To Know or Not to Know? Validating a Chinese Measure of Information Preferences

Min-Heng Wang[*] Greg Chih-Hsin Sheen[*†] Emily H. Ho[‡]
David Hagmann[§] Po-Yi Chen[¶] George Loewenstein[‖]

July 16, 2026

Abstract

The Information Preferences Scale (IPS) measures individuals' tendency to seek or avoid potentially useful but unpleasant information. Although the IPS has been validated in English-speaking Western samples, its use in Mandarin-speaking populations has been limited by the absence of a validated Chinese version. We translated the IPS into Traditional Chinese and evaluated its psychometric properties in Mandarin-speaking samples from Taiwan and Hong Kong (N = 1,465). Across studies, the Chinese IPS showed acceptable internal consistency, replicated the theorized multidimensional structure of the original scale, and demonstrated measurement invariance across cultures and languages. In a bilingual Hong Kong sample, IPS scores were moderately stable across English and Chinese administrations, and both versions predicted participants'willingness to learn potentially unflattering peer feedback. We further illustrate the scale's utility by showing domain-specific cultural differences in information preferences. By providing a validated Chinese version of the IPS and evidence of cross-language and cross-cultural comparability, this research enables future studies to examine information seeking and avoidance in Mandarin-speaking populations and to make more defensible cross-cultural comparisons of information preferences.

Keywords: information avoidance, information preferences, cross-cultural comparison, measurement invariance, judgment and decision-making

[*]Department of Marketing Management, The Rotterdam School of Management, Erasmus University.
[†]Institute of Political Science, Academia Sinica, Taiwan. Corresponding author. 128 Academia Road, Sec. 2, Nankang, 115 Taipei, Taiwan. E-mail: gregcsheen@gate.sinica.edu.tw.
[‡]Department of Medical Social Sciences, Northwestern University, USA.
[§]Department of Management, The Hong Kong University of Science and Technology, Hong Kong.
[¶]Department of Educational Psychology and Counseling, National Taiwan Normal University, Taiwan.
[‖]Department of Social and Decision Sciences, Carnegie Mellon University, USA.

1 Introduction

People sometimes choose to avoid information—even if it is free and could be potentially useful for decision making—when they worry that it might convey negative news (Foust and Taber, 2023). This avoidance of information is widespread and has a considerable impact on the quality of human judgment and decision-making across diverse domains (Golman et al., 2017). In the realm of health, for example, individuals sometimes avoid medical tests due to fear of receiving a positive diagnosis (Brashers et al., 2002; Ganguly and Tasoff, 2017; Oster et al., 2013). In business, acquiring companies sometimes fail to engage in sufficient due-diligence about the company they are acquiring, often when the acquisition has an internal 'momentum'—as may have been the case in Hewlett-Packard's disastrous 2011 acquisition of the software company Autonomy (Fischer, 2013). In the context of politics, individuals selectively expose themselves to news or information sources that conform to their ideology and political preferences, which can exacerbate polarization (Arceneaux et al., 2012; Druckman et al., 2013; Knobloch-Westerwick and Lavis, 2017; Knobloch-Westerwick et al., 2015; Pennycook and Rand, 2021; Sunstein, 2001). Information avoidance, therefore, poses significant challenges to public health, business, public policy, and political life.

Reflecting the dominant situationalist perspective according to which behavior is seen as guided mainly by situational factors, the phenomenon of information avoidance has primarily been studied through experimental manipulations conducted by behavioral economists or psychologists in both controlled laboratory settings and real-world field environments (Golman et al., 2022; Li et al., 2021). However, casual observation, as well as empirical research, suggest that there are meaningful individual differences in people's propensities to seek out or avoid potentially threatening but useful information. Ho et al. (2021) have, specifically, developed a psychometrically validated instrument, the Information Preferences Scale (IPS), that allows for the direct and robust assessment of individuals' information preferences across multiple domains (Ho et al., 2021). The IPS is a 13-item instrument developed to measure an individual's tendency to seek or avoid information that may be unpleasant to obtain

but could help them make informed decisions in the future (Ho et al., 2021). Presenting respondents with a series of scenarios involving health, finance, and social interactions, the instrument assesses preferences for obtaining or avoiding information in a series of scenarios with a 4-point Likert scale (1 = *Definitely don't want to know* to 4 = *Definitely want to know*).

The ability to measure information preferences opens up new avenues for research in psychology, behavioral economics, political science and other related social sciences, enabling a deeper understanding of human behavior in information-seeking or -avoidance situations. Because it was created in English, however, the IPS has primarily been utilized in European, North American, or United Kingdom contexts (Ash et al., 2023; Ho et al., 2021; Liew et al., 2023) where English is commonly spoken. Little work has been done in East Asia, especially among Mandarin-speaking people, despite the fact that Mandarin is, globally, the most commonly spoken language.

To address this gap, we translated the English version of IPS into Chinese. We followed best practice in the translation process including hiring professional translators to perform back translation (Brislin, 1970; Klotz et al., 2023), which meant that the items translated from English to Chinese were re-translated back to English to ensure that there was no loss of semantic or cultural meaning (Appendix A presents both the original and translated versions of the IPS).

We then conducted data collection in Taiwan (Study 1) and subsequently in Hong Kong (Study 2), both places that have Chinese/Mandarin as their official or primary language. Here, we present the results of analyses of these data sets that validate the Chinese version of the Information Preference Scale (IPS), map East-West cross-cultural variations in information preferences, and assess the translated scale's ability to predict real-world information acquisition decisions.

The aim of this research is threefold. First, this study validates the psychometric properties of the Chinese version of the Information Preference Scale (IPS). Using multi-group con-

firmatory factor analysis (MG-CFA) (Kline, 2023), we test for measurement invariance (MI) across U.S., Taiwan, and Hong Kong samples. Establishing MI ensures that the construct is interpreted consistently across cultures and languages, allowing for valid cross-cultural comparisons of IPS scores (Alacreu-Crespo et al., 2019; Welindt et al., 2023; Zhang et al., 2019). The detailed MG-CFA procedures and results are all presented in the Appendix.

Second, after establishing MI across groups/cultures, we test the predictive validity of the translated scale. Study 2 does so by testing whether both the English and Chinese versions of the IPS (conducted on the bilingual sample) predict real-world information-seeking decisions, and specifically whether it captures respondents' likelihood of seeking or avoiding feedback on whether one is perceived as likable or competent by peers.

The third and final goal of the research is to demonstrate the usefulness of the translated scale by investigating differences in patterns of information preferences between American and East Asian respondents after controlling for potential different scale use. Doing so, the analysis reveals that Taiwanese decision-makers display greater information avoidance than Americans within interpersonal contexts, but stronger information-seeking tendencies when it comes to financial affairs.

The paper concludes by discussing the implications of these findings for human judgment, decision-making, and behavioral economics.

2 Study 1

Study 1 was designed to validate the Chinese version of the IPS, and to examine differences in information preferences between American and Taiwanese samples. It was conducted by Rakuten Insight, an international survey firm, which administered the survey in Taiwan.[1] Participants received compensation for completing the survey. The study was approved by the Institutional Review Board of multiple authors' institutions.

[1]We chose Rakuten Insight because the stratified sampling executed by the survey firm allowed the sample to be approximately nationally representative.

2.1 Participants

The sample, which was recruited during January and February 2024, included the 1,198 (50.4% male) adults who passed the attention check (685 participants failed). The age-group distribution was: 20-29 years (17.0%), 30-39 years (18.5%), 40-49 years (21.5%), 50-59 years (19.6%), and 60 years or older (23.4%). See Appendices B and C for a detailed demographic breakdown. The USA samples collected by Ho et al. (2021) via Amazon Mechanical Turk included 52.89% male participants, with a mean age of 34.93 (SD = 9.99), after excluding 20 participants (out of 400) who either failed the attention check or did not complete the survey.[2]

2.2 Statistical Analysis

Since the IPS was being used for the first time with East Asian (Taiwanese) participants, we calculated Cronbach's alphas for the entire scale and each subscale. To assess factorial validity, the proposed four-factor structure (see Figure 1) was also tested through confirmatory factor analysis (CFA); we report the goodness-of-fit indices. Furthermore, to assess the relationship between information preferences (IP) and other pertinent constructs hypothesized to be related to this measure, we examine the correlation of the IPS to a widely-used risk attitude scale: the Domain-Specific Risk-Taking (DOSPERT) scale (Weber et al., 2002). Recent empirical evidence suggests that risk preferences/attitudes apply to our desire for information (Kobayashi et al., 2019; Moon et al., 2022; Somerville et al., 2017); we are taking a risk when we expose ourselves to the possibility of receiving negative news (Kőszegi and Rabin, 2006).

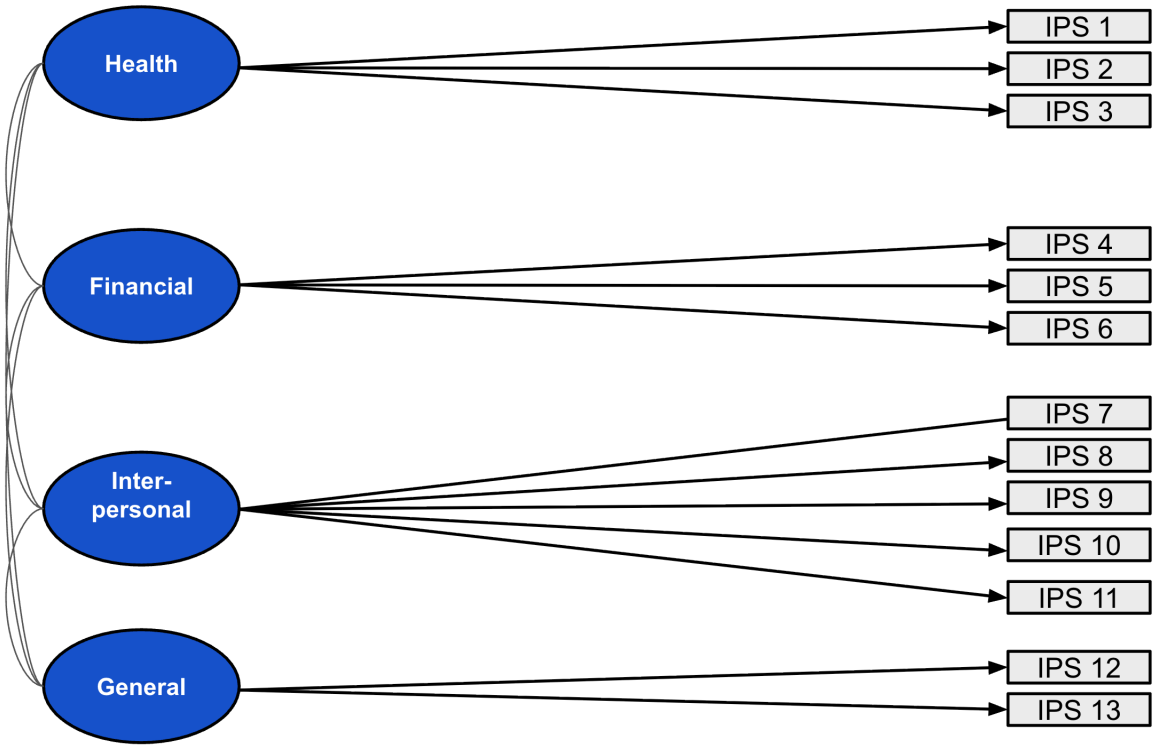
Simply put, before we could use the IPS to compare how Taiwanese and American people feel about information, we had to make sure of one critical baseline: if the IPS serves as a "ruler" measuring people's information preferences, we need to guarantee that we are using

[2]All the data and codes used in this paper, including both Studies 1 and 2, are available at the following OSF webpage `https://osf.io/j7h3r/`. Given the nature of the study is to corroborate the validity of the IPS scale, the studies were not preregistered.

the exact same "ruler" for both groups. Imagine if, when looking at the same scale item, a Taiwanese respondent thinks "Agree" means an 80-out-of-100, while an American respondent thinks it only means a 60. If their interpretation and scoring baselines differ, directly averaging and comparing their raw scores would be misleading. In statistics, this baseline check is called Measurement Invariance (MI). Because information preference is multidimensional rather than a single trait (Ho et al., 2021), we used a step-by-step statistical model called Multi-Group CFA. This allowed us to systematically test whether both cultural groups shared the same understanding and scoring standards. Methodologists note that to fairly compare group averages, we must at least achieve scalar invariance (Millsap and Olivera-Aguilar, 2012; Wang and Wang, 2019), indicating both the way participants understand the words (factor loadings) and their starting points for giving scores (intercepts) are identical. Additionally, since our scale uses a 4-point scale, we applied a specialized statistical estimator (ULSMV) that is uniquely tailored to handle this type of categorical data, ensuring our estimates remain highly accurate and stable (Forero et al., 2009).

Once we mathematically proved that our "ruler" was calibrated perfectly across both cultures/groups, we could confidently examine the actual differences (Cheung and Rensvold, 2002). To do this, we set the Taiwanese group as our baseline and allowed the American group's average to be freely estimated. This approach lets us see clearly, without statistical bias, whether American or Taiwanese respondents have a stronger true preference for seeking out specific information.

Figure 1: Multi-dimensional CFA Model of the Information Preferences Scale



Note. Ellipses denote the four latent IPS dimensions (Health, Financial, Interpersonal, and General); rectangles denote the observed scale items (IPS 1–IPS 13) loading onto their respective dimensions. Curved double-headed arrows represent the estimated correlations among the latent factors.

2.3 Results

2.3.1 Descriptive Statistics and Psychometric Properties of Translated IPS

In Table 1, the Cronbach's alphas for the IPS scale and its subscales are reported. The column marked TW reports the results for the Taiwanese sample of Study 1. From this column it can be seen that all four subscales demonstrate at least acceptable or good internal consistency, ranging from $\alpha = .61$ to $\alpha = .80$, in line with estimates from the original validation study (presented in the column marked USA). The entire scale's internal consistency ($\alpha = .86$) is

even more favorable than that of any of the subscales, and is identical to that displayed by the U.S. sample.

Table 1: Cronbach's α of the entire IPS and each subscale for the East Asian and USA Samples

Scale/Subscale	Number of Items	Item Number	TW	HK (CH)	HK (EN)	USA
IPS (Total)	13		.86	.76	.79	.86
IPS (Health)	3	1, 2, 3	.78	.62	.66	.79
IPS (Financial)	3	4, 5, 6	.61	.52	.62	.61
IPS (Interpersonal)	5	7, 8, 9, 10, 11	.80	.62	.69	.76
IPS (General)	2	12, 13	.64	.47	.33	.79

We also examined the factorial structure of the IPS among Taiwanese respondents; the hypothesized four-factor model showed good fit to the data.[3] Furthermore, consistent with the idea that sharing of information is in part indicative of risk taking, we found a small but highly significant positive correlation between IPS and DOSPERT scores for both risk attitudes ($r = 0.12,\ p < 0.001$) and risk perception ($r = 0.16,\ p < 0.001$). Individuals demonstrating a greater proclivity for information seeking also show a greater propensity to take (and perceive) risks.

2.3.2 Comparison of Taiwanese and American Scale Responses

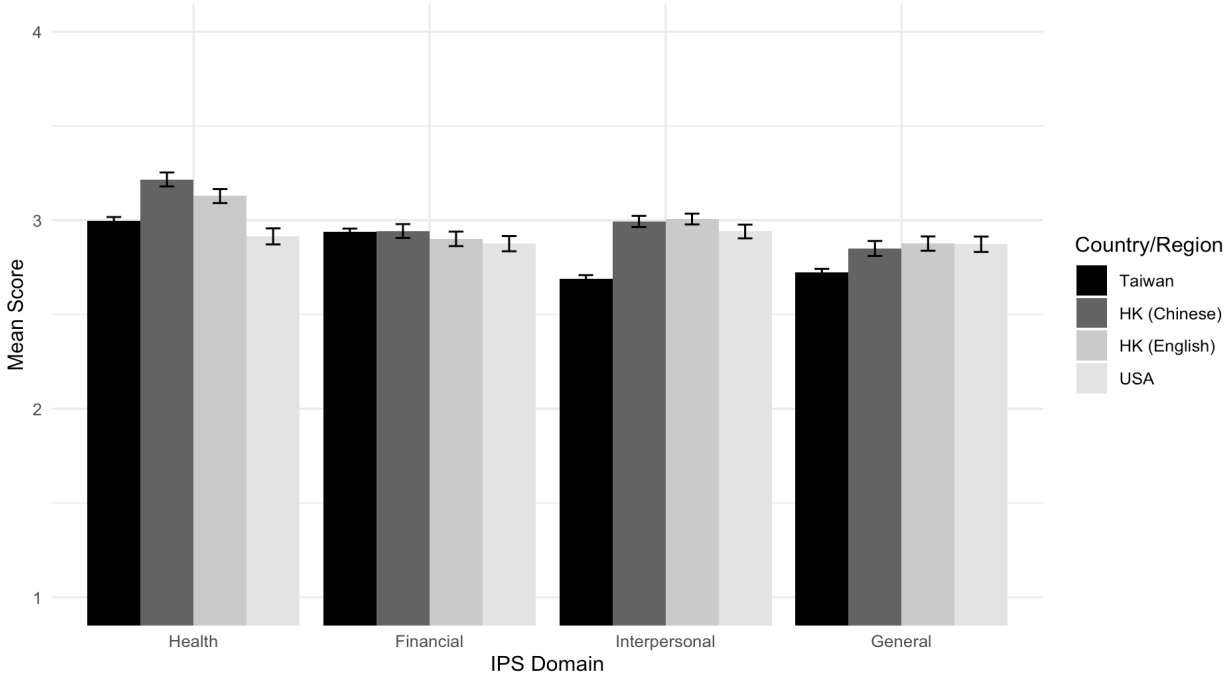
If measurement invariance is not maintained, directly comparing observed scores (as in Figure 2) across groups/cultures may lead to bias, as the scale use may vary across groups/cultures. This may result in unreliable conclusions. In such a condition, comparing latent means, rather than observed means, provides a more accurate and interpretable approach, ensuring that any observed variances reflect genuine group differences rather than measurement artifacts. Since measurement invariance has been established (see Appendix F), we can

[3]CFI = .94, TLI = .92, RMSEA = .08, SRMR = 0.05. The fit indices indicate that the factor structure is appropriate. All factor loadings exceed .55, and their directions align with expectations

now meaningfully compare differences in information preference dimensions between the Taiwanese and American samples.

The chi-square difference testing results indicated that three out of four domains exhibited statistically significant differences: *financial* ($p < 0.01$), *interpersonal* ($p < 0.01$), and *general* ($p < 0.01$). In the interpersonal and general domains, American participants demonstrated a higher preference for information compared to their East Asian (Taiwanese) counterparts. In contrast, in the financial domain, the pattern was reversed, with Taiwanese participants displaying a stronger preference for financial-related information than Americans. No significant differences were found in the *health* domain ($p > .05$). Since we also collected a sample from Hong Kong in Study 2, the comparison between Taiwanese and Hong Kong participants will be discussed in a later section. These findings illustrate the utility of having English and Mandarin versions of the scale that satisfy measurement invariance.

Figure 2: Measured Information Preference by Country/Region



Note. Bars represent mean observed IPS scores on the original 1–4 response metric, plotted separately by domain (Health, Financial, Interpersonal, and General) for each country/region. Error bars denote ±1 standard error of the mean.

3 Study 2

In Study 2 we directly compare the English and Chinese versions of the IPS using a bilingual sample from Hong Kong. Participants completed the Chinese[4] and English versions of the IPS in sessions separated by approximately a week, and, as a test of predictive validity, also made a consequential decision about whether to acquire potentially uncomfortable interpersonal/social information: how their peers perceive them in terms of likability, competence, or supportiveness. We also investigated the measurement invariance (MI) between individuals

[4]At the beginning of the survey, participants were asked whether they preferred to read the Traditional Chinese or Simplified Chinese version and were then directed to the corresponding interface. The two versions differed only in the Chinese characters used; all other content and procedures were identical.

in Hong Kong and Taiwan.

3.1 Methods

The experiment was conducted in two stages as part of a research participation requirement for undergraduate students at a business school in Hong Kong. All students were required to be proficient in English as a condition of admission, and all courses (excluding Chinese language courses) were taught in English.

We sent email invitations to 446 students across eight course sections. Of these, 435 completed the first stage of the study, and, of those completing the first stage, 374 also completed the second stage. The sample demographics were as follows: 50.3% identified as female, with an average age of 19.55 year (SD = 1.40).

In Stage 1, participants were asked whether they could read Chinese. Those who answered affirmatively (n = 267) were randomly assigned to first complete the English version of the IPS, then the Chinese version, or to complete both in the opposite order. Participants unable to read Chinese (n = 107) were assigned to an "English Only" condition, in which they completed the IPS in English twice, enabling us to measure test-retest reliability using the exact same (English version) of the scale.

After completing the IPS, participants engaged in a peer nomination task designed to set the stage for the real information-related decision they would be given. Each participant was presented with a list of students from their course section and asked to nominate three peers they considered most likable, most competent, and most supportive, for a total of nine unique peers. Participants could only nominate a peer for one category, and nominations were made by dragging names into designated boxes on the interface.[5] Stage 1 of the survey concluded with standard demographic questions.

Approximately a week after completing Stage 1, the same participants participated in

[5]Due to a coding error, one section required students to select six peers per category. However, as our analysis relies on relative rankings, this discrepancy does not affect the subsequent data, and responses from this section are included in all analyses.

Stage 2. Bilingual participants who had completed the IPS in English or Chinese during Stage 1 were now asked to complete the scale in the other language. Those who were not bilingual again completed the scale in English. Following this, participants were asked to indicate which of the three social dimensions—likability, competence, or supportiveness—was most important to them. Based on their response, participants were reminded of the peer nomination task from Stage 1 and were asked *if they wanted to know whether they ranked in the top 20% of their peers in the dimension they thought was most important.* This information-seeking decision served as our behavioral outcome variable.[6] Participants who opted to receive the information were then truthfully informed whether they ranked in this top group. The study concluded with a brief debriefing, explaining the purpose of the study.

Moreover, to assess MI between the Hong Kong and Taiwan samples, we used the Taiwanese sample from Study 1 (who completed the Chinese IPS) and the Hong Kong sample from Study 2 (Chinese-reading individuals). The analysis followed a process similar to that used in Study 1.

3.2 Results

3.2.1 Correlation and Reliability analysis

The correlation matrices for Study 2 are presented in Table 2 (all measured by the Chinese version of IPS) and Table 3 (all measured by the English version of IPS). Regardless of whether the IPS was administered in Chinese or English, the full IPS score was significantly and positively correlated with all subscales (which is not surprising, since they comprised the full score), and the subscales were also significantly and positively correlated with one another, though these correlations are sufficiently low to be consistent with the idea that they are measuring at least partly domain-specific traits. Across both language versions, the association between the full IPS value and the General, two-item, subscale was some-

[6]The top 20% rankings were determined by counting how many nominations each participant received in each category and selecting the top 20% with the most votes.

what lower than that of the three other subscales; but this is likely to reflect that they contributed only two items to the full scale, whereas the Health and Financial subscales both contributed 3, and the Interpersonal subscale contributed 5 (in fact, the correlations were higher in all cases for subscales that contributed more items to the total scale). When the intercorrelations among the IPS subscales reported in Tables 2 and 3 are compared with those reported in the original IPS validation study (Ho et al., 2021), the correlations obtained using the Chinese-language version (Table 2) are generally somewhat lower. In contrast, the correlations obtained using the English-language version (Table 3) are more similar to those reported in the original study, with some even being higher, such as the correlation between the health and interpersonal subscales. These comparison should, however, be interpreted cautiously because the original study reported latent factor correlations, whereas Tables 2 and 3 present Pearson correlations.

Table 2: Pearson Correlation Matrix for the Scales Measured in Chinese

Variable	1	2	3	4	5
1. Full IPS	-				
2. Health	.71***	-			
3. Financial	.70***	.30***	-		
4. Interpersonal	.80***	.38***	.41***	-	
5. General	.58***	.35***	.24***	.28***	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

Table 3: Pearson Correlation Matrix for the Scales Measured in English

Variable	1	2	3	4	5
1. Full IPS	-				
2. Health	.79***	-			
3. Financial	.74***	.44***	-		
4. Interpersonal	.85***	.51***	.52***	-	
5. General	.58***	.42***	.20***	.36***	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

Table 4 presents test-retest reliability analyses for both the English-English sample (comprising respondents who did not speak Chinese) and the English-Chinese sample (comprising bilingual respondents) for the full IPS and each of the four subscales. These reveal all-significant moderate to strong positive correlations between Time 1 and Time 2 for all measures and both samples, providing evidence of test–retest reliability. For participants who completed the English version at both time points, the correlations ranged from .60 to .75, the highest of which was, not surprisingly, for the full IPS. The test-retest correlations are, perhaps unsurprisingly, somewhat lower for participants who completed the scale in different languages, with correlations ranged from .42 to .68, and again the strongest for the full scale. Thus, the overall and domain-specific information preference scores were relatively stable even when the language of administration changed. Overall, these findings suggest that the IPS demonstrates satisfactory temporal stability and that most of its test–retest reliability is retained across the English and Chinese versions.

Table 4: Test-retest Reliability for English-English and Bilingual Versions

Test–retest comparison	Full IPS	Health	Financial	Interpersonal	General
English–English	.75***	.70***	.60***	.66***	.63***
English–Chinese	.68***	.57***	.58***	.62***	.42***

Note. Entries are Pearson correlations between Time 1 and Time 2 scores. The English–Chinese row combines participants assigned to the English-first and Chinese-first conditions. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

3.2.2 Measurement invariance

Measurement invariance (MI) was sequentially confirmed for the Hong Kong and Taiwan samples (see Appendix I), justifying a direct comparison of observed means. Results indicated that the Hong Kong sample exhibited a significantly stronger information-seeking propensity in the health, interpersonal, and general domains (all $p < .001$); the difference in the financial domain was not significant ($p = .242$).

3.2.3 Predictive validity

The predictive validity of Chinese and English versions of IPS are shown in Table 5. The full IPS scale predicts actual information-seeking behavior regardless of the language version of the scale. Meanwhile, the interpersonal subscale, regardless of language version or whether other domains of information preference are controlled for, still shows a positive predictive relationship with the dependent variable.

Table 5: Predictive Validity of the Chinese (CH) and English (EN) IPS Versions for Actual Information Seeking.

Outcome variable = Actual information seeking				
Predictor	Correlation		Multiple Regression	
	CH	EN	CH	EN
Full IPS	.27***	.21**		
Health	.15*	.06	.01 (.04)	-.07 (.04)
Financial	.23***	.19**	.09* (.04)	.07 (.04)
Interpersonal	.23***	.24***	.11* (.05)	.18** (.06)
General	.14*	.08	.03 (.03)	.02 (.04)
R-squared			.08	.08

Note. Starting from the second column, the first two columns show correlation coefficients (r), and the last two show partial regression coefficients (β) from the multiple regression analysis.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

4 Discussion

The Information Preferences Scale (IPS) assesses individual tendencies to seek or avoid information even when it is costless and potentially beneficial for future decision-making. In this research, we translated the IPS from English into Chinese and evaluated the reliability, factorial validity, measurement invariance, test–retest reliability, and predictive validity of the translated instrument. Overall, the findings provide initial evidence that the Chinese version of the IPS can be used to assess information preferences among Chinese-speaking respondents and can be meaningfully compared with the original English version under appropriate measurement conditions.

A central contribution of this research is the validation of the Chinese version of IPS. By testing measurement invariance between the original English version and the translated Chinese version, we examined whether the scale captures a comparable underlying construct across languages and cultural contexts. Establishing measurement invariance does not eliminate all concerns about cultural interpretation or response processes, but it provides an important psychometric foundation for comparing scores across cultures or groups (Jessee, 2021; Reise et al., 1993). Within this framework, our comparisons between U.S. and Taiwanese respondents suggest that information preferences may vary by domain. Taiwanese respondents showed lower information-seeking tendencies in the interpersonal and general domains, but higher information-seeking tendencies in the financial domain, while no significant difference was observed in the health domain. These patterns should be interpreted cautiously, however, as suggestive evidence rather than definitive conclusions about broad cultural differences.

One possible interpretation is that some domains of information preference may be more culturally sensitive than others. For example, lower interpersonal information seeking among Taiwanese respondents may be related to concerns about social evaluation, face, or relational harmony, which are often discussed in relation to interdependent self-construals (Markus and Kitayama, 1991). However, this explanation remains speculative. The observed differences could also reflect differences in sample composition, contextual salience of specific scenarios, response styles not fully captured by the measurement model, or other unmeasured cultural and institutional factors. Future research should therefore directly test these proposed mechanisms rather than treating the present cross-cultural differences as conclusive evidence of culture-specific information preferences.

Study 2 provided additional support for the validity of the IPS through a within-subject, bilingual design. Scores on the English and Chinese versions were moderately stable across administrations, and both versions predicted participants' willingness to learn potentially unflattering peer feedback. These results support the cross-language reliability and predic-

tive utility of the scale. At the same time, the predictive-validity evidence was based on a single behavioral outcome involving interpersonal feedback among Hong Kong university students. Thus, although the behavioral outcome was consequential, it remains unclear whether the IPS predicts information-seeking decisions equally well in health, financial, marketing, political, or other high-stakes contexts.

The finding that the financial subscale showed some association with interpersonal information-seeking behavior is also worth interpreting cautiously. One possibility is that financial information-preference items capture a broader willingness to confront potentially undesirable but useful information. Another possibility is that the result reflects features of the specific behavioral task or sample. Because this cross-domain association was not equally strong across language versions and was observed using only one interpersonal behavioral outcome, it should be viewed as preliminary. Future studies should examine whether domain-specific IPS subscales predict behavior primarily within their corresponding domains or whether some domains capture more general information-seeking tendencies.

Several limitations should be noted. First, the approximately one-week interval between administrations in Study 2 provides evidence of short-term test–retest reliability but does not establish the long-term stability of information preferences. Repeated exposure to similar items may also have influenced responses across the two sessions. Future research should use longer intervals and additional waves of data collection to examine the temporal stability of IPS scores. Second, the Hong Kong sample consisted of university students, which limits generalizability. Future work should validate the scale in broader and more diverse Mandarin-speaking populations, including community samples, older adults, and respondents from different educational and socioeconomic backgrounds.

Third, although measurement invariance analyses provide important support for cross-cultural comparison, they do not by themselves prove that all items are interpreted identically across groups. Some items may still carry different cultural meanings, especially those involving interpersonal evaluation, social feedback, or counterfactual financial out-

comes. Future research could combine quantitative measurement-invariance testing with qualitative methods, such as cognitive interviewing, to better understand how respondents from different cultural backgrounds interpret IPS items. Such work would be especially useful for identifying whether certain scenarios require cultural adaptation rather than direct translation.

Fourth, the present research focused primarily on Chinese-speaking populations in Taiwan and Hong Kong, alongside comparison with the original U.S. sample. These comparisons are informative but should not be taken as representing East Asian cultures as a whole. Taiwan, Hong Kong, and the United States differ on many dimensions beyond language and culture, including educational systems, financial environments, health-care institutions, and norms surrounding social feedback. Future studies should examine the IPS across additional societies and languages, ideally using samples that are more closely matched on demographic characteristics and research designs that can separate language, culture, and institutional context.

Taken together, the present findings establish the Chinese version of IPS as a promising instrument for studying information seeking and avoidance in Chinese-speaking populations. The cross-cultural patterns observed here illustrate the kinds of questions that become possible once a translated scale has been psychometrically validated, but they should be regarded as a starting point rather than a final account of cultural differences in information preferences. Future research can build on this foundation by examining the mechanisms underlying domain-specific information avoidance, testing behavioral prediction across multiple domains, and evaluating whether interventions designed to reduce harmful information avoidance need to be tailored to specific cultural and institutional contexts.

Funding Statement: This study was financially supported by the Research Project Grants provided by the National Science and Technology Council, Taiwan (NSTC 110-2410-H-001-100-MY3 and NSTC 113-2628-H-001-015-MY3).

Declaration of Competing Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability: All the data and codes used in this paper, including both Studies 1 and 2, are available at the following OSF webpage `https://osf.io/j7h3r/`.

IRB Approvals: The study was reviewed and approved by the authors' institutions: Academia Sinica (AS-IRB-HS-23048) and The Hong Kong University of Science and Technology (HREP-2024-0108).

Author Contributions: M.W., G.C.S., D.H., and G.L. conceptualized the study. All authors contributed to the methodology. M.W. performed the formal analyses. M.W., G.C.S., E.H.H., D.H., and G.L. conducted the investigation. E.H.H., P.C., and G.L. validated the analyses. M.W. and G.C.S. curated the data. G.C.S. and D.H. provided resources and administered the project. M.W. and G.C.S. wrote the original draft, and E.H.H. and G.L. reviewed and edited the manuscript. G.L. supervised the research. G.C.S. acquired funding.

References

Alacreu-Crespo, Adrián, María C Fuentes, Diana Abad-Tortosa, Irene Cano-Lopez, Esperanza González, and Miguel Ángel Serrano. 2019. "Spanish validation of General Decision-Making Style scale: Sex invariance, sex differences and relationships with personality and coping styles." *Judgment and Decision Making*, 14(6): 739–751.

Arceneaux, Kevin, Martin Johnson, and Chad Murphy. 2012. "Polarized political communication, oppositional media hostility, and selective exposure." *The Journal of Politics*, 74(1): 174–186.

Ash, Elliott, Daniel Sgroi, Anthony Tuckwell, and Shi Zhuo. 2023. "Mindfulness reduces information avoidance." *Economics Letters*, 224: 110997.

Brashers, Dale E, Daena J Goldsmith, and Elaine Hsieh. 2002. "Information seeking and avoiding in health contexts." *Human Communication Research*, 28(2): 258–271.

Brislin, Richard W. 1970. "Back-translation for cross-cultural research." *Journal of cross-cultural psychology*, 1(3): 185–216.

Byrne, Barbara M. 1994. *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming.* Sage.

Byrne, Barbara M, Richard J Shavelson, and Bengt Muthén. 1989. "Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance." *Psychological Bulletin*, 105(3): 456.

Chang, Yu-Wei, Wei-Kang Huang, and Rung-Ching Tsai. 2015. "DIF detection using multiple-group categorical CFA with minimum free baseline approach." *Journal of Educational Measurement*, 52(2): 181–199.

Cheung, Gordon W, and Roger B Rensvold. 2002. "Evaluating goodness-of-fit indexes for testing measurement invariance." *Structural equation modeling*, 9(2): 233–255.

Choi, Seung W, Laura E Gibbons, and Paul K Crane. 2011. "Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations." *Journal of Statistical Software*, 39(8): 1.

Druckman, James N, Erik Peterson, and Rune Slothuus. 2013. "How elite partisan polarization affects public opinion formation." *American Political Science Review*, 107(1): 57–79.

Fischer, Dov. 2013. "HP's Purchase of Autonomy:'Hubris at the Top'and Failure of Internal Controls." *Available at SSRN 2357327*.

Forero, Carlos G, Alberto Maydeu-Olivares, and David Gallardo-Pujol. 2009. "Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation." *Structural equation modeling*, 16(4): 625–641.

Foust, Jeremy L, and Jennifer M Taber. 2023. "Information avoidance: Past perspectives and future directions." *Perspectives on Psychological Science*, 17456916231197668.

Ganguly, Ananda, and Joshua Tasoff. 2017. "Fantasy and dread: The demand for information and the consumption utility of the future." *Management Science*, 63(12): 4037–4060.

Golman, Russell, David Hagmann, and George Loewenstein. 2017. "Information avoidance." *Journal of economic literature*, 55(1): 96–135.

Golman, Russell, George Loewenstein, Andras Molnar, and Silvia Saccardo. 2022. "The demand for, and avoidance of, information." *Management Science*, 68(9): 6454–6476.

Ho, Emily H, David Hagmann, and George Loewenstein. 2021. "Measuring information preferences." *Management Science*, 67(1): 126–145.

Jessee, Stephen A. 2021. "Estimating individuals'political perceptions while adjusting for differential item functioning." *Political Analysis*, 29(1): 1–18.

Kline, Rex B. 2023. *Principles and practice of structural equation modeling.* Guilford Publications.

Klotz, Anthony C, Brian W Swider, and Seo Hyun Kwon. 2023. "Back-translation practices in organizational research: Avoiding loss in translation." *Journal of Applied Psychology*, 108(5): 699.

Knobloch-Westerwick, Silvia, and Simon M Lavis. 2017. "Selecting serious or satirical, supporting or stirring news? Selective exposure to partisan versus mockery news online videos." *Journal of Communication*, 67(1): 54–81.

Knobloch-Westerwick, Silvia, Benjamin K Johnson, and Axel Westerwick. 2015. "Confirmation bias in online searches: Impacts of selective exposure before an election on political attitude strength and shifts." *Journal of Computer-Mediated Communication*, 20(2): 171–187.

Kobayashi, Kenji, Silvio Ravaioli, Adrien Baranès, Michael Woodford, and Jacqueline Gottlieb. 2019. "Diverse motives for human curiosity." *Nature Human Behaviour*, 3(6): 587–595.

Kőszegi, Botond, and Matthew Rabin. 2006. "A model of reference-dependent preferences." *The Quarterly Journal of Economics*, 121(4): 1133–1165.

Liew, Shi Xian, Jake R Embrey, and Ben R Newell. 2023. "The non-unitary nature of information preference." *Psychonomic Bulletin & Review*, 30(5): 1966–1974.

Little, Todd D. 2024. *Longitudinal structural equation modeling.* Guilford Publications.

Li, Yufeng, Juanjuan Meng, Changcheng Song, and Kai Zheng. 2021. "Information avoidance and medical screening: A field experiment in China." *Management Science*, 67(7): 4252–4272.

Markus, Hazel R, and Shinobu Kitayama. 1991. "Cultural variation in the self-concept." In *The self: Interdisciplinary approaches*. 18–48. Springer.

Millsap, Roger E. 2012. *Statistical approaches to measurement invariance.* Routledge.

Millsap, Roger E., and María Olivera-Aguilar. 2012. "Investigating measurement invariance using confirmatory factor analysis." In *Handbook of structural equation modeling.* , ed. Rick H. Hoyle, 380–392. The Guilford Press.

Moon, Won-Ki, Lee Ann Kahlor, Janet Zheng Yang, and Hayoung Sally Lim. 2022. "Risk perception, affect, and information avoidance during the 2016 US presidential election." *Journal of Risk Research*, 25(7): 860–873.

Oster, Emily, Ira Shoulson, and E Ray Dorsey. 2013. "Optimal expectations and limited medical testing: Evidence from Huntington disease." *American Economic Review*, 103(2): 804–830.

Pennycook, Gordon, and David G Rand. 2021. "The psychology of fake news." *Trends in Cognitive Sciences*, 25(5): 388–402.

Putnick, Diane L, and Marc H Bornstein. 2016. "Measurement invariance conventions and reporting: The state of the art and future directions for psychological research." *Developmental Review*, 41: 71–90.

Reise, Steven P, Keith F Widaman, and Robin H Pugh. 1993. "Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance." *Psychological bulletin*, 114(3): 552.

Sass, Daniel A, Thomas A Schmitt, and Herbert W Marsh. 2014. "Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators." *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2): 167–180.

Somerville, Leah H, Stephanie F Sasse, Megan C Garrad, Andrew T Drysdale, Nadine Abi Akar, Catherine Insel, and Robert C Wilson. 2017. "Charting the expansion of strategic exploratory behavior during adolescence." *Journal of Experimental Psychology: General*, 146(2): 155.

Steenkamp, Jan-Benedict EM, and Hans Baumgartner. 1998. "Assessing measurement invariance in cross-national consumer research." *Journal of Consumer Research*, 25(1): 78–90.

Steinmetz, Holger. 2013. "Analyzing observed composite differences across groups." *Methodology*.

Sunstein, Cass R. 2001. *Republic.com.* Princeton, NJ:Princeton University Press.

Tse, Winnie Wing-Yee, Mark HC Lai, and Yichi Zhang. 2023. "Does strict invariance matter? Valid group mean comparisons with ordered-categorical items." *Behavior Research Methods*, 1–23.

Vandenberg, Robert J, and Charles E Lance. 2000. "A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research." *Organizational Research Methods*, 3(1): 4–70.

Van de Schoot, Rens, Peter Lugtig, and Joop Hox. 2012. "A checklist for testing measurement invariance." *European Journal of Developmental Psychology*, 9(4): 486–492.

Wang, Jichuan, and Xiaoqian Wang. 2019. *Structural equation modeling: Applications using Mplus.* John Wiley & Sons.

Weber, Elke U, Ann-Renee Blais, and Nancy E Betz. 2002. "A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors." *Journal of Behavioral Decision Making*, 15(4): 263–290.

Welindt, Dillon, David M Condon, and Sara J Weston. 2023. "Measurement invariance of the Domain-Specific Risk-Taking (DOSPERT) scale." *Journal of Behavioral Decision Making*, 36(4): e2337.

Widaman, Keith F, and Steven P Reise. 1997. *Exploring the measurement invariance of psychological instruments: Applications in the substance use domain.* American Psychological Association.

Wu, Pei-Chen, and Tsai-Wei Huang. 2014. "Gender-related invariance of the Beck Depression Inventory II for Taiwanese adolescent samples." *Assessment*, 21(2): 218–226.

Xu, Yuning, and Samuel B Green. 2016. "The impact of varying the number of measurement invariance constraints on the assessment of between-group differences of latent means." *Structural Equation Modeling: A Multidisciplinary Journal*, 23(2): 290–301.

Yoon, Myeongsun, and Eun Sook Kim. 2014. "A comparison of sequential and nonsequential specification searches in testing factorial invariance." *Behavior Research Methods*, 46: 1199–1206.

Yoon, Myeongsun, and Roger E Millsap. 2007. "Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study." *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3): 435–463.

Zhang, Don C, Garett C Foster, and Michael G McKenna. 2019. "Is the DOSPERT gender invariant? A psychometric test of measurement invariance." *Journal of Behavioral Decision Making*, 32(2): 203–211.

Zumbo, Bruno D. 1999. "A handbook on the theory and methods of differential item functioning (DIF)." *Ottawa: National Defense Headquarters*, 160.

Appendix

A Items on the Information Preferences Scale: Original Scale and Traditional Chinese Translation

Table A.1: Items on the Information Preferences Scale (Side-by-Side)

Original Item	Chinese Translation	Dimension
1. As part of a semi-annual medical checkup, your doctor asks you a series of questions. The answers to these questions can be used to estimate your life expectancy (the age you are predicted to live to). Do you want to know how long you can expect to live?	在每半年一次的健康檢查，醫師詢問了一系列的問題，以便估計你的預期壽命（你可以活到幾歲）。請問你會想知道自己預期能活多久嗎？	H
2. You provide some genetic material to a testing service to learn more about your ancestors. You are then told that the same test can, at no additional cost, tell you whether you have an elevated risk of developing Alzheimer's. Do you want to know whether you have a high risk of developing Alzheimer's?	你提供了基因樣本給一項檢測服務，想藉此更了解自己的祖先。而你也得知透過同一項檢測，不需要額外的費用，就可以知道自己罹患阿茲海默症的風險如何。你會想知道自己是否屬於阿茲海默症的高風險群嗎？	H
3. At your annual checkup, you are given the option to see the results of a diagnostic test which can identify, among other things, the extent to which your body has suffered long-term effects from stress. Do you want to know how much lasting damage your body has suffered from stress?	在你的年度健康檢查中，你有機會能選擇要不要看一個檢測結果，該檢測結果能告訴你一些訊息，包括你的身體長期受壓力損害的程度。請問你是否想知道你的身體因爲壓力受到長期傷害的程度嗎？	H
4. Ten years ago, you had the opportunity to invest in two retirement funds: Fund A and Fund B. For the past 10 years, you have invested all your retirement savings in Fund A. Do you want to know the balance you would have, if you had invested in Fund B instead?	你在十年前有個機會去投資兩種退休基金（A 基金或 B 基金）。在過去十年你選擇將自己的退休儲蓄全部都放在 A 基金。你會想知道，當初"如果"把錢放在 B 基金的收益嗎？	F

Original Item	Chinese Translation	Dimension
5. You decide to go to the theater for your birthday and give your close friend (or partner) your credit card so they can purchase tickets for the two of you, which they do. You aren't sure, but suspect that the tickets may have been expensive. Do you want to know how much the tickets cost?	你決定跟你的好友（或伴侶）去劇場看表演來慶祝自己生日，並把信用卡交給他們去刷卡買兩張票，你事先並不確定價格，但票價可能很貴。請問你會想知道這劇票的價格嗎？	F
6. You bought an electronic appliance at a store at what seemed like a reasonable, though not particularly low, price. A month has passed, and the item is no longer returnable. You see the same appliance displayed in another store with a sign announcing 'SALE.' Do you want to know the price you could have bought it for?	你在店裡購買了一項價格不算特別低，但還算合理的電子產品。結果一個月後，退貨期已過，你無意間在另一家店看到同樣的電子產品在特價拍賣。請問你會想知道它現在賣多少錢嗎？	F
7. You gave a close friend one of your favorite books for her birthday. Visiting her apartment a couple of months later, you notice the book on her shelf. She never said anything about it; do you want to know if she liked the book?	好友生日，你送了一本自己最喜愛的書給她作爲生日禮物，幾個月之後你拜訪她家，發現那本書放在她的書櫃上。你的朋友從未提到這本書，但你會想要知道她是否喜歡這本書嗎？	I
8. Someone has described you as quirky, which could be interpreted in a positive or negative sense. Do you want to know which interpretation he intended?	有人形容你是個"奇人": 這個詞可以詮釋成正面的（例如想法不落俗套），也可能是負面的意思（例如很古怪）。你會想要知道眞正的意思嗎？	I
9. You gave a toast at your best friend's wedding. Your best friend says you did a good job, but you aren't sure if he or she meant it. Later, you overhear people discussing the toasts. Do you want to know what people really thought of your toast?	你在摯友的婚禮上致賀詞。你的摯友說你做得非常好，但你不是很確定他是否眞的這麼認爲。之後，你無意中聽到其他人在談論你的賀詞內容，你會想要知道人們心裡面是怎麼評價的嗎？	I
10. As part of a fund-raising event, you agree to post a picture of yourself and have people guess your age (the closer they get, the more they win). At the end of the event, you have the option to see people's guesses. Do you want to learn how old people guessed that you are?	作爲募款活動的一個環節，你同意附上自己的照片並讓大家猜測你的年齡（愈接近眞實年齡則贏得愈多）。在活動尾聲時，你可以選擇是否要看大家對你年齡多大的猜測。你會想要知道這項資訊嗎？	I

Original Item	Chinese Translation	Dimension
11. You have just participated in a psychological study in which all the participants rate one-anothers' attractiveness. The experimenter gives you an option to see the results for how people rated you. Do you want to know how attractive other people think you are?	你參加了一個心理學研究，在研究過程中，每個人都要對其他參與者的外貌進行評分。實驗人員最後會給你一個機會，讓你自由選擇是否要看其他人對你外貌的評價。請問你是否會想看其他人眼中您多有吸引力嗎？	I
12. Some people seek out information even when it might be painful. Others avoid getting information that they suspect might be painful, even if it could be useful. How would you describe yourself?	即使可能帶來不愉快，一些人仍會持續搜尋更多資訊；另外一些人則相反，會因爲擔心帶來不愉快，而避免獲得更多資訊，即使這些資訊可能是有用的。請問你是哪一類型的人呢？	G
13. If people know bad things about my life that I don't know, I would prefer not to be told	如果別人知道一些關於我的不好的消息（但我自己不知道），那我寧願不要被告知。	G

Note. H = Healthy, F = Financial, I = Interpersonal, and G = General.

B Demographics of the Taiwan sample

- Gender

Gender (coding)	N	Percent
Male	604	50.4
Female	594	49.6
Total	1198	100

- Age

Age (coding)	N	Percent
20-29	204	17.0
30-39	221	18.5
40-49	257	21.5
50-59	235	19.6
60 or older	281	23.4
Total	1198	100

- Education

Education (coding)	N	Percent
Senior high or below	166	13.9
Associate degree	157	13.1
Bachelor's Degree	628	52.4
Master's Level or above	247	20.6
Total	1198	100

C Demographics of the U.S. Sample

- Gender

Gender (coding)	N	Percent
Male	201	52.9
Female	179	47.1
Total	380	100

- Age

Age (coding)	N	Percent
20-29	134	35.3
30-39	147	38.7
40-49	59	15.5
50-59	30	7.9
60 or older	10	2.6
Total	380	100

- Education

Education (coding)	N	Percent
Senior high or below	116	30.5
Associate degree	67	17.6
BA	155	40.8
MA or above	34	8.9
Professional Degree	8	2.1
Total	380	100

D Descriptive Statistics for the Items of IPS (USA & Taiwan Samples)

	Taiwan	n = 1198	United States	n = 380
Item No.	Mean	SD	Mean	SD
IPS 1	2.63	0.92	2.58	1.02
IPS 2	3.11	0.81	3.14	.95
IPS 3	3.25	0.76	3.02	1.00
IPS 4	2.76	0.91	2.50	1.16
IPS 5	3.19	0.78	3.34	.92
IPS 6	2.86	0.84	2.79	1.05
IPS 7	2.57	0.85	3.08	0.91
IPS 8	2.71	0.82	2.93	0.95
IPS 9	2.67	0.85	2.76	1.03
IPS 10	2.75	0.86	3.17	0.98
IPS 11	2.76	0.85	2.76	1.07
IPS 12	2.83	0.72	2.89	0.86
IPS 13	2.62	0.88	2.86	0.89

E Measurement Invariance

When translating a validated scale into a new language for participants from another culture, it is essential to examine **measurement invariance (MI)** across scales and populations. MI ensures that the translated scale assesses the same construct as the original, possessing the identical structure, meaning, and item functioning as the original population (Steenkamp and Baumgartner, 1998; Vandenberg and Lance, 2000). When measurement invariance holds, score differences between different cultures/groups can be interpreted as true differences in the underlying latent construct between the responding samples rather than artifacts of translation, cultural, or contextual differences (Millsap, 2012). It means that people with the same level of the underlying trait respond in comparable ways, the scale's structure is stable, and comparisons across groups are meaningful.

Given its importance, researchers have collectively proposed commonly accepted procedures to systematically establish different levels of MI for a given instrument (Steenkamp and Baumgartner, 1998; Van de Schoot et al., 2012). The MI results for Study 1 are provided in Appendix F.

F Measurement Invariance for Study 1

Table F.1 presents the overall fit of each invariance model. It is apparent that across all levels of invariance tested (i.e., configural, metric, and scalar), CFI and TLI are all above the recommended criterion of .90 (Byrne, 1994; Little, 2024), the values of RMSEA are all on or below than the appropriate cutoff of .08 (Little, 2024), indicating good fit, and all values of SRMR are below the recommended criterion of .08 (Byrne, 1994).

For our empirical results, full metric invariance was verified ($p = .12$). However, full scalar invariance did not hold, as indicated by the significant χ^2 test of difference between scalar and metric invariance ($p < .05$); therefore, we focus on establishing a model of *partial* scalar invariance, which permits certain parameters to be non-invariant across groups (Little, 2024).

Table F.1: Baseline Models and Measurement Invariance Fit Statistics

Level of MI	CFI	TLI	RMSEA	SRMR	χ^2 test of diff
Baseline Model (USA)	.97	.95	.06	.05	—
Baseline Model (Taiwan)	.94	.92	.08	.05	—
Configural Invariance	.95	.93	.07	.05	—
Metric Invariance	.96	.95	.06	.05	$p = .12$
Scalar Invariance	.92	.91	.08	.05	$p < .001$

Partial Scalar Invariance

To achieve partial scalar invariance, a series of parameter constraints are progressively removed. The sequence of removed parameter constraints is detailed in Appendix H. The qualitative inspection of the items revealed that East Asian respondents mainly tended to employ higher thresholds (intercepts) compared to Americans when assessing items related to others' opinions or attitudes toward them. This suggests that, even with the same level of information preference, East Asians tend to give lower ratings (a higher threshold corresponds to lower ratings, while a lower threshold corresponds to higher ratings).

In contrast, when respondents rated items unrelated to others' opinions or attitudes toward them, the pattern reversed, with American respondents using higher thresholds (intercepts) than their East Asian counterparts. The detailed information is also presented in the appendix. In Appendix G, we present sensitivity analyses from logistic regression-based method aiming at detecting item bias (Choi et al., 2011; Zumbo, 1999). The results showed that they are mainly consistent with the finding discussed above.

G Item Bias

To identify Differential Item Functioning (DIF) using the Logistic Regression-based method (Zumbo, 1999), we formulated a series of proportional odds models following the approach outlined by Choi et al. (2011):

$$
\begin{aligned}
M1 &: \text{logit}(P(x_i)) = \tau_0 + \tau_1\theta \\
M2 &: \text{logit}(P(x_i)) = \tau_0 + \tau_1\theta + \tau_2 Z \\
M3 &: \text{logit}(P(x_i)) = \tau_0 + \tau_1\theta + \tau_2 Z + \tau_3\theta Z
\end{aligned}
$$

Significant differences between M2 and M1 indicate uniform DIF, whereas significant differences between M3 and M2 indicate nonuniform DIF. We utilized the *lordif* package to conduct DIF analyses for the instrument measuring information preferences, with the group variable being respondents from the USA and Taiwan and conducted a series of likelihood ratio (LR) χ^2 tests.

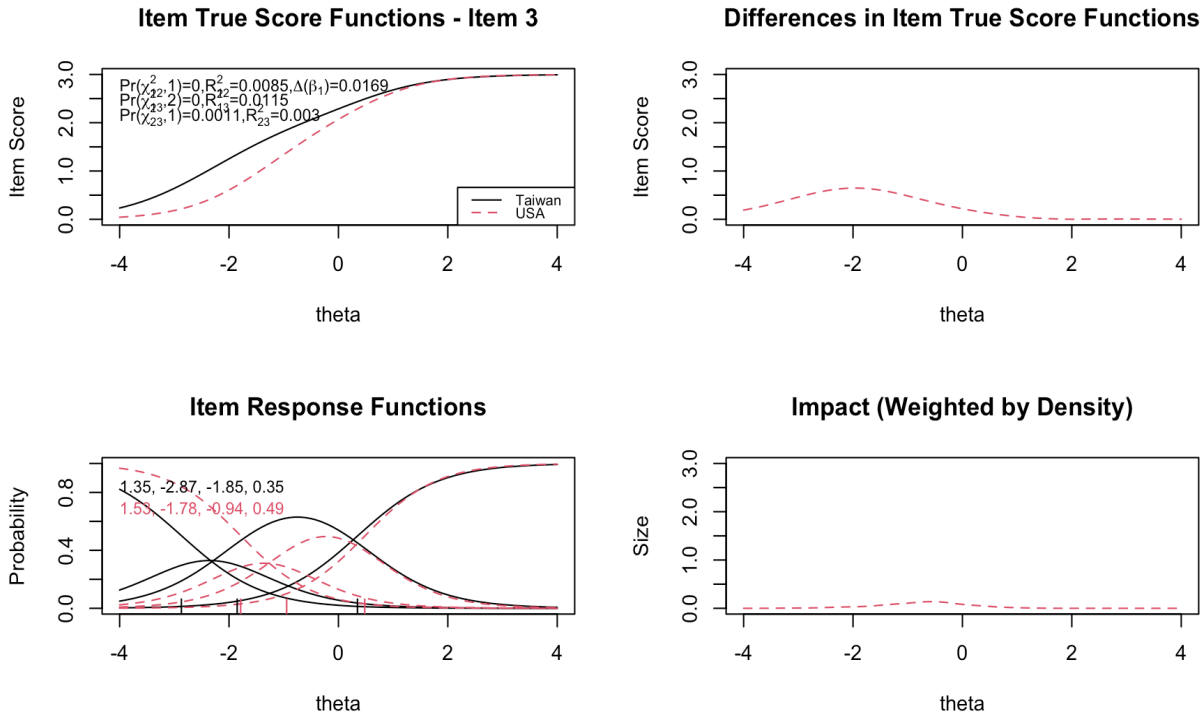
Empirically, items *3, 4, 5, 7, 8, 10*, and *13* on the IPS display non-uniform DIF (see Figures G.1–G.7) between American and East Asian (Taiwanese) respondents. Descriptions of these items with DIF are provided in Table G.1.

Table G.1: Item Number and Description of DIF Items.

Item Number	Item Description
3	***At your annual checkup, you are given the option to see the results of a diagnostic test, which can identify, among other things, the extent to which your body has suffered long-term effects from stress. Do you want to know how much lasting damage your body has suffered from stress?***
4	***Ten years ago, you had the opportunity to invest in two retirement funds: Fund A and Fund B. For the past 10 years, you have invested all your retirement savings in Fund A. Do you want to know the balance you would have if you had invested in Fund B instead?***
5	*You decide to go to the theater for your birthday and give your close friend (or partner) your credit card so they can purchase tickets for the two of you, which they do. You aren't sure but suspect that the tickets may have been expensive. Do you want to know how much the tickets cost?*
7	*You gave a close friend one of your favorite books for her birthday. Visiting her apartment a couple of months later, you notice the book on her shelf. She never said anything about it; do you want to know if she liked the book?*
8	*Someone has described you as quirky, which could be interpreted in a positive or negative sense. Do you want to know which interpretation he intended?*
10	*As part of a fundraising event, you agree to post a picture of yourself and have people guess your age (the closer they get, the more they win). At the end of the event, you have the option to see people's guesses. Do you want to learn how old people guessed that you are?*
13	*If people know bad things about my life that I don't know, I would prefer not to be told.*

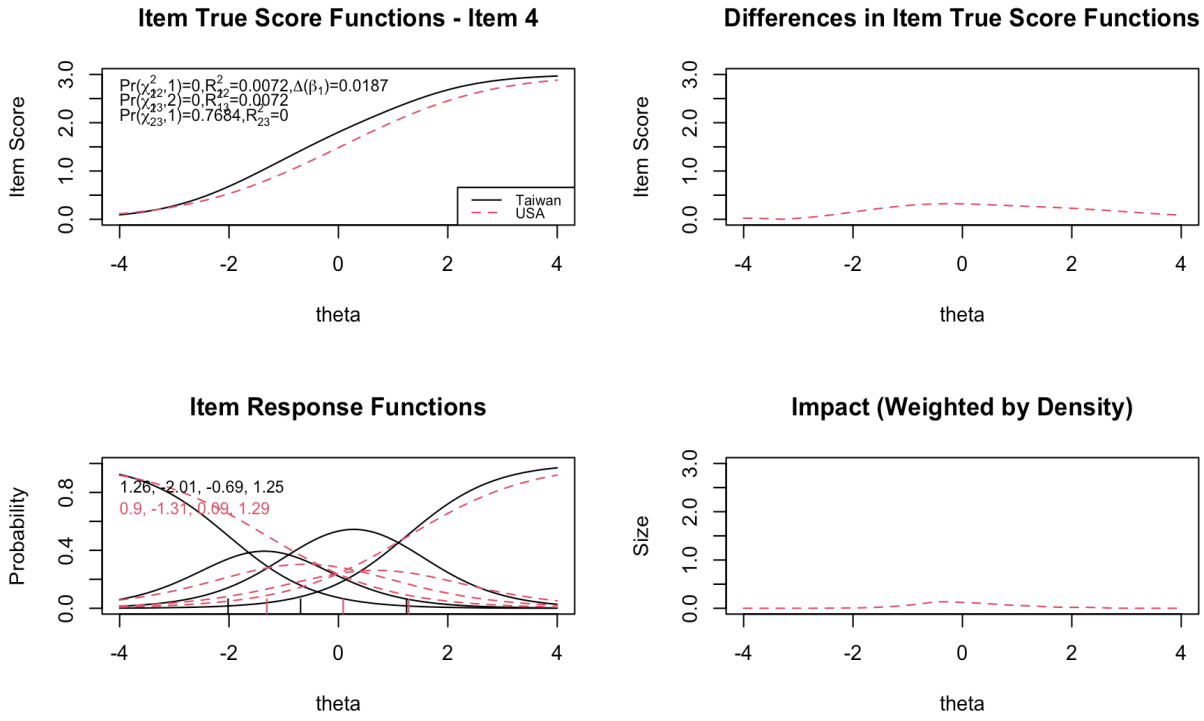
Note. The use of ***bold*** *font in the item descriptions indicates that these items relate to information that does not involve opinions or evaluations from others.*

Figure G.1: Item with DIF: Item 3



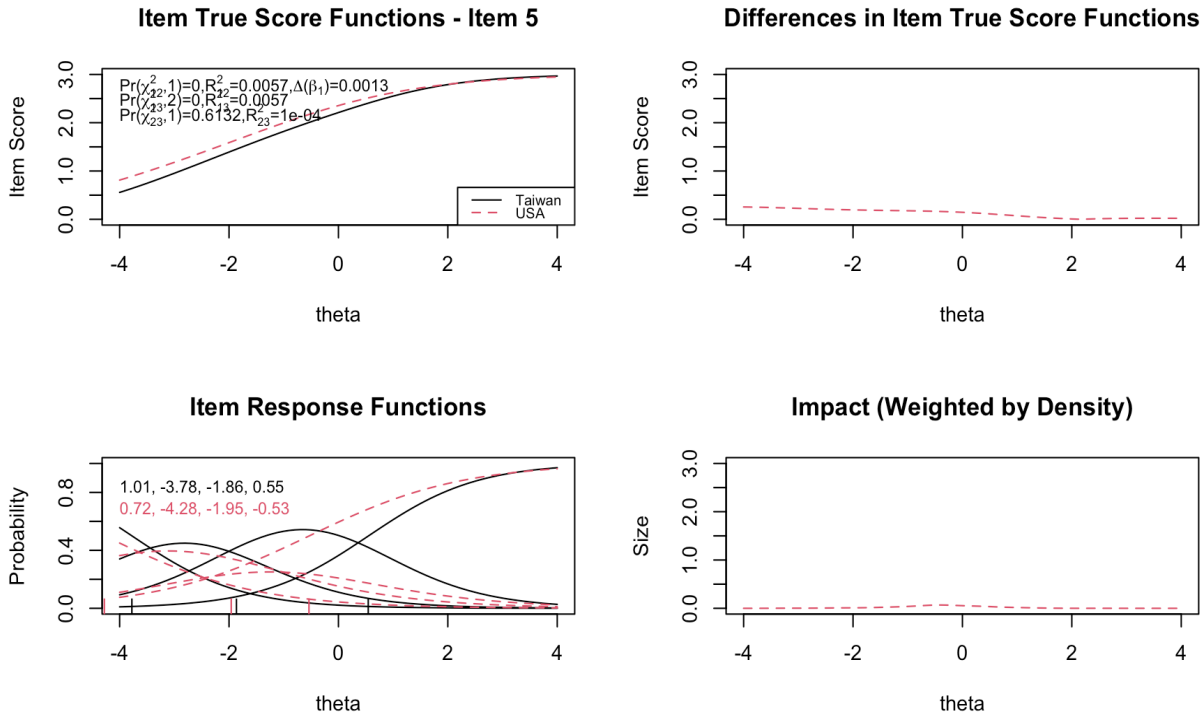
Note. Differential item functioning for Item 3 between the Taiwanese (solid line) and U.S. (dashed line) samples, estimated with the *lordif* package. The panels display the group-specific item true score functions, the difference between them, the item response category functions, and the density-weighted impact, each plotted against the latent trait θ.

Figure G.2: Item with DIF: Item 4



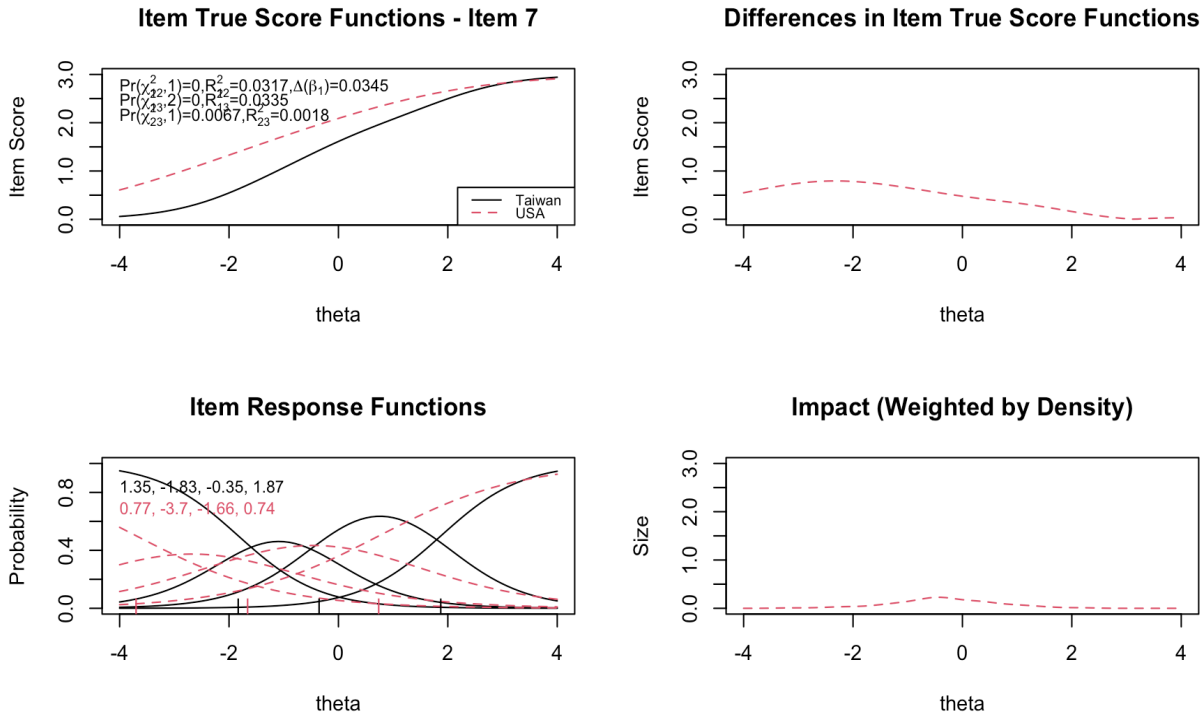
Note. Differential item functioning for Item 4 between the Taiwanese (solid line) and U.S. (dashed line) samples, estimated with the *lordif* package. The panels display the group-specific item true score functions, the difference between them, the item response category functions, and the density-weighted impact, each plotted against the latent trait θ.

Figure G.3: Item with DIF: Item 5



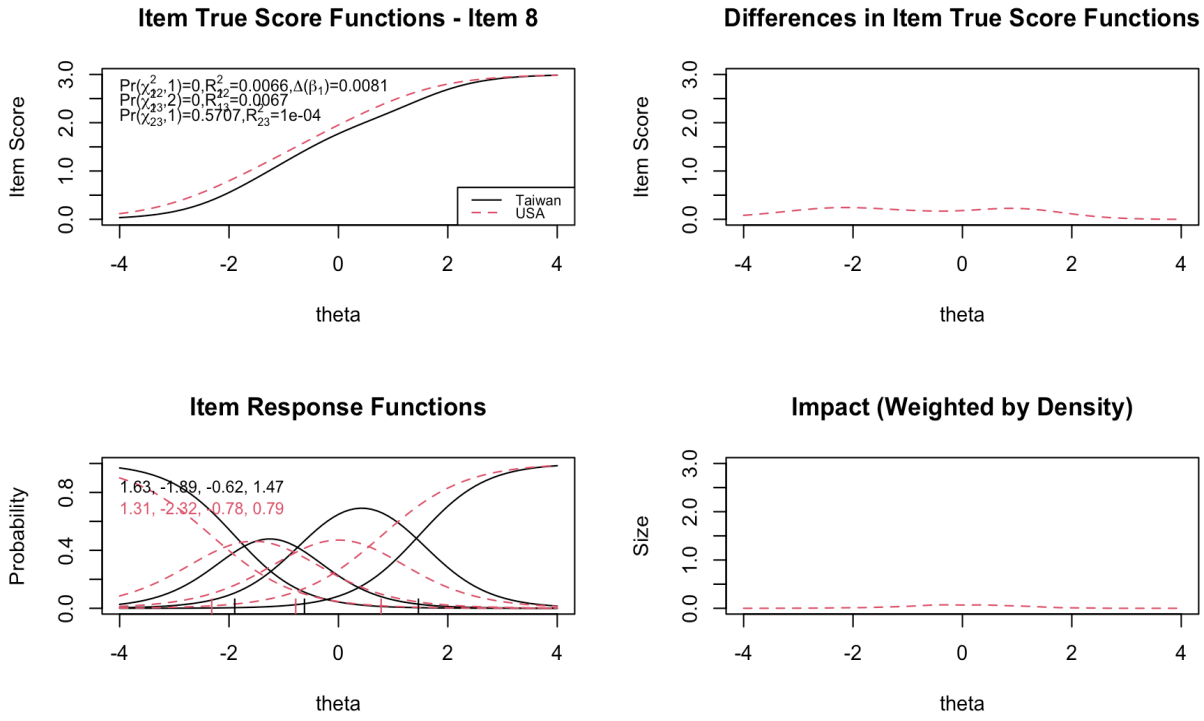
Note. Differential item functioning for Item 5 between the Taiwanese (solid line) and U.S. (dashed line) samples, estimated with the *lordif* package. The panels display the group-specific item true score functions, the difference between them, the item response category functions, and the density-weighted impact, each plotted against the latent trait θ.

Figure G.4: Item with DIF: Item 7



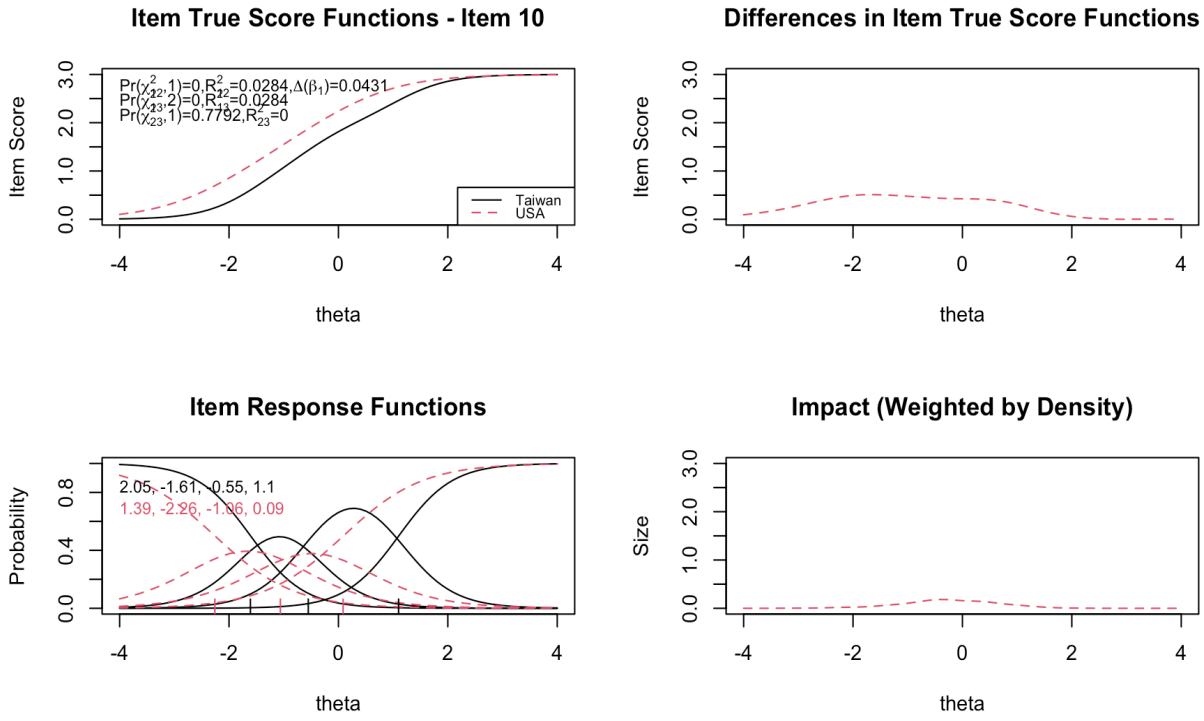
Note. Differential item functioning for Item 7 between the Taiwanese (solid line) and U.S. (dashed line) samples, estimated with the *lordif* package. The panels display the group-specific item true score functions, the difference between them, the item response category functions, and the density-weighted impact, each plotted against the latent trait θ.

Figure G.5: Item with DIF: Item 8



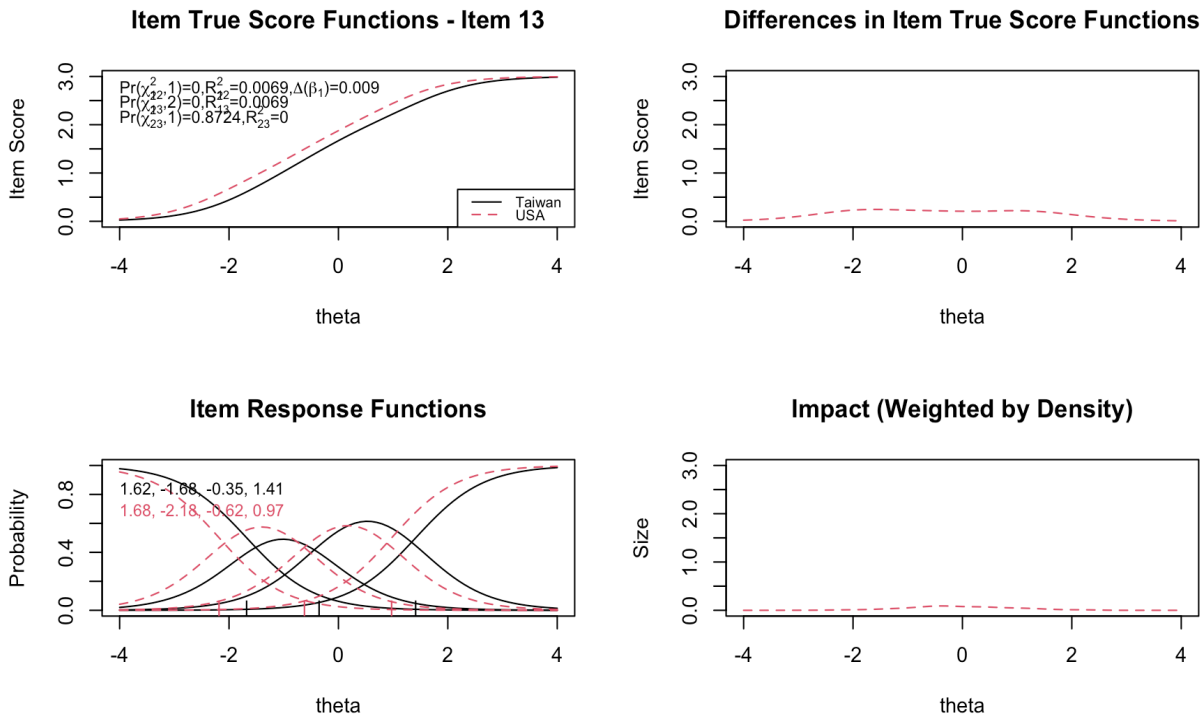
Note. Differential item functioning for Item 8 between the Taiwanese (solid line) and U.S. (dashed line) samples, estimated with the *lordif* package. The panels display the group-specific item true score functions, the difference between them, the item response category functions, and the density-weighted impact, each plotted against the latent trait θ.

Figure G.6: Item with DIF: Item 10



Note. Differential item functioning for Item 10 between the Taiwanese (solid line) and U.S. (dashed line) samples, estimated with the *lordif* package. The panels display the group-specific item true score functions, the difference between them, the item response category functions, and the density-weighted impact, each plotted against the latent trait θ.

Figure G.7: Item with DIF: Item 13



Note. Differential item functioning for Item 13 between the Taiwanese (solid line) and U.S. (dashed line) samples, estimated with the *lordif* package. The panels display the group-specific item true score functions, the difference between them, the item response category functions, and the density-weighted impact, each plotted against the latent trait θ.

H Sequential Backward Specification Search Method for achieving Partial Scalar Invariance across American and Taiwanese samples

Table H.1: Measurement invariance testing results of the Information Preferences Scale across American and Taiwanese samples.

Model (Item, Threshold)	χ^2	df	$\Delta\chi^2$	CFI	TLI	RMSEA
Metric Invariance	520.53	127	–	.96	.95	.06
Partial Scalar Invariance (IPS5, 3)	872.15	148	293.37***	.93	.92	.08
Partial Scalar Invariance (IPS7, 1)	820.81	147	245.72***	.93	.93	.08
Partial Scalar Invariance (IPS7, 2)	760.37	146	205.67***	.94	.93	.07
Partial Scalar Invariance (IPS10, 3)	701.35	145	165.13***	.94	.94	.07
Partial Scalar Invariance (IPS11, 2)	675.35	144	144.72***	.95	.94	.07
Partial Scalar Invariance (IPS6, 3)	653.69	143	129.01***	.95	.94	.07
Partial Scalar Invariance (IPS2, 3)	631.43	142	111.44***	.95	.95	.07
Partial Scalar Invariance (IPS13, 1)	618.62	141	96.992***	.95	.95	.07
Partial Scalar Invariance (IPS7, 3)	607.30	140	84.237***	.95	.95	.07
Partial Scalar Invariance (IPS8, 3)	591.55	139	70.318***	.95	.95	.06
Partial Scalar Invariance (IPS4, 3)	580.75	138	57.993***	.95	.95	.06
Partial Scalar Invariance (IPS3, 2)	568.47	137	47.188***	.96	.95	.06
Partial Scalar Invariance (IPS3, 1)	568.95	136	46.428***	.96	.95	.06
Partial Scalar Invariance (IPS6, 3)	555.65	135	32.51***	.96	.95	.06
Partial Scalar Invariance (IPS11, 1)	541.16	134	19.69**	.96	.95	.06
Partial Scalar Invariance (IPS8, 3)	532.171	133	10.555	.96	.95	.06

Note. Parentheses contain information on item number and corresponding threshold. $\Delta\chi^2$ indicated the variation between each partial scalar invariance model and the metric invariance model. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table H.2: Non-invariant Items.

Subscale	Item	Threshold	USA	Taiwan	Subgroup with Higher Threshold
Health	2	3	0.118	0.434	Taiwan
Health	3	1	-1.291	-1.740	USA
Health	3	2	-0.734	-1.191	USA
Financial	4	2	-0.258	-0.418	USA
Financial	4	3	0.428	0.797	Taiwan
Financial	5	3	-0.488	0.297	Taiwan
Financial	6	3	0.202	0.808	Taiwan
Interpersonal	7	1	-2.061	-1.153	Taiwan
Interpersonal	7	2	-0.866	-0.202	Taiwan
Interpersonal	7	3	0.690	1.204	Taiwan
Interpersonal	8	3	0.796	1.072	Taiwan
Interpersonal	9	2	-0.053	-0.346	USA
Interpersonal	10	3	0.319	0.927	Taiwan
Interpersonal	11	1	-0.930	-1.310	USA
Interpersonal	11	2	-0.002	-0.464	USA
General	13	1	-1.636	-1.178	Taiwan

Note. Non-invariant Items with different thresholds between the USA and Taiwanese samples in the Information Preferences Scale's final partial scalar invariance model.

I MI Analysis between HK and TW samples

Table I.1: Results of Measurement Invariance Between Hong Kong and Taiwan Samples.

Level of MI	χ^2	Df	CFI	TLI	RMSEA	SRMR	ΔCFI
Baseline Model (HK)	100.424	59	.94	.92	.05	.05	—
Configural Invariance	497.898	118	.96	.94	.05	.05	—
Metric Invariance	444.537	127	.96	.95	.05	.05	.007
Scalar Invariance	424.253	149	.97	.97	.05	.05	.005

Note. HK = Hong Kong; TW = Taiwan; MI = Measurement Invariance. ΔCFI values indicate the change relative to the preceding, less restrictive model.

J Correlation Matrices for Study 2

We also show the correlation for the conditions of "English First" and of "Chinese First" in Appendix J. For the "English First" condition, the entire IPS at both Time 1 and Time 2 was predictive of these business school students' actual information-seeking decisions ($r = 0.27$ and 0.18 respectively). When further examining the predictive validity for actual information-seeking decisions, only health and general subscales are insignificant. the financial and interpersonal domains at both Time 1 and Time 2 were also predictive (financial: $r = 0.33$ and 0.18 respectively; interpersonal: $r = 0.26$ and 0.20 respectively); For the "Chinese First" condition, the total IPS score at both time points was predictive of these business school students' information-seeking decisions in the real world. Regarding the subscales, the interpersonal domains at both Time 1 and Time 2 were significant predictors ($r = 0.19$ and 0.28, respectively). In contrast, the health and general subscales again showed weaker associations with actual information-seeking decisions, appearing either non-significant at both time points or significant at only one.

Table J.1: Pearson Correlations for the IPS and Its Subscales at Time 1

Variable	M	SD	1	2	3	4	5
1. Full IPS	3.04	0.41	-				
2. Health	3.24	0.59	.71***	-			
3. Financial	2.99	0.64	.69***	.29***	-		
4. Interpersonal	3.03	0.49	.81***	.44***	.38***	-	
5. General	2.88	0.66	.54***	.28***	.21***	.25***	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

N = 267 (Hong Kong bilingual sample). Entries are Pearson correlations among the IPS total score and its four subscales at Time 1; M and SD are in the 1–4 response metric. Scores combine both language administrations.

Table J.2: Pearson Correlations for the IPS and Its Subscales at Time 1 and 2 (Test-retest Reliability)

Variable	M	SD	1	2	3	4	5	6	7	8	9	10
1. Full IPS T1	3.05	0.40	-									
2. Health T1	3.24	0.59	.71***	-								
3. Financial T1	2.99	0.64	.68***	.28***	-							
4. Interpersonal T1	3.03	0.49	.81***	.43***	.37***	-						
5. General T1	2.88	0.67	.54***	.28***	.21***	.25***	-					
6. Full IPS T2	2.96	0.40	.68***	.48***	.47***	.53***	.41***	-				
7. Health T2	3.11	0.62	.51***	.57***	.31***	.29***	.31***	.77***	-			
8. Financial T2	2.86	0.58	.50***	.21***	.58***	.33***	.25***	.71***	.36***	-		
9. Interpersonal T2	2.97	0.46	.58***	.31***	.32***	.62***	.27***	.82***	.45***	.48***	-	
10. General T2	2.85	0.60	.33***	.26***	.12	.21***	.42***	.54***	.39***	.16**	.26***	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

N = 267 (bilingual sample). Pearson correlations among the IPS total and subscales at Time 1 (variables 1–5) and Time 2 (variables 6–10); M and SD in the 1–4 metric. Scores combine both language administrations.

Table J.3: Pearson Correlations for the IPS and Its Subscales: Chinese-Language Version

Variable	M	SD	1	2	3	4	5
1. Full IPS	3.01	0.40	-				
2. Health	3.22	0.61	.71***	-			
3. Financial	2.94	0.60	.70***	.30***	-		
4. Interpersonal	2.99	0.48	.80***	.38***	.41***	-	
5. General	2.85	0.65	.58***	.35***	.24***	.28***	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

N = 267. Correlations are computed on the bilingual sample using each participant's Chinese-language administration of the IPS.

Table J.4: Pearson Correlations for the IPS and Its Subscales: English-Language Version

Variable	M	SD	1	2	3	4	5
1. Full IPS	2.99	0.41	-				
2. Health	3.13	0.60	.77***	-			
3. Financial	2.90	0.63	.71***	.35***	-		
4. Interpersoanl	3.01	0.47	.83***	.52***	.45***	-	
5. General	2.88	0.62	.50***	.32***	.14*	.24***	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

N = 267. Correlations are computed on the bilingual sample using each participant's English-language administration of the IPS.

Table J.5: Means, Standard Deviations, and Pearson Correlations for All Study Variables: English-First, Chinese-Second Order (E1C2).

Pearson Correlations for Variables in the IPS Experiment: C2E1

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Information Seeking	0.87	0.34	-										
2. Whole IPS T1	2.97	0.42	.27**	-									
3. Health T1	3.12	0.63	.09	.75***	-								
4. Financial T1	2.91	0.56	.33***	.74***	.39***	-							
5. Interpersonal T1	2.97	0.48	.26**	.82***	.39***	.52***	-						
6. General T1	2.86	0.62	.11	.62***	.45***	.28**	.33***	-					
7. Whole IPS T2	3.04	0.43	.18*	.70***	.53***	.53***	.57***	.40***	-				
8. Health T2	3.17	0.60	.02	.49***	.59***	.27**	.30***	.30***	.76***	-			
9. Financial T2	3.00	0.64	.18*	.49***	.34***	.55***	.34***	.22*	.72***	.37***	-		
1. Interpersonal T2	3.04	0.50	.20*	.59***	.29***	.42***	.66***	.29***	.84***	.52***	.44***	-	
2. General T2	2.91	0.66	.10	.42***	.36***	.28**	.26**	.40***	.54***	.31***	.22*	.28**	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

Pearson correlations among actual information seeking and the IPS total and subscales at Time 1 (English administration; variables 2–6) and Time 2 (Chinese administration; variables 7–11). Information Seeking (variable 1) indicates whether the participant chose to receive feedback on their peer-rated standing (1 = sought, 0 = declined). M and SD are in the 1–4 metric, except Information Seeking, for which M is the proportion who sought information. T1 = Time 1; T2 = Time 2.

Table J.6: Means, Standard Deviations, and Pearson Correlations for All Study Variables: Chinese-First, English-Second Order (C1E2).

Pearson Correlations for Variables in the IPS Experiment: C1E2

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Information Seeking	0.85	0.36	-										
2. Whole IPS T1	3.05	0.38	.25**	-									
3. Health T1	3.31	0.57	.20*	.66***	-								
4. Financial T1	2.98	0.64	.14	.65***	.21*	-							
5. Interpersonal T1	3.03	0.48	.19*	.78***	.35***	.30***	-						
6. General T1	2.84	0.67	.15	.55***	.27**	.20*	.22**	-					
7. Whole IPS T2	2.94	0.38	.24**	.67***	.48***	.46***	.47***	.39***	-				
8. Health T2	3.09	0.60	.08	.50***	.58***	.27**	.28**	.25**	.79***	-			
9. Financial T2	2.80	0.60	.20*	.48***	.19*	.62***	.24**	.21*	.68***	.33***	-		
1. Interpersonal T2	2.97	0.43	.28***	.58***	.34***	.30***	.57***	.28**	.83***	.52***	.44***	-	
2. General T2	2.84	0.58	.05	.27**	.23**	.01	.13	.44***	.45***	.33***	.04	.18*	-

Note. Significance levels: *p < .05, **p < .01, ***p < .001.

Pearson correlations among actual information seeking and the IPS total and subscales at Time 1 (Chinese administration; variables 2–6) and Time 2 (English administration; variables 7–11). Information Seeking (variable 1) indicates whether the participant chose to receive feedback on their peer-rated standing (1 = sought, 0 = declined). M and SD are in the 1–4 metric, except Information Seeking, for which M is the proportion who sought information. T1 = Time 1; T2 = Time 2.

K The Procedure for Establishing Measurement Invariance

Within the MG-CFA framework, testing for measurement invariance (MI) proceeds in four steps (Kline, 2023). It helps to return to the "ruler" analogy from Section 2.2: before we can compare two groups' scores, we must be sure we are measuring both groups with the same ruler. The four steps check progressively stricter senses in which the two rulers must match, and each step adds its constraints on top of the previous one (see Table K.1).

Table K.1: The Four Levels of Measurement Invariance

Level	**What it checks**	**Ruler analogy**
Configural	The same items form the same underlying dimensions in each group	The ruler measures the same things
Metric	The factor loadings (item–dimension strengths) are equal across groups	The ruler's units are the same size
Scalar	The item thresholds (response-category cutoffs) are equal across groups	The ruler's zero point is in the same place
Strict	The item-specific measurement error is equal across groups	The ruler is equally precise

Note. Each level adds its constraints on top of the previous one. Scalar invariance is the minimum required for valid comparison of group means; strict invariance is rarely achieved and generally not required.

The first step, **configural invariance**, asks whether the ruler measures the same things in each group; that is, whether the same items cluster into the same underlying dimensions (for example, whether the health items still form a distinct "health" factor) in both the original and the translated version. If this basic layout differs across groups, no further comparison is meaningful.

The second step, **metric invariance**, asks whether the ruler's units are the same size across groups. Technically, it tests whether the factor loadings, how strongly each item reflects its underlying dimension, are equivalent across versions (Widaman and Reise, 1997). When they are, the same amount of the underlying trait shows up as the same amount of score movement in each group.

The third step, **scalar invariance**, asks whether the ruler's starting point, i.e., its "zero", sits in the same place for each group. Technically, it tests whether the item thresholds, or in other words, the points on the underlying trait at which a respondent moves from one response category to the next (e.g., from "Very Unsatisfied" to "Very Satisfied"), are equivalent across groups. This is the decisive step for comparing means: if one group needs a higher level of the underlying trait before choosing a higher response category, that group's "zero" is effectively shifted, and a raw difference in average scores could reflect this shift rather than any true difference in the trait (Steinmetz, 2013). Little (2024) likewise identifies scalar invariance as the key threshold for treating observed scores as "comparable" across groups.

The fourth step, **strict invariance**, further requires that the ruler be equally *precise* in each group, that the item-specific measurement error (item uniqueness) be identical across groups. This

strictest level is rarely achieved in practice (Millsap and Olivera-Aguilar, 2012; Widaman and Reise, 1997) and is generally not required for valid comparisons of means (Putnick and Bornstein, 2016; Tse et al., 2023).

At each step, we test whether imposing the new equality constraints meaningfully worsens model fit, using the χ^2 difference test (Sass et al., 2014). A significant result means the constraint does not hold, meaning that the two groups' rulers differ in that respect. Importantly, when full scalar invariance is rejected, comparability is not necessarily lost: a **partial invariance** model may instead be adopted (Xu and Green, 2016; Yoon and Millsap, 2007), in which most items are held equal across groups while a few are allowed to differ (Byrne et al., 1989). In ruler terms, we keep the bulk of the ruler identically calibrated and let a handful of mismatched marks float.

To decide which constraints to relax, we use a **sequential backward specification search** guided by modification indices. They are diagnostic values that flag which single constraint, if freed, would most improve model fit (Yoon and Kim, 2014; Yoon and Millsap, 2007). Concretely, we start from the metric-invariance model, free the one constraint with the largest modification index, re-estimate, and repeat—freeing one constraint at a time—until the $\Delta\chi^2$ comparing the partial scalar model against the metric model is no longer significant, indicating that the remaining constraints are tenable.

Establishing partial scalar invariance offers two advantages. First, it lets us estimate latent mean differences accurately without discarding the non-invariant items altogether (Steenkamp and Baumgartner, 1998; Xu and Green, 2016). Second, the freed parameters are themselves informative: because the intercepts (here, the item thresholds) in MG-CFA partly reflect response styles (Chang et al., 2015; Kline, 2023), examining which items differ across groups and in which direction, can reveal how the two groups use the scale differently (Wu and Huang, 2014).