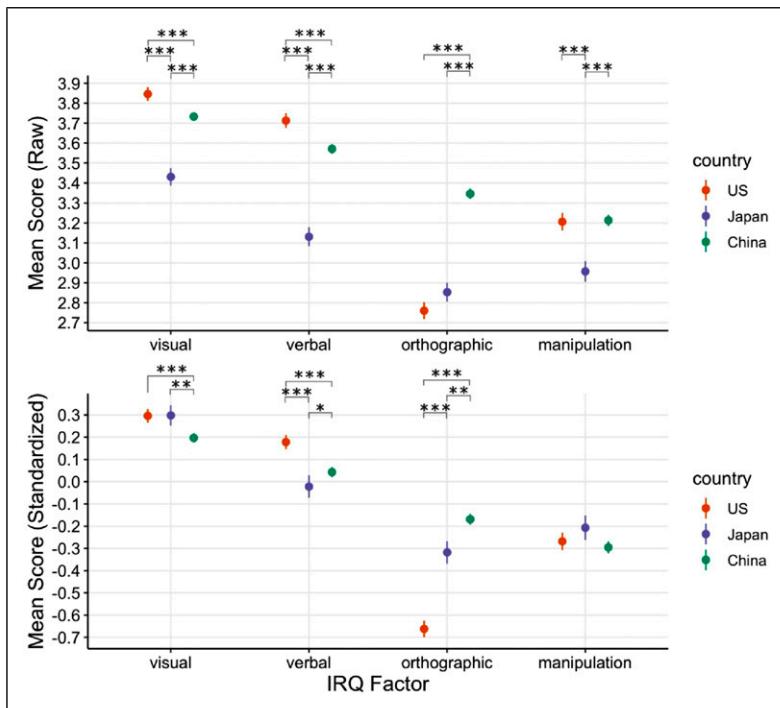


Figure 2. Comparison of raw means and within-culture standardized means of item responses grouped according to the factor structure extracted in Roebuck and Lupyán (2020). The US values are computed from data published by Roebuck and Lupyán (2020). Error bars are standard errors, and statistical significance levels derived from pairwise Welch's t-tests are indicated by asterisks (*: $p < .05$; **: $p < .01$; ***: $p < .001$)

Comparison of Standardized Scores. Within-culture standardized responses revealed cultural differences that are more readily interpretable than the raw score comparisons (Table 1 and Figure 2, bottom panel). Within each population, the mean score was set to 0 and the standard deviation was scaled to 1. All 3 groups yielded the highest scores on items that load onto the visual imagery factor, followed by items that load onto the internal verbalization factor. Scores for both of these factors were higher than scores for representational manipulation across all groups.

The greatest cross-cultural variation was observed in the orthographic imagery factor: US scores for these items were about half a standard score lower than the Chinese scores and about a third of a standard score lower than

the Japanese scores (US, $z = -0.66$; China: $z = -0.17$; Japan, $z = -0.32$). Orthographic imagery scores were thus particularly high in the two Asian samples compared to the US sample, and Chinese scores were slightly higher than Japanese scores. The US sample had noticeably higher scores on items that load onto the internal verbalization factor compared to the Japanese and Chinese samples. There are several other differences that can be statistically detected, but the findings above are relatively pronounced patterns that can be readily discerned from the standardized data. Taken at face value, these results suggest that orthographic imagery occupies a more prominent role in the inner mental life of Chinese and Japanese participants more than it does for participants in the US. The results also suggest that participants in the US may make greater use of internal verbalization than Chinese and Japanese participants.

There is no clear answer to what standardization procedure is most adequate in an analysis like this one, and the within-culture standardization approach that we adopt here is among the common methods employed in analysis of cross-cultural questionnaires (Fischer, 2004). We also tried an alternative method in which scores are standardized within individuals, yielding what are known as ipsative scores (Baron, 1996), but the change in mean scores for the factors was on the order of 0.002 to 0.02 standard scores, negligible for practical purposes.

Internal Reliability of IRQ Factors

To test the internal reliability of the factor structure extracted by [Roebuck and Lupyan \(2020\)](#) in the Chinese and Japanese data, we measured Cronbach's alpha ([Table 2](#)). The alpha coefficients measured in these new data are presented together with the values reported in the original US study for reference ([Roebuck & Lupyan, 2020](#)). The reliabilities of the Chinese sample

Table 2. Cronbach's Alpha Measure of Internal Consistency for Each of the 4 Factors Extracted From a US Sample in the Original Study. Internal Consistency for US Data is Computed From the Raw Data Published Online by [Roebuck and Lupyan \(2020\)](#)

Cronbach's alpha				
	Visual	Verbal	Orthographic	Manipulation
Chinese	0.55	0.74	0.70	0.68
Japanese	0.77	0.85	0.71	0.75
US	0.86	0.86	0.72	0.79

were overall lower than the US values, although mostly falling within a conventionally acceptable range ($\alpha > 0.7$). A noticeably low alpha coefficient was found for the visual imagery factor, whose value of 0.55 was far below that of the same factor in the US study, as well as below the conventional threshold. Reliability in the Japanese sample was at a very similar level to the US sample with the possible exception of the visual imagery factor, which was somewhat lower.

The analysis revealed some items whose removal increased internal reliability. Such increases were largely on the order of $\Delta\alpha = +0.01$, with the exception of one item linked to the visual imagery factor (item 10) in the Chinese sample, whose removal resulted in a large increase of 0.06. This item corresponded to the statement “If I imagine my memories visually they are more often static than moving”, suggesting a particularly poor fit of this item with respect to the other visual imagery items for Chinese participants.

Confirmatory Factor Analysis

A confirmatory factor analysis of the IRQ factors of [Roebuck and Lupyan \(2020\)](#) with the Chinese and Japanese data offered mixed results. Several goodness-of-fit indices demonstrated inadequate model fit ([Table 3](#)). In the Chinese sample, the criteria recommended by Hu and Bentler (45) (e.g., $\text{RMSEA} < .06$; $\text{SRMR} < .08$; $\text{CFI} > .95$; $\text{TLI} > .95$), a commonly cited reference, were met for RMSEA and SRMR but not for CFI or for TLI. In the Japanese sample, none of the indices met the recommended criteria. However, this was no worse than the US data set from which the model was initially constructed ([Roebuck & Lupyan, 2020](#)), where none of the fit indices successfully met the criteria (although a follow-up analysis with a larger sample,

Table 3. Goodness-of-Fit Indices for Confirmatory Factor Analyses With the Factor Structure Extracted by [Roebuck and Lupyan \(2020\)](#). US Values are Computed From Raw Data Published Online by Roebuck and Lupyan for a Preliminary Data Set. The Recommended Criteria for Each Index ([Hu & Bentler, 1999](#)) are Displayed on the Bottom Row

Goodness-of-fit indices							
Sample	N	χ^2	df	RMSEA	SRMR	CFI	TLI
Chinese	467	1187.88	588	0.047	0.061	0.788	0.773
Japanese	100	1022.88	588	0.086	0.1	0.661	0.636
US	222	1201.18	588	0.069	0.098	0.773	0.757
Recommended criteria				<0.06	<0.08	>0.95	>0.95

reported in Roebuck and Lupyan, exhibits better somewhat fit than this earlier sample). Goodness-of-fit measures for the US sample were better than the Japanese sample but worse than the Chinese sample. It was therefore unclear whether the IRQ factor model was a comparatively worse fit for the two new Asian samples compared to the data from the published US sample. Across all 3 samples, CFI and TLI were far from meeting the recommended criteria, while RMSEA and SRMR were fairly close to the threshold even when they fell short, such that under some other more lenient criteria (e.g., .46), they would pass as acceptable. A combination of low CFI/TLI and acceptable RMSEA/SRMR likely reflects low correlations among the variables, resulting in a condition where the specified factor model does not sufficiently improve the fit of model to data compared to a null model that includes only variances.

We tested measurement invariance to examine whether combining data from multiple countries worsens the fit of the factor model from the original US study (Table 4), using the criteria recommended by [Rutkowski and Svetina \(2014\)](#). The aggregated data comprising the US, Chinese, and Japanese samples demonstrated slightly poorer fit than the US data alone, but the difference did not exceed the criteria for configural invariance, thereby suggesting invariance of factor structure between the groups. A test for metric invariance was then conducted by constraining the factor loadings to be equal across the 3 groups. Again, the difference in the fit measures did not exceed the recommended criteria. We followed this with a test for scalar invariance, by constraining the item intercepts to be equal across groups. In this case the change in fit statistics exceeded the criteria due to a large change in CFI, although the change in RMSEA remained small and sub-threshold. The analysis indicated that factor structure (configural invariance) and factor loadings (metric invariance) but not intercepts (scalar invariance) were invariant across the 3 samples under the adopted criteria.

The same series of tests were also conducted in pairwise fashion for the US and Chinese samples and also for the US and Japanese samples. Similarly to the aggregate analysis with all 3 samples, metric but not scalar invariance was established for each of these groupings. The US–Chinese pair demonstrated slightly better fit measures for configural invariance compared to the US-only sample, and the US–Japanese pair demonstrated slightly worse fit. In sum, the fit of the factor loadings extracted in the original study does not noticeably decrease in the Chinese and Japanese samples, although the fit of the item intercepts do, thereby suggesting that the IRQ measures the same constructs across the three sampled cultures, but is limited in the degree to which actual item responses can be directly compared across cultures.

To further assess the adequacy of the original IRQ factor structure for the Chinese and Japanese samples, we inspected the intercorrelations among the

Table 4. Tests of Measurement Invariance for Three Combinations of Samples: {US, China, Japan}, {US, China}, and {US, Japan}, With Conventional Goodness-of-Fit Indices

Measurement invariance			Conventional Goodness-of-Fit Indices						
Included groups	Invariance test	χ^2	df	RMSEA	SRMR	CFI	TLI	Δ CFI	Δ RMSEA
US, China, Japan	Configural	3411.94	1764	0.060	0.074	0.758	0.741		
	Metric	3608.27	1828	0.061	0.082	0.739	0.730	-0.019	0.001
	Scalar	4293.69	1892	0.069	0.090	0.647	0.648	-0.091	0.009
	Configural	2389.05	1176	0.055	0.071	0.781	0.765		
	Metric	2512.24	1208	0.056	0.079	0.764	0.754	-0.016	0.001
	Scalar	2864.00	1240	0.062	0.084	0.706	0.702	-0.058	0.006
	Configural	2224.06	1176	0.074	0.096	0.758	0.718		
	Metric	2322.78	1208	0.076	0.097	0.739	0.708	-0.017	0.001
	Scalar	2650.24	1240	0.084	0.107	0.647	0.640	-0.074	0.008

factors in the 3 samples. The intercorrelations among the factors were high in the Chinese and Japanese sample, often by a factor of 2 compared to the intercorrelations reported by [Roebuck and Lupyan \(2020\)](#) ([Table 5](#)). This higher degree of similarity among the factors suggests that the IRQ factor structure extracted from the US sample is not appropriately capturing the variance present within the Chinese and Japanese data.

In summary, we uncovered mixed evidence about the extent to which the IRQ factors fit the Chinese and Japanese data. Several measures showed inadequate fit, but goodness-of-fit was not particularly worse than for the original US sample. Metric invariance was obtained for the three samples in aggregate as well as in a pairwise manner, but scalar invariance was not. The high factor intercorrelations for the Chinese and Japanese sample suggest that the IRQ factors are not nearly as well-separated for the Asian samples as they are in the US sample.

Exploratory Factor Analysis

Analytical Specifications. Due to ambiguous fit of the IRQ factor model with respect to the Chinese and Japanese samples, it is desirable to conduct an exploratory factor analysis to find an alternative factor structure that better captures the pattern of the data from the two East Asian societies. A better model may point us toward meaningful cross-cultural differences in the structure of internal representations. However, the Japanese sample ($N = 100$) was considerably smaller than the Chinese sample ($N = 467$), and a sample size of 100 falls short of many recommendations for sample size in exploratory factor analysis. The exact recommendation varies and often depends on other aspects of the analysis such as the number of items, the distribution of communalities, and factor loadings, but the minimum sample size appropriate for the intended analysis is about 200 participants ([Fabrigar et al., 1999](#); [Fabrigar & Wegener, 2012](#); [Hair et al., 2009](#); [Matsunaga, 2010](#)). We therefore conducted exploratory factor analysis only for the Chinese sample.

To first select the number of factors to be retained in an exploratory factor analysis of the Chinese data, we employed 3 selection methods—“optimal coordinates” ([Raîche et al., 2013](#)), “parallel analysis” ([Horn, 1965](#)), and “comparison data” ([Ruscio & Roche, 2012](#))—which were the 3 best-performing methods in [Ruscio & Roche \(2012\)](#) comparative analysis of methods for selecting number of factors. All of these methods indicated that retainment of 3 factors was optimal. A multivariate Shapiro-Wilk test for normality indicated that the Chinese data were not normally distributed ($W = 0.885, p < 0.0001$), so following the recommendation of [Costello and Osborne \(2005\)](#), we employed the principal axis factoring method. Although the data

Table 5. Intercorrelations Among the IRQ Factors for the Three Samples. The US Intercorrelations are Taken From Table I of Roeckbuck and Luyyan (2020)

IRQ factor	Chinese			Japanese			US		
	Visual	Verbal	Ortho.	Visual	Verbal	Ortho.	Visual	Verbal	Ortho.
Visual	~			~			~		
Verbal	0.80	~		0.69	~		0.47	~	
Orthographic	0.70	0.78	~	0.70	0.79	~	0.35	0.38	~
Manipulation	0.60	0.50	0.53	0.82	0.55	0.69	0.42	0.29	0.31

were not normal, they did satisfy both the Kaiser-Meyer-Olkin criterion ($KMO = 0.87$) and Bartlett's test for sphericity ($\chi^2(630) = 3358.62; p < .001$), where each of these results indicates adequacy for factor analysis. For factor rotation we followed [Roebuck and Lupyan \(2020\)](#), who used oblique factor rotation due to factor correlations, and employed oblimin, a standard method for oblique rotation.

The extracted factor structure had similarities with the original IRQ factors but also notable differences, and the factor loadings were generally low compared to the original US study ([Table 6](#)). The low factor loadings are likely to be partly due to a difference in procedure between this study and Roebuck and Lupyan—while the original study progressively narrowed down the number of questionnaire items from 81 to 36 based on their factor loadings and correlations with other items, the present study started with this finalized set of 36 items. Due to these relatively low factor loadings, we set the item loading cutoff to ± 0.3 . This cutoff is more lenient than Roebuck and Lupyan's criterion of ± 0.4 and more lenient than common recommendations for exploratory factor analysis, but also consistent with recommendations such as [Costello and Osborne \(2005\)](#). In the present case, a cutoff of ± 0.3 greatly enhances interpretability of the factors, allowing for a more meaningful comparison of the factor structure extracted from the Chinese data to the factor structure from the original US sample. Following the methodology of Roebuck and Lupyan, we included items whose factor loading exceeded the cutoff on one and only one factor, and excluded any items whose removal increased internal consistency (i.e., Cronbach's alpha) for the factor that it loads on. 12 items were dropped in total. Among these, 10 items were dropped due to not reaching the cutoff on any factor, 1 item was dropped due to exceeding the cutoff on more than one factor (item 11), and 1 item was dropped due to its removal increasing internal consistency (item 10). The final factor structure is shown in [Table 7](#), and the items in their original English rendition with both their US and Chinese factor names are listed in [Table 9](#).

Extracted Factor Structure. Factor 1 was loaded on by many of the items that were coded as internal verbalization in the original IRQ study, but it also included all of the orthographic imagery items, suggesting that these two modalities are not statistically separable in the Chinese population.

Factor 2 was the only factor that was loaded on by visual imagery items, and thus appears to primarily be a visual factor, although there were a number of items coded as internal verbalization that also loaded on this factor. Although it requires further study, the splitting of internal verbalization items between Factors 1 and 2 may be occurring along the lines of discursive vs. non-discursive items ([Alderson-Day et al., 2018](#); [McCarthy-Jones &](#)

Table 6. Factor Loadings of the 3-Factor Exploratory Factor Analysis With the Chinese Data. The Column Labeled “R&L Factor” Indicates the Corresponding Factor for That Item in the Original Study by [Roebuck and Lupyan \(2020\)](#). h^2 Indicates Communality, and u^2 Indicates Uniqueness, the Complement of Communality. “Drop” Indicates Whether the Item was Dropped From the Final Factor Structure on the Basis of the Criteria Noted in the Text

Item	R&L factor	Factor 1	Factor 2	Factor 3	h^2	u^2	Drop
1	Visual	−0.02	0.38	−0.01	0.14	0.86	No
2	Visual	0.06	0.32	0.18	0.19	0.81	No
3	Visual	0.14	0.37	0.14	0.25	0.75	No
4	Visual	0.29	0.12	0.12	0.17	0.83	Yes
5	Visual	0.00	0.32	−0.03	0.10	0.90	No
6	Visual	0.02	0.48	0.01	0.24	0.76	No
7	Visual	0.18	0.10	−0.06	0.05	0.95	Yes
8	Visual	0.14	0.25	−0.02	0.11	0.89	Yes
9	Visual	0.28	0.30	0.05	0.25	0.75	No
10	Visual	0.45	−0.32	−0.19	0.22	0.78	Yes
11	Verbal	0.34	0.31	−0.02	0.29	0.71	Yes
12	Verbal	0.38	0.20	−0.04	0.23	0.77	No
13	Verbal	0.34	0.20	−0.16	0.19	0.81	No
14	Verbal	0.06	0.43	−0.06	0.20	0.80	No
15	Verbal	0.45	0.06	0.00	0.23	0.77	No
16	Verbal	0.33	0.23	−0.03	0.22	0.78	No
17	Verbal	0.31	0.28	−0.11	0.22	0.78	No
18	Verbal	0.21	0.41	0.12	0.33	0.67	No
19	Verbal	−0.26	0.44	0.03	0.18	0.82	No
20	Verbal	0.51	−0.05	0.01	0.24	0.76	No
21	Verbal	0.28	0.24	−0.09	0.17	0.83	Yes
22	Verbal	0.28	0.27	0.05	0.23	0.77	Yes
23	Orthographic	0.65	−0.05	0.14	0.45	0.55	No
24	Orthographic	0.31	0.24	0.03	0.22	0.78	No
25	Orthographic	0.54	−0.02	0.03	0.29	0.71	No
26	Orthographic	0.43	0.03	0.08	0.23	0.77	No
27	Orthographic	0.52	0.11	0.05	0.34	0.66	No
28	Orthographic	0.27	0.13	0.06	0.13	0.87	Yes
29	Manipulation	0.09	0.00	0.73	0.56	0.44	No
30	Manipulation	0.37	0.08	0.17	0.24	0.76	No
31	Manipulation	0.07	−0.03	0.61	0.39	0.61	No
32	Manipulation	0.04	0.09	0.24	0.09	0.91	Yes
33	Manipulation	−0.14	0.00	0.62	0.37	0.63	No

(continued)

Table 6. (continued)

Item	R&L factor	Factor 1	Factor 2	Factor 3	h^2	u^2	Drop
34	Manipulation	0.27	0.18	0.24	0.26	0.74	Yes
35	Manipulation	0.17	0.13	0.22	0.14	0.86	Yes
36	Manipulation	0.01	0.19	0.29	0.16	0.85	Yes
	Variance explained	0.11	0.07	0.05			

(Fernyhough, 2011), where items with a discursive or reasoning-like quality load onto Factor 1 (e.g., item 15, “I tend to think things through verbally when I am relaxing”) whereas items that lack an explicit reasoning-like component (e.g., item 14, “My inner speech helps my imagination”) load onto Factor 2.

Factor 3 comprised only 3 items but with high loading. These items were all from the representational manipulation factor, and they were a subset that specifically concerned spatial manipulation of geometric constructs. The other items in the original representational manipulation factor pertained to other, non-spatial modalities—in particular verbal, gustatory, and auditory representations, so this factor appears to be strictly selective for spatial manipulation.

The intercorrelations of the factors for this exploratory factor analysis (Table 8) were substantially lower than the intercorrelations in the confirmatory factor analysis using the US factor structure (Table 5), suggesting that the new factors were comparatively well-separated. Internal reliability was reasonably good, and at a similar level to the confirmatory factor analysis (Table 8).

In sum, although the factor loadings were lower in the present analysis than they were in the original study by Roebuck and Lupyan (2020), several unique findings emerged: (1) A large portion of the items concerning internal verbalization loaded onto the same factor as the orthographic imagery items. (2) The visual imagery items clustered together, although they were also associated with a number of items related to internal verbalization that may be defined by their absence of the discursive component noted above. (3) Spatial manipulation of geometric constructs constituted its own factor. To encapsulate these provisional findings, we refer to Factor 1 as the “ortho-verbal” factor, Factor 2 as the “visuo-verbal” factor, and Factor 3 as the “spatial manipulation” factor (Table 7). The validity of these constructs is not yet clear without a confirmatory factor analysis on a new sample. A comparison of the factor labels from the original US study and the new factor labels provided in the present analysis is given in Table 9. Finally, the intercorrelations revealed that these extracted factors were much better separated than the US

Table 7. Factor Loadings for the Factor Structure Given by the Exploratory Factor Analysis, After Items Have Been Dropped. Factor Loadings Below the Cutoff of 0.3 Have Been Removed for Ease of Interpretation. The IRQ Factor Column Indicates the Corresponding Factor in the Original Study by [Roebuck and Lupyan \(2020\)](#). Factor 1 is Dubbed the “Ortho-verbal” Factor; Factor 2 is Dubbed the “Visuo-Verbal” Factor; Factor 3 is Dubbed the “Spatial Manipulation” Factor, See Text

Item	IRQ factor	Factor 1 ortho-verbal	Factor 2 visuo-verbal	Factor 3 spatial manipulation
1	Visual		0.38	
2	Visual		0.32	
3	Visual		0.37	
5	Visual		0.32	
6	Visual		0.48	
9	Visual		0.30	
12	Verbal	0.38		
13	Verbal	0.34		
14	Verbal		0.43	
15	Verbal	0.45		
16	Verbal	0.33		
17	Verbal	0.31		
18	Verbal		0.41	
19	Verbal		0.44	
20	Verbal	0.51		
23	Orthographic	0.65		
24	Orthographic	0.31		
25	Orthographic	0.54		
26	Orthographic	0.43		
27	Orthographic	0.52		
29	Manipulation			0.73
30	Manipulation	0.37		
31	Manipulation			0.61
33	Manipulation			0.62

sample-derived factors as tested in the confirmatory factor analysis, and internal reliability was mostly adequate ([Table 8](#)).

Association of Factor Scores With Participant Characteristics

Using the factor structure extracted by the exploratory factor analysis, we computed factor scores for each participant in the Chinese sample. A factor

Table 8. Factor Intercorrelations and Internal Reliability. Cronbach's Alpha are for Values After Items Were Dropped According to the Criterion Noted in the Text

Factor	1	2	3	α
1	~			0.79
2	0.38	~		0.66
3	0.24	0.28	~	0.71

score is a weighted average of a participant's responses across the items that measure a given factor, and is a more accurate measure of the participant's placement on that factor than sum scores of observed responses. In order to gain further insight into the participant characteristics of the Chinese sample that explain variation in internal representations, we conducted a multiple regression analysis with demographic variables as predictors and factor scores as outcomes (Table 10). Age and time spent reading were log-transformed due to a heavy positive skew in each, and all variables were standardized except for gender. For identification of gender we offered participants the options "male", "female", and "other", but all participants in the sample selected either male or female, rendering the variable dichotomous. One participant was excluded from this analysis due to their reported years of education being an impossible number.

We found a gender effect for the spatial manipulation factor: male factor scores were on average 0.24 standard deviation units higher than female factor scores. Although self-report is prone to biases in self-evaluation, this outcome is consistent with the widely replicated finding that male participants have an advantage over female participants in spatial cognition tasks such as mental rotation (Levine et al., 2016). For factor scores across all 3 factors, there was a positive effect of the reported (log-transformed) hours per week spent on intensive reading (on books and newspapers rather than, e.g., social media). Moreover, the magnitude of association was roughly equal across the 3 factors (ortho-verbal, $\beta = 0.11$; visuo-verbal, $\beta = 0.13$; spatial manipulation, $\beta = 0.10$), suggesting that reading is associated with the self-perceived strength of internal representations regardless of modality.

In this regression we included two variables designed to index the participant's immersion in the English language. One is English thinking, which encodes Likert responses to a Chinese statement that corresponds to, "I frequently think in English." The other is English usage, which similarly encodes Likert responses but to a statement that corresponds to "I frequently use English in daily life (such as reading English texts, watching English films, engaging in English conversations, etc.)." We included these variables as a

Table 9. IRQ Items With Factor Labels From Both the Original US Study and the Exploratory Factor Analysis in the Original Study. Items 19 and 33 Were Reverse-Coded. Blank Cells in the Chinese Factor Column are Items That Were Dropped Based on the Procedure Described in the Text. IRQ Items are Redrawn, With Permission, from [Roebuck & Lupyan \(2020\)](#)

Item	US factor	Chinese factor	Statement
1	Visual	Visuo-verbal	I often enjoy the use of mental pictures to reminisce
2	Visual	Visuo-verbal	I can close my eyes and easily picture a scene that I have experienced
3	Visual	Visuo-verbal	My mental images are very vivid and photographic
4	Visual		The old saying “A picture is worth a thousand words” is certainly true for me
5	Visual	Visuo-verbal	When I think about someone I know well, I instantly see their face in my mind
6	Visual	Visuo-verbal	I often use mental images or pictures to help me remember things
7	Visual		My memories are mainly visual in nature
8	Visual		When traveling to get to somewhere I tend to think more visually than verbally
9	Visual	Visuo-verbal	If I talk to myself in my head it is usually accompanied by visual imagery
10	Visual		If I imagine my memories visually they are more often static than moving
11	Verbal		I think about problems in my mind in the form of a conversation with myself
12	Verbal	Ortho-verbal	If I am walking somewhere by myself, I often have a silent conversation with myself
13	Verbal	Ortho-verbal	If I am walking somewhere by myself, I frequently think of conversations that I've recently had
14	Verbal	Visuo-verbal	My inner speech helps my imagination
15	Verbal	Ortho-verbal	I tend to think things through verbally when I am relaxing
16	Verbal	Ortho-verbal	When thinking about a social problem, I often talk it through in my head
17	Verbal	Ortho-verbal	I like to give myself some down time to talk through thoughts in my mind
18	Verbal	Visuo-verbal	I hear words in my “mind’s ear” when I think
19	Verbal	Visuo-verbal	I rarely vocalize thoughts in my mind

(continued)

Table 9. (continued)

Item	US factor	Chinese factor	Statement
20	Verbal	Ortho-verbal	I often talk to myself internally while watching TV
21	Verbal		My memories often involve conversations I've had
22	Verbal		When I read, I tend to hear a voice in my "mind's ear"
23	Orthographic	Ortho-verbal	When I hear someone talking, I see words written down in my mind
24	Orthographic	Ortho-verbal	I see words in my "mind's eye" when I think
25	Orthographic	Ortho-verbal	When I am introduced to someone for the first time, I imagine what their name would look like when written down
26	Orthographic	Ortho-verbal	A strategy I use to help me remember written material is imagining what the writing looks like
27	Orthographic	Ortho-verbal	I hear a running summary of everything I am doing in my head
28	Orthographic		I rehearse in my mind how someone might respond to a text message before I send it
29	Manipulation	Spatial manipulation	I can easily imagine and mentally rotate three-dimensional geometric figures
30	Manipulation	Ortho-verbal	I can easily choose to imagine this sentence in my mind pronounced unnaturally slowly
31	Manipulation	Spatial manipulation	In school, I had no problems with geometry
32	Manipulation		It is easy for me to imagine the sensation of licking a brick
33	Manipulation	Spatial manipulation	I find it difficult to imagine how a three-dimensional geometric figure would exactly look like when rotated
34	Manipulation		I can easily imagine someone clearly talking, and then imagine the same voice with a heavy cold
35	Manipulation		I think I have a large vocabulary in my native language compared to others
36	Manipulation		I can easily imagine the sound of a trumpet getting louder

Table 10. Regression Analyses Using Demographic and Background Variables of the Chinese Participants to Predict Their Factor Scores Across the 3 Factors Extracted Above From the Exploratory Factor Analysis. All Variables Other Than Gender are Standardized, and Age and Dense Reading are Log-Transformed. Gender is Dichotomous, and Coded as 1 = Male, 2 = Female

	Ortho-verbal			Visuo-verbal			Spatial manipulation			
	β	SE	t	β	SE	t	β	SE	t	p
English usage	0.03	0.05	0.64	0.520	0.07	0.51	0.190	0.12	0.05	0.032
English thinking	0.29	0.05	5.57	<0.001	0.17	0.05	3.23	0.001	0.09	0.072
Intensive reading	0.11	0.04	2.88	0.004	0.13	0.04	3.26	0.001	0.10	0.017
Age	-0.01	0.04	-0.30	0.765	0.04	0.04	0.97	0.335	0.18	0.45
Gender	-0.10	0.08	-1.15	0.251	-0.04	0.08	-0.54	0.587	-0.24	0.08
Education	-0.08	0.06	-1.41	0.158	-0.05	0.06	-0.77	0.439	0.07	0.06
R^2	0.153			0.097			0.084			

proxy for familiarity with Alphabetic writing systems, despite their likely confounding with other variables such as socio-economic status. English thinking and English usage are highly correlated (Pearson's $r = 0.66$), but their variance inflation factors are sufficiently low (English thinking, VIF = 1.81; English usage, VIF = 1.86), suggesting that collinearity is not an immediate problem.

For the ortho-verbal (Factor 1) and visuo-verbal (Factor 2) factors, factor scores were predicted by English thinking (ortho-verbal, $\beta = 0.29$, $p < .001$; visuo-verbal, $\beta = .17$, $p < .001$) but not English usage (ortho-verbal, $\beta = 0.03$, $p = 0.52$; visuo-verbal, $\beta = .07$, $p = 0.19$). English thinking may be predicting ortho-verbal and visuo-verbal scores simply by functioning as additional measurement items of these latent factors—a suspicion that is supported by a test of internal reliability, where Cronbach's alpha slightly increased when English thinking was included as a measurement item and remained constant for English usage. Under this scenario, “I frequently think in English” may just be another statement about internal verbalization in general. In contrast to the other two factors, spatial manipulation (Factor 3) was associated with a weak and marginally non-significant effect of English thinking ($\beta = .09$; $p = .072$), and a similar albeit significant effect of English usage ($\beta = .12$; $p = .032$). When the same regression was conducted without English thinking, English usage predicted factor scores with roughly equal magnitude across the three factors (ortho-verbal, $\beta = .22$; visuo-verbal, $\beta = .18$; spatial manipulation, $\beta = .18$; all $p < 0.01$), thereby suggesting that English usage, like reading, is associated with the strength of internal representations in general, regardless of modality. Therefore, for the ortho-verbal and visuo-verbal factors, the effect of English usage is masked by English thinking, where the latter may in fact be functioning just as a measurement item of these factors as noted above.

In sum, the analysis of factor scores revealed an effect of gender for the spatial manipulation factor, and what plausibly appear to be general effects of reading and English usage across all three factors, despite the masking of English usage by English thinking in the ortho-verbal and visuo-verbal factors. The impact of English immersion (thinking and usage) on factor scores is not yet clear.

Discussion

Cultural psychology has revealed substantial cross-cultural variation in perceptual processing, especially for vision, and has demonstrated the correspondence of such perceptual differences with other cultural variables such as social interdependence—*independence* (Kitayama et al., 2009). This body of research has supplied compelling evidence that the organization of the human

mind is permeable to cultural influence, but has often focused on analytic–holistic cognitive style at the expense of other possible dimensions of variation. To extend the scope of cross-cultural psychological inquiry, we employed the Internal Representations Questionnaire (IRQ) (Roebuck & Lupyan, 2020), an instrument designed to probe individual differences in qualitative modalities of thinking. Although there is a large body of research on modalities of thinking such as the visualizer–verbalizer continuum (Kirby et al., 1988; Mayer & Massa, 2003), the IRQ is a unique, bottom-up approach to the investigation of the structure of internal representations.

By administering the questionnaire to new cultural populations, we investigated both cross-cultural and within-culture individual differences in internal representation. In particular, we studied people in Japan and the People's Republic of China, under the hypothesis that variation in writing systems (Handel, 2019) may account for meaningful variation in internal representations across cultures.

Summary of Outcomes

A simple comparison of raw and standardized scores using the factor structure extracted from the US sample in the original study (Roebuck & Lupyan, 2020) revealed substantive differences between cultures (Table 1; Figure 2). After using within-culture standardization to reduce the effect of culture-specific response styles (Fischer, 2004), Chinese and Japanese scores were considerably higher than US scores on the orthographic imagery factor, and US scores were higher than Chinese and Japanese scores on the internal verbalization factor. There were other cross-cultural differences as well, including between the Chinese and Japanese samples, but the magnitude of these findings were smaller.

In our preregistration prior to collecting the Chinese data, we had predicted that the Chinese scores for orthographic imagery would be similar to or higher than the Japanese scores for the same IRQ factor, and that the Chinese scores for internal verbalization would be similar to or lower than the Japanese scores for the same factor. Prior to preregistering, we had observed that the Japanese participants reported higher orthographic imagery and lower internal verbalization than the US participants, and reasoned that Chinese participants should exhibit the same contrast but in a more pronounced manner, due to written Chinese being a purer logographic system while written Japanese can be considered as intermediate between written English and written Chinese due to its combination of logographic and phonetic writing systems. In the standardized comparison (Figure 2, bottom), our prediction about Chinese orthographic imagery scores turned out to be accurate. The results for internal

verbalization were less notable, as mean internal verbalization scores were only 0.02 standard scores smaller in the Chinese sample than they were in the Japanese sample, but were nonetheless consistent with the preregistered prediction. However, the mean age of the Japanese sample was 22 years older than the mean age of the Chinese sample, and a more balanced comparison of these two groups would require age-matched samples, especially given reports about age-related differences in mental imagery (Floridou et al., 2022; Kemps & Newson, 2005).

We observed mixed results about whether the original factor structure of Roebuck and Lupyan, derived from their US sample, was a good fit for the Chinese and Japanese data. A confirmatory factor analysis yielded ambiguous goodness-of-fit measures and factor intercorrelations in the Chinese and Japanese data that were considerably higher than in the US data. We therefore conducted an exploratory factor analysis with the Chinese data but not the Japanese data, due to a limitation in sample size for the latter. The analysis revealed a 3-factor structure: (1) an “ortho-verbal” factor that comprises orthographic imagery as well as some internal verbalization items that may have in common a discursive character, (2) a “visuo-verbal” factor that comprises visual imagery as well as some internal verbalization items that may have in common a non-discursive character, and (3) a “spatial manipulation” factor that is a subset of the representational manipulation factor of the original IRQ study, containing items related to the manipulation of geometric objects but excluding other modalities of representational manipulation.

Using the extracted 3-factor structure to further analyze the Chinese data, a comparison of factor scores with demographic variables revealed a number of findings. Among individuals who reported more time spent reading or immersed in the English language (i.e., English thinking or English usage), higher scores were observed across all 3 factors. This may indicate that engagement with linguistic material—whether in the form of immersion in a foreign language or in reading—is associated with more vivid internal representations overall. Relatedly, Roebuck and Lupyan (2020) had found that mean responses are correlated across factors rather than exhibiting a tradeoff between the different factors, despite research on cognitive styles (e.g., visualizer–verbalizer) often assuming a tradeoff (Mayer & Massa, 2003).

Male participants had higher factor scores on the spatial manipulation factor than female participants, as predicted based on a large body of past research (Levine et al., 2016). A gender effect was present only for this one factor.

Interpretation of the Chinese Factors

The factor structure extracted from the Chinese sample differed from the factor structure reported by Roebuck and Lupyan (2020) for their US sample. The

structure revealed here may serve to point us toward qualitative differences in the organization of internal representations between Chinese and US individuals.

Ortho-Verbal Conjunction. The joining of orthographic imagery and internal verbalization mirrors the notion that alphabetic reading involves extensive conversion of visuo-orthographic input into phonological representations, while Chinese reading involves more sustained activation of both orthographic and phonological representations (Perfetti et al., 2013; Xu et al., 1999). This difference is proposed to be due to a structural property of Chinese characters, namely how they primarily encode semantic meaning and only subordinately phonemic information, in contrast to a more direct phonological encoding in alphabetic symbols.

In Chinese orthography, tens of thousands of units of meaning (on the order of thousands for everyday use) are each represented with a dedicated character, and this large array is in turn mapped onto a much narrower set of just several hundred toned syllables. This is unlike in English, where isolated graphemes usually do not represent meaning in themselves, but only sounds. This mapping of a large set of logograms onto a smaller set of sounds results in a high density of homophony, where any given phonemic (syllabic) representation is likely to map onto multiple characters and hence multiple meanings (Figure 3(A)). This ambiguity incentivizes the development of a direct route of cognitive access from orthography to meaning that is unmediated by phonology (Figure 3(B)), as orthography carries considerably more information than phonology in such a writing system (Perfetti et al., 2013; Tan et al., 2005; Wu et al., 2012).

Neuroimaging studies reveal a developmental divergence in cortical responses to orthographic input when comparing Chinese- and English-reading children. These data suggest that Chinese readers, compared to English readers, exhibit more sustained activation of visuo-orthographic representations in parallel with phonological representations, and that this sustained activation is subserved by cortical regions such as the superior parietal lobule, the inferior temporal gyrus, and the middle occipital gyrus, all of which are areas involved in visuo-orthographic analysis (Cao et al., 2009, 2010, 2014). This cross-cultural neurocognitive divergence is thus best explained as a difference in the processing demands of the two writing systems, resulting in different learning trajectories. A genetic explanation for this divergence is far less plausible, due to the historical recency of literacy.

These structural differences between the two writing systems also explains the discrepancy between the orthographic imagery factor that was extracted from the US sample by Roebuck and Lupyan and the composite ortho-verbal

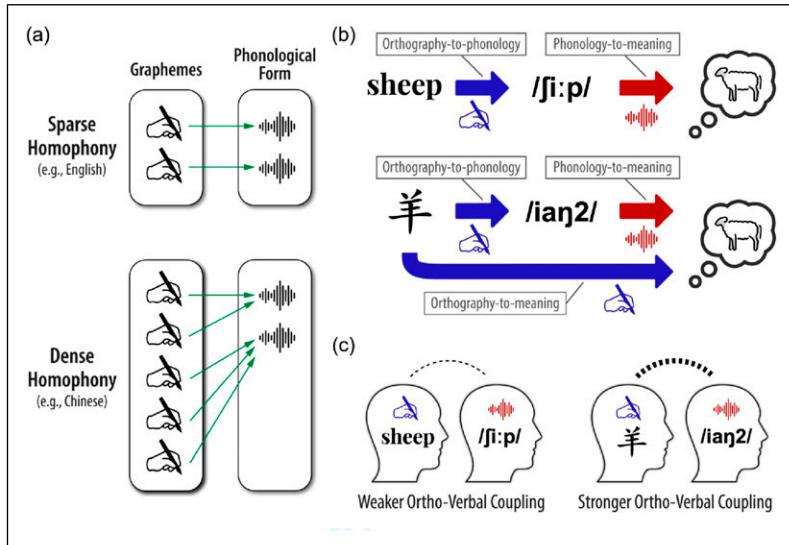


Figure 3. (A) Scripts with dense homophony (e.g., Chinese) – unlike scripts with sparse homophony (e.g., written English) – entail a many-to-few mapping from graphemes to phonological form, thus yielding ambiguity if semantic meaning is decoded solely from phonological representations of written language. (B) This structural difference between writing systems plausibly explains existing neuroimaging evidence for stronger parallel encoding of phonological and orthographic representations during reading in Chinese than in English readers (see text for details). Illustrated here are orthographic, phonetic, and semantic representations of “sheep” in English and Chinese. (C) Parallel encoding of semantic meaning may explain our finding of the statistical inseparability of orthographic and verbal imagery (i.e., “ortho-verbal conjunction”) in Chinese but not English readers

factor that was extracted from the Chinese sample in the present study. For the Chinese sample, there was no statistical separation between orthographic imagery and at least some components of internal verbalization, suggesting a stronger coupling between these two modalities of representation (“ortho-verbal conjunction”) compared to the US sample (Figure 3(C)). On the present explanatory account, this coupling arises from the parallel encoding of phonological and orthographic representations during reading, which arises as a learned neurocognitive adaptation to a literacy environment with dense homophony. More broadly, the account predicts that cultural variation in information environments influences cultural variation in the structure of internal representations (Kroupin et al., 2025).

Discursive Versus Non-Discursive Verbalization. It is not self-evident why the internal verbalization-related items subsumed by the ortho-verbal factor appear to share a discursive character. However, previous research using the Varieties of Inner Speech Questionnaire (VISQ and VISQ-R) found that self-reports about inner speech can be decomposed into multiple factors, one of which is a factor for “dialogic” inner speech (Alderson-Day et al., 2018; McCarthy-Jones & Fernyhough, 2011) that overlaps considerably with the discursive items that loaded onto the ortho-verbal factor in the Chinese sample. This proposed sub-division of internal verbalization is consistent with the factor structure extracted in the present study. It also means that it may be possible in a future study to use items from the VISQ to distinguish between dialogic and non-dialogic items, and explore the robustness of the apparent separation of these items in the extracted factor structure.

If internal representations of orthographic symbols are directly linked to atomic units of semantic meaning in Chinese-readers (through the “direct” pathway discussed above), then it is plausible that in the same population, internal representations of higher-order orthographic structures such as sentences or paragraphs are directly linked to higher-order units of meaning such as discourse and narrative. On this scenario, Chinese readers would be able to comprehend discursive meaning with relatively less reliance on internal verbalization, whereas in English readers, discursive meaning is obligatorily tied to internal verbalization. The discursive items of the questionnaire may be occupying the same factor as orthographic items in the Chinese factor structure as another downstream consequence of this direct processing pathway.

The visuo-verbal factor also subsumed a number of items that were, in the original US study, associated with internal verbalization. These items appeared to have in common a lack of the discursive component. Speculatively, they may pertain more to the immediate sensation or action of vocalization, rather than discourse, but the number of these items was too small to allow any measured judgment of their collective properties. One of these items, item 14 (“My inner speech helps my imagination”), carries an implicit connection to the visual modality insofar as imagination is commonly construed visually, but the other two do not do so in any obvious manner. It is not clear why visual imagery would be merged with internal verbalization, whether discursive or not, nor whether this is a robust finding in the first place. The organization of this factor will require further study.

Spatial Manipulation. The representational manipulation factor of Roebuck and Lupyan (2020) was reduced to a subset pertaining specifically to spatial manipulation. The coherence of this subset was strong, with the items loading

on this spatial manipulation factor having the highest factor loadings among all the questionnaire items. Although the cause of this pattern is unclear, it is plausible that some component of the Chinese cultural environment, such as the educational curriculum, tends to decouple spatial manipulation from other modalities of representational manipulation when compared to the US population.

Broader Outlook and Future Directions

The present study compares the structure of internal representations across samples in the United States, Japan, and the People’s Republic of China, under the hypothesis of a causal role played by writing systems. Although we find preliminary evidence supporting our orthographic hypothesis, additional studies are required for more robust validation and qualification. For example, investigation of Asian populations that employ alphabetic scripts—like Vietnamese, Malaysian Malay, most Indonesian, and Mongolian populations—can help resolve the role of cultural differences other than writing systems as potential confounds. The direction of causality itself also requires validation, as it remains possible that cross-cultural differences in internal representations—driven by some factor other than orthography, for instance genetics or social organization—may account for the form of writing systems, rather than the converse. Investigation of non-literate or minimally literate sub-populations would partly help resolve this ambiguity, as well as confer valuable insights regarding the gradations of the orthographic effect on imagery.

Beyond differences in internal representation induced by logographic vs. alphabetical writing systems, we may observe subtle differences within each system. For example, the “deeper” orthographies of English and French may exhibit signs of ortho-verbal conjunction to a greater extent than “shallower” orthographies like Finnish and Italian, due to denser homophony in the former (Seymour et al., 2003).

Finally, our framework suggests a structural coupling between writing systems and internal representations, thus offering a conceptual inroad toward the inference of population-level changes in mental imagery driven by historical changes in orthography (Han et al., 2022; Kelly et al., 2021; Morin & Koshevoy, 2024). In addition to this possibility of historical inference, the framework also supports the prediction of ongoing and future changes. For example, the widespread adoption of digital interfaces for reading and writing has reportedly precipitated a “character amnesia” among users of Chinese script in recent years (Huang et al., 2021), with plausible consequences for the structure of their internal representations. Artificial intelligence is likely to

instigate further, possibly dramatic, changes in literacy practices, thus potentially driving systematic changes in imagery and internal representations (Clark, 2025; Oakley et al., 2025). Although only a first step, our study sketches out this functional relationship between cultural technologies of literacy and our internal mental lives.

Conclusion

Administering the Internal Representations Questionnaire (IRQ) to Chinese and Japanese samples, we obtained evidence about cross-cultural differences in the structure of internal representations. These populations were appropriate for testing the hypothesis that variation in writing systems induces variation in internal representations. A naive comparison of item responses using the factor structure extracted from the original US study demonstrated that respondents from the two east Asian cultures had higher scores for orthographic imagery and lower scores for internal verbalization compared to respondents from the US sample, a finding that is aligned with basic features of their respective writing systems. A confirmatory factor analysis raised doubt about whether the US factors were appropriate for the two new samples, so we performed an exploratory factor analysis on the Chinese data and extracted a 3-factor structure that exhibited notable differences from the US factor structure. The extracted factor structure indicated differences in the organization of internal representations between Chinese and US participants, revealing findings that are consistent with data from cross-cultural studies on the psycholinguistics and functional neuroimaging of reading. In particular, some components of internal verbalization were statistically inseparable from orthographic imagery, suggesting that the two are closely tied together in Chinese but not US participants. This may be a downstream consequence of differences in learned neurocognitive adaptations to their respective writing systems, a process that would reveal the potency of cultural transmission in shaping basic aspects of human psychology are not readily observed in behavior.

Acknowledgements

The author would like to thank Michael Muthukrishna for help with conceptual development, Gary Lupyan and Annabel Chen for comments on earlier drafts, and Gandalf Li for native Chinese translation and comments.

ORCID iD

Ryutaro Uchiyama  <https://orcid.org/0000-0003-2444-9980>

Ethical Considerations

Ethical approval for this study was granted by the London School of Economics Research Ethics Committee (REC), which has U.S. Department of Health and Human Services IORG (IRB organization) status (IRB00004908, and Federal Wide Assurance FWA00025801). The REC case number for the present study is 19,583. All aspects of the study were conducted in accordance with relevant guidelines and regulations endorsed by the REC.

Consent to Participate

Informed consent was obtained from all respondents in their native languages, through an initial consent statement at the start of the questionnaire that described the purpose of the study, the possibility of their anonymized and aggregated data being published in academic venues, and the right to withdraw from the study at any point in time. Participants were transferred to the main questionnaire only if they had read through these terms and chose to accept them by selecting the acceptance option within the local survey management company's user interface.

Author Contributions

The sole author is responsible for all aspects of the paper.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

Preregistrations and data that support the findings of this study are available at Open Science Framework (OSF) at the URL <<https://osf.io/nxmg2/>> and under OSF project identifier <nxmg2>.

References

Alderson-Day, B., Mitrenga, K., Wilkinson, S., McCarthy-Jones, S., & Fernyhough, C. (2018). The varieties of inner speech questionnaire – Revised (VISQ-R): Replicating and refining links between inner speech and psychopathology. *Consciousness and Cognition*, 65, 48–58. <https://doi.org/10.1016/j.concog.2018.07.001>

Allinson, C. W., & Hayes, J. (1996). The cognitive style index: A measure of intuition-analysis for organizational research. *Journal of Management Studies*, 33(1), 119–135. <https://doi.org/10.1111/j.1467-6486.1996.tb00801.x>

Ansari, D., Donlan, C., Thomas, M. S. C., Ewing, S. A., Peen, T., & Karmiloff-Smith, A. (2003). What makes counting count? Verbal and visuo-spatial contributions to typical and atypical number development. *Journal of Experimental Child Psychology*, 85(1), 50–62. [https://doi.org/10.1016/S0022-0965\(03\)00026-2](https://doi.org/10.1016/S0022-0965(03)00026-2)

Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, 41(5), 319–329. <https://doi.org/10.1016/j.evolhumbehav.2020.07.015>

Baron, H. (1996). Strengths and limitations of ipsative measurement. *Journal of Occupational and Organizational Psychology*, 69(1), 49–56. <https://doi.org/10.1111/j.2044-8325.1996.tb00599.x>

Barrett, H. C. (2020). Deciding what to observe: Thoughts for a post-WEIRD generation. *Evolution and Human Behavior*, 41(5), 445–453. <https://doi.org/10.1016/j.evolhumbehav.2020.05.006>

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., Fitzpatrick, S., Gurven, M., Henrich, J., Kanovsky, M., Kushnick, G., Pisor, A., Scelza, B. A., Stich, S., von Rueden, C., Zhao, W., & Laurence, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, 113(17), 4688–4693. <https://doi.org/10.1073/pnas.1522070113>

Blazhenkova, O., & Kozhevnikov, M. (2010). Visual-object ability: A new dimension of non-verbal intelligence. *Cognition*, 117(3), 276–301. <https://doi.org/10.1016/j.cognition.2010.08.021>

Bruder, J., & Zehra, M. (2025). Aphantasia, hyperphantasia and sensory imagery in a multi-cultural sample. *Journal of Cultural Cognitive Science*, 9(3), 465–481. <https://doi.org/10.1007/s41809-025-00184-8>

Cao, F., Brennan, C., & Booth, J. R. (2014). The brain adapts to orthography with experience: Evidence from English and Chinese. *Developmental Science*, 18(5), 785–798. <https://doi.org/10.1111/desc.12245>

Cao, F., Lee, R., Shu, H., Yang, Y., Xu, G., Li, K., & Booth, J. R. (2010). Cultural constraints on brain development: Evidence from a developmental study of visual word processing in Mandarin Chinese. *Cerebral Cortex*, 20(5), 1223–1233. <https://doi.org/10.1093/cercor/bhp186>

Cao, F., Peng, D., Liu, L., Jin, Z., Fan, N., Deng, Y., & Booth, J. R. (2009). Developmental differences of neurocognitive networks for phonological and semantic processing in Chinese word reading. *Human Brain Mapping*, 30(3), 797–809. <https://doi.org/10.1002/hbm.20546>

Chen, C., Lee, S., & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science*, 6(3), 170–175. <https://doi.org/10.1111/j.1467-9280.1995.tb00327.x>

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, 125(1), 47–63. <https://doi.org/10.1037/0033-2909.125.1.47>

Clark, A. (2025). Extending minds with generative AI. *Nature Communications*, 16(1), 4627. <https://doi.org/10.1038/s41467-025-59906-9>

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(1), 7. <https://doi.org/10.7275/JYJ1-4868>

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, 16(4), 234–244. <https://doi.org/10.1038/nrn3924>

Edelman, S. (2008). *Computing the mind: How the mind really works*. Oxford University Press.

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989x.4.3.272>

Fischer, R. (2004). Standardization to account for cross-cultural response bias: A classification of score adjustment procedures and review of research in JCCP. *Journal of Cross-Cultural Psychology*, 35(3), 263–282. <https://doi.org/10.1177/0022022104264122>

Floridou, G. A., Peerdeman, K. J., & Schaefer, R. S. (2022). Individual differences in mental imagery in different modalities and levels of intentionality. *Memory & Cognition*, 50(1), 29–44. <https://doi.org/10.3758/s13421-021-01209-7>

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Prentice Hall.

Han, S. J., Kelly, P., Winters, J., & Kemp, C. (2022). Simplification is not dominant in the evolution of Chinese characters. *Open Mind*, 6, 264–279. https://doi.org/10.1162/opmi_a_00064

Handel, Z. (2019). *Sinography: The borrowing and adaptation of the Chinese script*. Brill.

Henrich, J., Blasi, D. E., Curtin, C. M., Davis, H. E., Hong, Z., Kelly, D., & Kroupin, I. (2023). A cultural species and its cognitive phenotypes: Implications for philosophy. *Review of Philosophy and Psychology*, 14(2), 349–386. <https://doi.org/10.1007/s13164-021-00612-y>

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., & Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *The Behavioral and brain sciences*, 28(6), 795–815. <https://doi.org/10.1017/S0140525X05000142>

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and brain sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>

House, B. R., Kanngiesser, P., Barrett, H. C., Broesch, T., Cebioglu, S., Crittenden, A. N., Erut, A., Lew-Levy, S., Sebastian-Enesco, C., Smith, A. M., Yilmaz, S., & Silk, J. B. (2020). Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behaviour*, 4(1), 36–44. <https://doi.org/10.1038/s41562-019-0734-z>

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>

Huang, S., Zhou, Y., Du, M., Wang, R., & Cai, Z. G. (2021). Character amnesia in Chinese handwriting: A mega-study analysis. *Language Sciences*, 85, 101383. <https://doi.org/10.1016/j.langsci.2021.101383>

Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Brace, N., Duuren, M., Pike, G., Grice, S., & Campbell, R. (2004). Exploring the Williams syndrome face-processing debate: The importance of building developmental trajectories. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 45(7), 1258–1274. <https://doi.org/10.1111/j.1469-7610.2004.00322.x>

Kelly, P., Winters, J., Miton, H., & Morin, O. (2021). The predictable evolution of letter shapes: An emergent script of West Africa recapitulates historical change in writing systems. *Current Anthropology*, 62(6), 669–691. <https://doi.org/10.31235/osf.io/eg489>

Kemps, E., & Newson, R. (2005). Patterns and predictors of adult age differences in mental imagery. *Aging, Neuropsychology, and Cognition*, 12(1), 99–128. <https://doi.org/10.1080/13825580590925152>

Kirby, J. R., Moore, P. J., & Schofield, N. J. (1988). Verbal and visual learning styles. *Contemporary Educational Psychology*, 13(2), 169–184. [https://doi.org/10.1016/0361-476X\(88\)90017-3](https://doi.org/10.1016/0361-476X(88)90017-3)

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at new look. *Psychological Science*, 14(3), 201–206. <https://doi.org/10.1111/1467-9280.02432>

Kitayama, S., Ishii, K., Imada, T., Takemura, K., & Ramaswamy, J. (2006). Voluntary settlement and the spirit of independence: Evidence from Japan's "northern frontier.". *Journal of Personality and Social Psychology*, 91(3), 369–384. <https://doi.org/10.1037/0022-3514.91.3.369>

Kitayama, S., Park, H., Sevincer, A. T., Karasawa, M., & Uskul, A. K. (2009). A cultural task analysis of implicit independence: Comparing North America, Western Europe, and East Asia. *Journal of Personality and Social Psychology*, 97(2), 236–255. <https://doi.org/10.1037/a0015999>

Kosslyn, S. M., Cacioppo, J. T., Davidson, R. J., Hugdahl, K., Lovallo, W. R., Spiegel, D., & Rose, R. (2002). Bridging psychology and biology: The analysis of individuals in groups. *American Psychologist*, 57(5), 341–351. <https://doi.org/10.1037/0003-066X.57.5.341>

Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework of cognitive style. *Psychological Bulletin*, 133(3), 464–481. <https://doi.org/10.1037/0033-2909.133.3.464>

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition*, 33(4), 710–726. <https://doi.org/10.3758/BF03195337>

Kraemer, D. J. M., Hamilton, R. H., Messing, S. B., DeSantis, J. H., & Thompson-Schill, S. L. (2014). Cognitive style, cortical stimulation, and the conversion hypothesis. *Frontiers in Human Neuroscience*, 8, Article 15. <https://doi.org/10.3389/fnhum.2014.00015>

Kraemer, D. J. M., Rosenberg, L. M., & Thompson-Schill, S. L. (2009). The neural correlates of visual and verbal cognitive styles. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(12), 3792–3798. <https://doi.org/10.1523/JNEUROSCI.4635-08.2009>

Kroupin, I., Davis, H. E., & Henrich, J. (2025). Beyond Newton: Why assumptions of universality are critical to cognitive science, and how to finally move past them. *Psychological Review*, 132(2), 291–310. <https://doi.org/10.1037/rev0000480>

Leibniz, G. W. (1697). *Novissima sinica*. Joachim Georgii.

Levine, S. C., Foley, A., Lourenco, S., Ehrlich, S., & Ratliff, K. (2016). Sex differences in spatial cognition: Advancing the conversation: Sex differences in spatial cognition. *Wiley Interdisciplinary Reviews. Cognitive Science*, 7(2), 127–155. <https://doi.org/10.1002/wcs.1380>

Lupyan, G., Uchiyama, R., Thompson, B., & Casasanto, D. (2023). Hidden differences in phenomenal experience. *Cognitive Science*, 47(1), Article e13239. <https://doi.org/10.1111/cogs.13239>

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8(3), 108–114. <https://doi.org/10.1016/j.tics.2004.01.003>

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253. <https://doi.org/10.1037/0033-295x.98.2.224>

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. <https://doi.org/10.1037/0022-3514.81.5.922>

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, 3(1), 97–110. <https://doi.org/10.21500/20112084.854>

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology*, 95(4), 833–846. <https://doi.org/10.1037/0022-0663.95.4.833>

McBride, C. A. (2016). Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educational Psychology Review*, 28(3), 523–549. <https://doi.org/10.1007/s10648-015-9318-2>

McCarthy-Jones, S., & Fernyhough, C. (2011). The varieties of inner speech: Links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness and Cognition*, 20(4), 1586–1593. <https://doi.org/10.1016/j.concog.2011.08.005>

McLuhan, M. (1962). *The Gutenberg galaxy: The making of typographic man*. University of Toronto Press.

Morin, O., & Koshevoy, A. (2024). A cultural evolutionary model for the law of abbreviation. *Topics in Cognitive Science*, tops.12782. <https://doi.org/10.1111/tops.12782>

Na, J., Grossmann, I., Varnum, M. E. W., Kitayama, S., Gonzalez, R., & Nisbett, R. E. (2010). Cultural differences are not always reducible to individual differences. *Proceedings of the National Academy of Sciences of the United States of America*, 107(14), 6192–6197. <https://doi.org/10.1073/pnas.1001911107>

Noppeney, U., Penny, W. D., Price, C. J., Flandin, G., & Friston, K. J. (2006). Identification of degenerate neuronal systems based on intersubject variability. *NeuroImage*, 30(3), 885–890. <https://doi.org/10.1016/j.neuroimage.2005.10.010>

Noppeney, U., Price, C. J., Penny, W. D., & Friston, K. J. (2006). Two distinct neural mechanisms for category-selective responses. *Cerebral Cortex*, 16(3), 437–445. <https://doi.org/10.1093/cercor/bhi123>

Oakley, B., Johnston, M., Chen, K., Jung, E., & Sejnowski, T. (2025). The memory paradox: Why our brains need knowledge in an age of AI. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5250447>

Perfetti, C., Cao, F., & Booth, J. (2013). Specialization and universals in the development of reading skill: How Chinese research informs a universal science of reading. *Scientific Studies of Reading: The Official Journal of the Society for the*

Scientific Study of Reading, 17(1), 5–21. <https://doi.org/10.1080/10888438.2012.689786>

Raîche, G., Walls, T. A., Magis, D., Riopel, M., & Blais, J.-G. (2013). Non-graphical solutions for cattell's scree test. *Methodology*, 9(1), 23–29. <https://doi.org/10.1027/1614-2241/a000051>

Richerson, P. J., Boyd, R., & Henrich, J. (2010). Gene-culture coevolution in the age of genomics. *Proceedings of the National Academy of Sciences of the United States of America*, 107(Supplement_2), 8985–8992. <https://doi.org/10.1073/pnas.0914631107>

Roebuck, H., & Lupyan, G. (2020). The internal representations questionnaire: Measuring modes of thinking. *Behavior Research Methods*, 52(5), 2053–2070. <https://doi.org/10.3758/s13428-020-01354-y>

Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, 24(2), 282–292. <https://doi.org/10.1037/a0025697>

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, 74(1), 31–57. <https://doi.org/10.1177/0013164413498257>

Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, 366(6466), eaau5141. <https://doi.org/10.1126/science.aau5141>

Senzaki, S., Masuda, T., Takada, A., & Okada, H. (2016). The communication of culturally dominant modes of attention from parents to children: A comparison of Canadian and Japanese parent-child conversations during a joint scene description task. *PLoS One*, 11(1), Article e0147199. <https://doi.org/10.1371/journal.pone.0147199>

Seymour, P. H. K., Aro, M., Erskine, J. M., & collaboration with COST Action A8 network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174. <https://doi.org/10.1348/00072603321661859>

Talhelm, T., & English, A. S. (2020). *Historically rice-farming societies have tighter social norms in China and worldwide* (p. 201909909). Proceedings of the National Academy of Sciences. <https://doi.org/10.1073/pnas.1909909117>

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. <https://doi.org/10.1126/science.1246850>

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping*, 25(1), 83–91. <https://doi.org/10.1002/hbm.20134>

Tasaki, K., & Shin, J. (2017). Nihonjin no kaitou bias: Response style no shubetsukan/bunkakan hikaku [Japanese response bias: Cross-level and cross-national

comparisons on response styles]. *Shinrigaku Kenkyu: Japanese Journal of Psychology*, 88(1), 32–42. <https://doi.org/10.4992/jjpsy.88.15065>

Wang, Q. (2021). The cultural foundation of human memory. *Annual Review of Psychology*, 72(1), 151–179. <https://doi.org/10.1146/annurev-psych-070920-023638>

Wei, L., English, A. S., Talhelm, T., Zhang, Y., Tan, X., Zhu, J., & Wang, J. (2025). People in relationally mobile cultures report higher well-being. *Emotion*, 25(3), 541–555. <https://doi.org/10.1037/emo0001439>

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *ETS Research Bulletin Series*, 47(1), 1–64. <https://doi.org/10.1002/j.2333-8504.1975.tb01065.x>

Wu, C.-Y., Ho, M.-H. R., & Chen, S.-H. A. (2012). A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. *NeuroImage*, 63(1), 381–391. <https://doi.org/10.1016/j.neuroimage.2012.06.047>

Xu, Y., Pollatsek, A., & Potter, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 838–857. <https://doi.org/10.1037//0278-7393.25.4.838>

Yanaoka, K., Foster, R., Michaelson, L. E., Saito, S., & Munakata, Y. (2024). The power of cultural habits: The role of effortless control in delaying gratification. *Current Opinion in Psychology*, 60, 101903. <https://doi.org/10.1016/j.copsyc.2024.101903>

Zhang, Z., Van Lieshout, L. L. F., Colizoli, O., Li, H., Yang, T., Liu, C., Qin, S., & Bekkering, H. (2024). A cross-cultural comparison of intrinsic and extrinsic motivational drives for learning. *Cognitive, Affective, & Behavioral Neuroscience*, 25(1), 25–44. <https://doi.org/10.3758/s13415-024-01228-2>

Author Biography

Ryutaro Uchiyama is Assistant Professor of Social and Cognitive Sciences at the Singapore University of Technology and Design, and Principal Investigator of the Human Cognitive Ecology Lab.