

Supplementary Information

Audience evaluation of heritage-derived cultural-creative products: a multilevel analysis of rater-level and product-level variance

Zetian Wang, Fangming Zhao, Wentao Tang, Mingfeng Xiao and Lingling Zheng

Supplementary Note 1. Exploratory literature audit

An exploratory audit covered journal studies published from 11 September 2021 to 11 September 2026; screening was recorded on 11 September 2026. The archived search used Crossref REST metadata and DOI checks, public web searches, publisher HTML/PDF and PMC full text, and reference lists and similar titles to find additional candidates, with DOI deduplication. Scopus, Web of Science and CNKI searches were not carried out. Eligible studies were empirical journal articles in heritage, tourism, design or related material aesthetics in which the same participant rated at least two concrete stimuli or presentation conditions on subjective scales. Multiple items about one stimulus did not qualify as multiple stimuli; studies limited to eye tracking, route choice or forced image selection without per-stimulus subjective scales were excluded. The record contains 43 unique DOIs: seven were excluded, four were unclear, and 32 met the repeated-rating criterion; one of these was removed after an Expression of Concern, leaving 31 in the main audit. Coding used the published full text, particularly methods, results and tables. The archive states that there was no independent double screening, but does not record the actual number of coders. Authors' analysis scripts were not available.

Note. The final column records the origin of each conclusion. The L09 variance values are conditional random-intercept components, not variance shares.

ID	First author	Year	Journal	DOI	Title	Category	Analysis estimates participant and stimulus variance components	Conclusion source
L01	Wang et al.	2025	npj Heritage Science	10.1038/s40494-025-01760-5	Design of panoramic virtual museum interactive interface based on entropy weight TOPSIS and PSO-SVR	Stimulus means	No (analysis of stimulus means)	Classification rule
L02	Cai et al.	2025	Scientific Reports	10.1038/s41598-025-94862-w	Research on wheelchair form design based on Kansei engineering and GWO-BP neural network	Stimulus means	No (analysis of stimulus means)	Classification rule
L03	Guan et al.	2025	Scientific Reports	10.1038/s41598-025-18149-w	Using entropy weight method and quality function deployment to develop emotion-driven dashboard interfaces for NEV	Stimulus means	No (analysis of stimulus means)	Classification rule
L04	Yu et al.	2026	Scientific Reports	10.1038/s41598-026-45602-1	Emotion-driven front-end design of NEVs using an improved LSTM with LDA-based emotion mining	Stimulus means	No (analysis of stimulus means)	Classification rule
L05	Kang et al.	2024	Heritage Science	10.1186/s40494-024-01326-x	Assessment of visual quality and social perception of cultural landscapes: application to Anyi traditional villages, China	Stimulus means	No (analysis of stimulus means)	Classification rule
L07	Zou et al.	2025	npj Heritage Science	10.1038/s40494-025-01937-y	Impact of post-earthquake restoration on aesthetic preferences and heritage value of Sparkling Lake in Jiuzhaigou	Single-level individual ratings	No (single-level analysis)	Classification rule
L08	Fei et al.	2025	Scientific Reports	10.1038/s41598-025-00841-6	A comparative study of the perception of traditional villages between different media	Single-level individual ratings	No (single-level analysis)	Classification rule
L09	Campagna et al.	2024	Scientific Reports	10.1038/s41598-024-63925-9	How material sensory properties and individual differences influence the haptic aesthetic appeal of visually presented stimuli	Accounted for dependence among a participant's ratings	Yes — conditional random-intercept variances for participants (0.50) and materials (0.09); shares not reported	Manual check, 2026-09-28
L10	Li et al.	2024	Heritage Science	10.1186/s40494-024-01328-9	Artistic heritage conservation: the relevance and cultural value of Guangzhou clan building paintings to traditional rituals from a kinship perspective through perceptual assessment and data mining	Stimulus means	No (analysis of stimulus means)	Classification rule
L12	Kang et al.	2025	Humanities and Social Sciences Communications	10.1057/s41599-025-05388-5	Design optimization of wood-carved window grilles in historical architectures using stable diffusion model and intuitionistic Fuzzy VIKOR	Stimulus means	No (analysis of stimulus means)	Classification rule
L13	Weihua et al.	2026	Scientific Reports	10.1038/s41598-026-58178-7	A Kansei-mediated protocol for transforming painting palettes into product aesthetics	Accounted for dependence among a participant's ratings	No (paired tests; manual check)	Manual check, 2026-09-28
L14	Cheng et al.	2025	Scientific Reports	10.1038/s41598-025-30524-1	Research on an intelligent fan cultural sustainability design method based on CPO-CNN-LSSVM	Stimulus means	No (analysis of stimulus means)	Classification rule
L15	Gui et al.	2025	Scientific Reports	10.1038/s41598-025-31226-4	Emotional design of new energy vehicle center consoles for elderly users based on entropy weighted TOPSIS and BPNN	Stimulus means	No (analysis of stimulus means)	Classification rule
L16	Huo et al.	2023	Humanities and Social Sciences Communications	10.1057/s41599-023-01761-4	A user experience map design method based on emotional quantification of in-vehicle HMI	Accounted for dependence among a participant's ratings	No (paired tests; manual check)	Manual check, 2026-09-28
L21	Wu et al.	2023	Humanities and Social Sciences Communications	10.1057/s41599-023-02399-y	Visual attention predictive model of built colonial heritage based on visual behaviour and subjective evaluation	Single-level individual ratings	No (single-level analysis)	Classification rule

ID	First author	Year	Journal	DOI	Title	Category	Analysis estimates participant and stimulus variance components	Conclusion source
L22	Xu et al.	2024	Heritage Science	10.1186/s40494-024-01397-w	Research on visual experience evaluation of fortress heritage landscape by integrating SBE–SD method and eye movement analysis	Stimulus means	No (analysis of stimulus means)	Classification rule
L24	Sun et al.	2025	npj Heritage Science	10.1038/s40494-025-01650-w	Effects of reflecting water on viewers' visual attention, spatial perception, and emotional responses: a case study in Chinese heritage garden	Accounted for dependence among a participant's ratings	No (repeated-measures comparisons; screening record)	Classification rule
L26	Cheng et al.	2026	Buildings	10.3390/buildings16010190	Visual Preferences and Place Attachment Construction of Generation Z Tourists at Sacred Heritage Landscapes Based on Eye-Tracking and Questionnaire	Accounted for dependence among a participant's ratings	No (repeated-measures comparisons; screening record)	Classification rule
L27	Ye et al.	2024	Sensors	10.3390/s24144459	Predicting Emotional Experiences through Eye-Tracking: A Study of Tourists' Responses to Traditional Village Landscapes	Other	No (person-level data)	Classification rule
L28	Zhu et al.	2025	Symmetry	10.3390/sym17020306	Research on Evaluation Methods of Complex Product Design Based on Hybrid Kansei Engineering Modeling	Stimulus means	No (analysis of stimulus means)	Classification rule
L29	Gong et al.	2022	Sustainability	10.3390/su141710556	Using Kansei Engineering for the Design Thinking Framework: Bamboo Pen Holder Product Design	Stimulus means	No (analysis of stimulus means)	Classification rule
L30	Du et al.	2024	Sustainability	10.3390/su16208837	A Product's Kansei Appearance Design Method Based on Conditional-Controlled AI Image Generation	Stimulus means	No (analysis of stimulus means)	Classification rule
L31	Gao et al.	2024	Sustainability	10.3390/su16135546	Kansei Drives Sustainable Material Innovation—An Approach to Enhance the Added Value of Biomass Materials	Stimulus means	No (analysis of stimulus means)	Classification rule
L32	Zhang et al.	2023	Applied Sciences	10.3390/app13179577	Product Styling Cognition Based on Kansei Engineering Theory and Implicit Measurement	Stimulus means	No (analysis of stimulus means)	Classification rule
L34	Huang et al.	2025	BioResources	10.15376/biores.20.3.7107-7133	Fusion of rough set theory, genetic algorithm-backpropagation neural networks and Shapley additive explanations for the design of bamboo furniture	Stimulus means	No (analysis of stimulus means)	Classification rule
L35	Liu et al.	2025	BioResources	10.15376/biores.20.4.8611-8631	Generative design of bamboo furniture combining game theory and AI-generated content	Stimulus means	No (analysis of stimulus means)	Classification rule
L36	Hao et al.	2025	BioResources	10.15376/biores.20.2.2711-2727	Children's preferences for the styling of consultation room furniture based on scenic beauty estimation and Kansei engineering.	Stimulus means	No (analysis of stimulus means)	Classification rule
L39	Gao et al.	2024	Complexity	10.1155/2024/5524490	Cultural Echoes in Modern Design: Assessing Young Consumers' Perceptions of Traditional Bamboo Weaving Patterns	Other	No (unit of analysis not stated; no multilevel model reported)	Classification rule
L40	Wang et al.	2024	Heliyon	10.1016/j.heliyon.2024.e28564	Correlation research on visual behavior and cognition of impression against localized church facades in China	Stimulus means	No (analysis of stimulus means)	Classification rule
L41	Xing et al.	2025	Frontiers of Architectural Research	10.1016/j.foar.2024.10.003	Cognitive preferences for architectural renovation strategies in traditional villages combining subjective evaluation and eye tracking	Stimulus means	No (analysis of stimulus means)	Classification rule
L42	Calderón-Fajardo et al.	2024	Tourism & Management Studies	10.18089/tms.20240305	Neurotourism Insights: Eye Tracking and Galvanic Analysis of Tourism Destination Brand Logos and AI Visuals	Accounted for dependence among a participant's ratings	No (paired tests; manual check)	Manual check, 2026-09-28

Studies not included in the main audit (12 records)

ID	DOI	Title	Status	Recorded reason (English)
L06	10.1038/s41598-024-85009-4	The influence of architectural heritage and tourists' positive emotions on behavioral intentions using eye-tracking study	Unclear	It was unclear whether the 214 questionnaire respondents in the OLS analysis each rated multiple stimuli; the multi-image eye-tracking experiment does not establish that structure for the questionnaire regression.
L11	10.1186/s40494-024-01456-2	Cultural rituality and heritage revitalization values of ancestral temple architecture painting art from the perspective of relational sociology theory	Excluded	The article presents cultural interpretation and relational-sociology discussion; quantitative audience perception ratings were proposed for future work rather than collected in a repeated-stimulus study.
L17	10.1038/s41598-024-82281-2	The application of artificial intelligence-assisted technology in cultural and creative product design	Unclear	Expert and user scores, means and ANOVA were reported, but participant counts and assignment of designs to individual raters were insufficient to verify repeated ratings.
L18	10.1038/s40494-025-01662-6	Developing a cultural sustainability assessment framework for environmental facilities in urban communities	Excluded	Experts chiefly rated the importance of indicators to derive weights; a multi-item questionnaire does not establish ratings of multiple concrete stimuli by each expert.
L19	10.1057/s41599-026-07116-z	Evaluating and translating cultural symbol design: An AHP-TOPSIS evaluation framework based on the Chaoshan five-element Gable Walls	Unclear	It was not confirmed that each user rated both tea sets; AHP pairwise comparisons of criteria do not establish repeated product ratings.
L20	10.1057/s41599-026-06869-x	Risk decision-making behavior in heritage tourism destinations based on eye-tracking experiments in	Excluded	The eye-tracking and risk-route-choice study was later aggregated by streetscape and did not meet the

		Gulangyu, China		subjective repeated-rating criterion.
L23	10.1057/s41599-022-01411-1	What do we visually focus on in a World Heritage Site? A case study in the Historic Centre of Prague	Excluded	The eye-tracking study used multiple Prague heritage images, but per-stimulus subjective ratings were not confirmed.
L25	10.3390/buildings12071058	Investigating the Visual Behavior Characteristics of Architectural Heritage Using Eye-Tracking	Excluded	The same participant viewed multiple architectural images, but the main analysis was eye tracking and clear evidence of per-stimulus subjective ratings was not found.
L33	10.3390/buildings14113529	A Study on the Visual Perception of Cultural Value Characteristics of Traditional Southern Fujian Architecture Based on Eye Tracking	Excluded	The main analysis concerned eye-tracking responses to architectural cultural values; per-stimulus subjective scale ratings were not confirmed.
L37	10.15376/biores.20.4.10082-10105	Redesign of the Ming-style ‘Warped Table’ based on Kansei needs of contemporary users using an SD–FA–QFD method	Unclear	The paper listed 16 images and means but described rating one sample per questionnaire; it was unclear whether a person actually rated multiple designs.
L38	10.1371/journal.pone.0318715	Research on the aesthetic sensitivity evaluation of tourism mascots based on semantic differential method	Removed after Expression of Concern	Eligible design, removed after an Expression of Concern dated 17 February 2026 (DOI 10.1371/journal.pone.0343103).
L43	10.18089/tms.20250206	Exploring Visual Engagement in UNESCO Creative Cities: An Eye-Tracking Analysis of Social Media Images	Excluded	Participants selected a preferred image in each group and gave oral reasons; eye-tracking and descriptive analyses did not provide per-stimulus scale ratings.

Supplementary Note 2 Questionnaire items and stimulus-board captions

All scale items used seven points. Agreement items were anchored 1 = strongly disagree (完全不同意) and 7 = strongly agree (完全同意). Semantic-differential pairs were scored 1 at the left pole and 7 at the right pole. Items were presented in Chinese; English translations are given for reference.

Descriptive perceptions (semantic differential)

Construct	Left pole (1)	Right pole (7)
Modernity	传统的 Traditional	现代的 Modern
	古朴的 Vintage	时尚的 Fashionable
	复古的 Retro	潮流的 Trendy
Refinement	朴拙的 Rustic	精致的 Refined
	粗糙的 Coarse	细腻的 Delicate
	平凡的 Ordinary	高级的 Premium
Dynamism	静态的 Static	动感的 Dynamic
	沉稳的 Stable	灵动的 Lively
	严肃的 Serious	活泼的 Playful
Distinctiveness	普通的 Common	独特的 Unique
	浅显的 Superficial	厚重的 Profound
	抽象的 Abstract	具象的 Concrete

Evaluative judgements (agreement items)

Construct	Item (Chinese)	Item (English)
Design perception	该产品的色彩搭配（黑/红/黄）能够传达民族特色	The product's colour scheme (black/red/yellow) conveys ethnic character.
	该产品的造型让我能联想到鸡枞菌或陀螺	The product's form reminds me of the jizong mushroom or the spinning top.
	该产品的纹样和装饰元素丰富有辨识度	The product's patterns and decorative elements are rich and recognisable.

Construct	Item (Chinese)	Item (English)
	该产品的工艺质感（仿掐丝珐琅）令人印象深刻	The product's craft texture (imitation cloisonné) is impressive.
	该产品的文化符号传达清晰、易于理解	The product's cultural symbols are conveyed clearly and are easy to understand.
	该产品整体设计协调统一	The product's overall design is harmonious and unified.
Perceived cultural authenticity	该产品准确反映了佤族文化的精髓	The product accurately reflects the essence of Wa culture.
	该产品让我感受到云南山林文化的氛围	The product conveys the atmosphere of Yunnan's mountain-forest culture.
	该产品在文化呈现上是可信的、不浮夸的	The product presents its culture credibly and without exaggeration.
Aesthetic evaluation	我认为该产品在视觉上很吸引我	I find the product visually appealing.
	我会向朋友推荐这款产品	I would recommend this product to friends.
	总体而言，我喜欢这款产品	Overall, I like this product.

Respondent-level items

Measure	Item (Chinese)	Item (English)
Cultural familiarity	我对佤族文化比较了解	I know a fair amount about Wa culture.
	我对云南少数民族非物质文化遗产比较了解	I know a fair amount about the intangible cultural heritage of Yunnan's ethnic minorities.
	我对带有民族风/国潮元素的产品有浓厚兴趣	I am very interested in products with ethnic or guochao elements.
Visit	我曾经到过云南少数民族聚居区	I have visited ethnic-minority areas of Yunnan.

Background items recorded age group, gender, ethnicity, education, occupation, monthly income, region of residence and the number of cultural-creative products purchased in the preceding year. Two attention-check items, one direct (correct response 7) and one reverse-keyed (correct response 1), were included as quality indicators.

Stimulus-board captions

Each board carried the header 设计提案 / Design Proposal, the product name and the caption below. Line drawings were labelled 设计平面图 (design drawing).

Product	Caption (Chinese)	Caption (English)
P1 Brooch	胸针作为单个装饰，使用“动态”陀螺，整体更丰富。	As a single ornament, the brooch uses the "dynamic" top for a richer overall effect.
P2 Earrings	耳饰不易太过复杂，“静态”更显大气简约。黑色耳环，镶嵌红、黄两颗钻，呼应佤族代表色。	Earrings should not be too complex; the "static" top looks more elegant and simple. A black hoop set with a red and a yellow stone echoes the representative colours of the Wa.
P3 Necklace pendant	以“动态”为主视觉，“静态”也可使用为项链吊坠，项链的形式可以有多种变化。	The "dynamic" top is the main visual; the "static" top can also serve as the pendant, and the necklace can take many forms.
P4 Ring	戒指主体以黑色，较宽，鸡纵陀螺图形与红黄钻为辅，图形搭配戒指需更小，但也更显精致。与红黄较细戒指以搭配，彰显民族风格也兼具现代时尚表达。	The ring body is black and fairly wide, accented by the jicong top motif and red and yellow stones; the motif is scaled down for the ring and appears more refined. Paired with thinner red and yellow bands, it shows ethnic style together with a contemporary fashion expression.
P5 Bookmark	应用：书签。在陀螺视觉基础上加上佤族牛头元素。	Application: bookmark. A Wa buffalo-head element is added to the top visual.
P6 Plush pendant	应用：毛绒玩偶挂件。利用陀螺下方圆柱特点，手握扭转可使玩偶的手动起来。	Application: plush pendant. Using the cylindrical handle beneath the top, twisting it by hand makes the figure's arms move.

Supplementary Note 3 Simulation of pooled and decomposed estimators in repeated-stimulus rating designs

Purpose

This note reports a Monte Carlo simulation of designs in which each respondent rates m stimuli. It asks two questions. When the within-respondent association between a predictor and an outcome is zero and the between-respondent association is not, how often do estimators that return a single slope report a significant positive association? And under the same conditions, does the within-between decomposition recover the within-respondent association? The simulation concerns analyses of individual ratings; analyses of stimulus means are outside its scope.

Data-generating process

Ratings $j = 1, \dots, m$ are nested in respondents $i = 1, \dots, N$. The predictor is $x_{ij} = \bar{x}_i + w_{ij}$, with respondent means $\bar{x}_i \sim N(0, \sigma^2_b)$ and deviations $w_{ij} \sim N(0, \sigma^2_w)$. The outcome is

$$y_{ij} = \beta_w (x_{ij} - \bar{x}_i) + \beta_b \bar{x}_i + u_i + e_{ij},$$

with respondent intercepts $u_i \sim N(0, \tau^2)$ and residuals $e_{ij} \sim N(0, \sigma^2_e)$. The predictor ICC, $ICC_x = \sigma^2_b / (\sigma^2_b + \sigma^2_w)$, and the residual outcome ICC, $ICC_y = \tau^2 / (\tau^2 + \sigma^2_e)$, were set with total variances fixed at 1. The between-respondent association was $\beta_b = 0.5$.

The factorial grid crossed $ICC_x \in \{0.2, 0.4, 0.6, 0.8\}$, $ICC_y \in \{0.2, 0.4, 0.6, 0.8\}$, $m \in \{2, 3, 5, 10\}$, $N \in \{100, 228, 500\}$ and $\beta_w \in \{0, 0.125, 0.25, 0.5\}$: 768 cells, each with 1,000 replicates. Six anchor cells reproduced the present study for each of the three associations among the evaluative judgements (design perception $\rightarrow$ perceived cultural authenticity, perceived cultural authenticity $\rightarrow$ aesthetic evaluation, design perception $\rightarrow$ aesthetic evaluation), with $m = 2$, $N = 228$, $\beta_w = 0$ and $ICC_x = 0.817$, the respondent-level share of the predictor in the frozen data. The outcome ICC was set in two ways: by inserting the observed ICC as the residual ICC, and by calibrating the generating process so that the marginal outcome ICC matched the observed value. Each cell used the seed base + cell identifier.

Estimators

Four estimators were fitted to each replicate:

- A, ordinary least squares with a single slope for x ;
- B, the same model with respondent-clustered standard errors;
- C, a random-intercept model with a single slope for x ;
- D, a within-between model with the respondent mean $\bar{x}_i$ and the deviation $x_{ij} - \bar{x}_i$ as separate predictors and a random intercept for respondents.

For D, the coefficient of the deviation estimates β_w , and the difference between the two coefficients tests $\beta_w = \beta_b$. For A–C, the single slope was read as an estimate of β_w , as it is when a pooled association is interpreted as describing how the qualities of a stimulus relate for the person who rates it.

The outcome measures were: the proportion of replicates with a two-sided $P < 0.05$ and a positive estimate (nominal rate 2.5% when $\beta_w = 0$); the two-sided rejection rate; bias and root mean squared error; 95% interval coverage; for D, the rate at which $\beta_w = \beta_b$ was rejected; and convergence. Models were fitted in R; package versions are recorded in the archived renv lock file.

Results

The single-slope estimators estimate a weighted combination of the within- and between-respondent associations. With $\beta_w = 0$, the mean estimate from A was close to $\beta_b \times ICC_x$: at $m = 2$ and $N = 228$ it rose from 0.10 at $ICC_x = 0.2$ to 0.40 at $ICC_x = 0.8$, whatever the outcome ICC.

Across the 192 cells with $\beta_w = 0$, a significant positive single slope appeared in a median of 100% of replicates for A (range 26.8% to 100%), 99.8% for B (24.7% to 100%) and 68.5% for C (3.2% to 100%). The rate rose with ICC_x and with N . For C it fell as m increased, since the random-intercept slope moves towards the within-respondent slope as each respondent contributes more ratings. At $m = 2$ and $N = 228$, the rate for A was 53.0% to 57.3% at $ICC_x = 0.2$ and at least 96.7% from $ICC_x = 0.4$ upwards; for C it was 10.9% to 40.0% at $ICC_x = 0.2$ and 99.4% to 100% at $ICC_x = 0.8$ (Supplementary Fig. 1a).

Estimator D held its nominal properties throughout. With $\beta_w = 0$, a significant positive within-respondent slope appeared in a median of 2.5% of replicates (1.2% to 4.1%), and the two-sided rejection rate averaged 5.0% (2.9% to 7.0%). Coverage of the 95% interval for β_w lay between 91.9% and 97.3% across all 768 cells (mean 94.9%). With $\beta_w = 0$, the difference between the within- and between-respondent coefficients was detected in a median of 99.2% of replicates, from 21.5% in cells with a low predictor ICC and few respondents. All four estimators converged in every replicate.

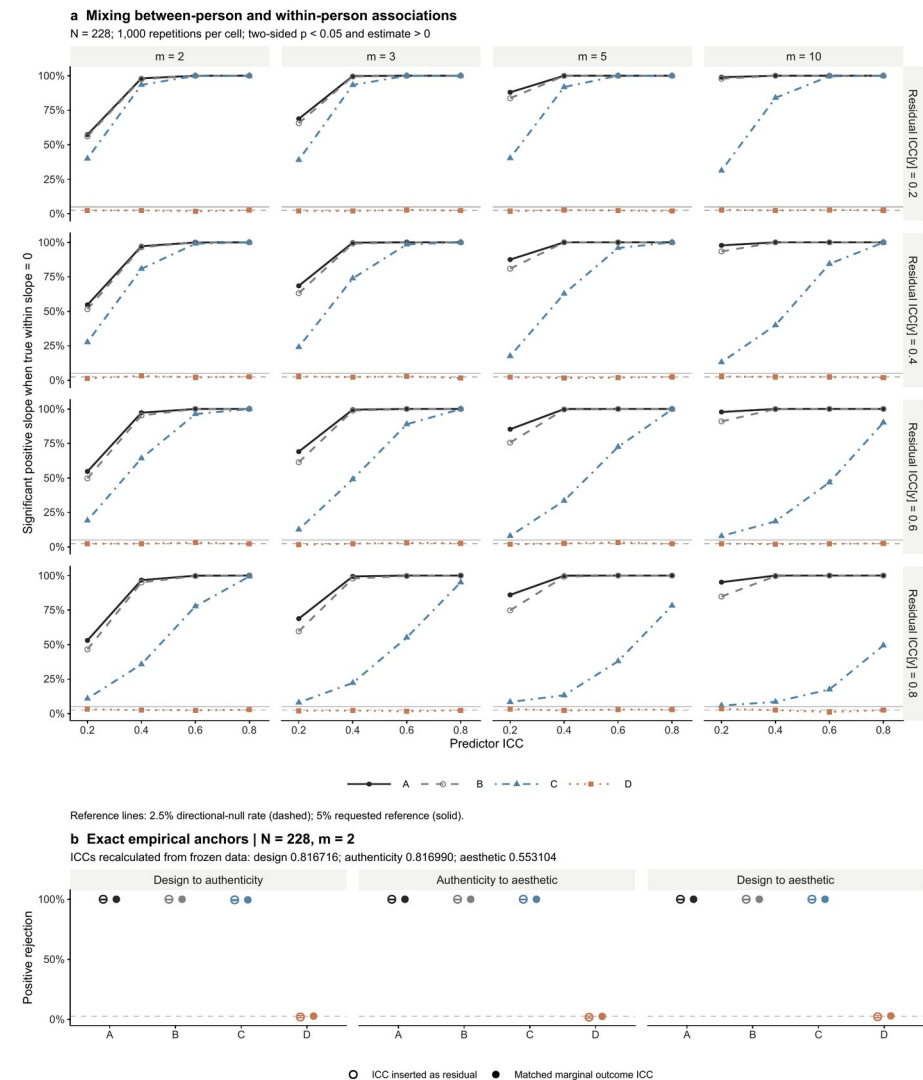
At the six anchor cells, A and B returned a significant positive single slope in 100% of replicates for all three associations, and C in 99.5% to 100%. D returned a significant positive within-respondent slope in 1.8% to 2.9% of replicates, with two-sided rejection rates of 3.8% to 6.0% and coverage of 94.0% to 96.2%, and detected the within-between difference in 93.9% to 99.7% of replicates (Supplementary Fig. 1b). The two outcome-ICC calibrations gave the same pattern.

Interpretation

The single slopes of A–C are sound estimates of their own pooled target. They cannot be read as within-respondent associations when the predictor varies mainly between respondents, as the evaluative judgements in the present study do. Under the settings of this study, a within-respondent association of zero appeared as a significant positive single slope in nearly every replicate, whereas the within-between decomposition recovered it at the nominal rate and detected the difference between the two levels in most replicates.

Simulation scripts, seeds, session information and the renv lock file are archived with the analysis scripts. Bias and coverage surfaces for all cells are given in Supplementary Data 1.

Supplementary Fig. 1



Supplementary Fig. 1 Significant positive single slopes when the within-respondent association is zero. a Proportion of 1,000 replicates in which each estimator returned a slope with two-sided $P < 0.05$ and a positive sign, for $N = 228$, by predictor ICC (horizontal axis), ratings per respondent m (columns) and residual outcome ICC (rows). A, ordinary least squares; B, ordinary least squares with respondent-clustered standard errors; C, random-intercept model with a single slope; D, within-between model (within-respondent coefficient). The dashed line marks the 2.5% nominal rate for a directional result and the solid line the 5% two-sided level. b The same proportion at the six anchor cells reproducing the present study ($m = 2$, $N = 228$, $ICC_x = 0.817$) for the three associations among the evaluative judgements; open symbols insert the observed outcome ICC as the residual ICC, filled symbols match the observed marginal outcome ICC.

Supplementary Table 1. Exclusion criteria and sequential screening flow

Note. The sequential exclusion order is the prespecified protocol order: under 18 years, missing required responses, duration <240 seconds, then at least 35 identical responses. Individual hits and intersections overlap and are not added to obtain the sequential total. Only nonzero intersections are shown.

Section	Criterion or intersection	Respondents	Evaluations	Remaining respondents	Remaining evaluations
A: individual criterion hits	underage	15	30	—	—
A: individual criterion hits	missing_required	0	0	—	—
A: individual criterion hits	duration_lt_240	190	380	—	—
A: individual criterion hits	straight_lining_ge_35	19	38	—	—
B: sequential flow	none	0	—	432	864
B: sequential flow	underage	15	—	417	834
B: sequential flow	missing_required	0	—	417	834
B: sequential flow	duration_lt_240	183	—	234	468
B: sequential flow	straight_lining_ge_35	6	—	228	456
C: nonzero intersections	underage+duration_lt_240	7	—	—	—
C: nonzero intersections	duration_lt_240+straight_lining_ge_35	13	—	—	—

Supplementary Table 2. Screening sensitivity of crossed variance models

Note. Panel A reports REML crossed random-intercept models. Panel B uses the 127-rater sample passing both attention checks; the family difference has a 1,000-resample respondent-cluster percentile interval. M1/M2 entries are proportional reductions of rater-level variance. Profile intervals are on the product-variance scale. A dash means not applicable.

Panel	Scenario or model	Construct	Raters	Evaluations	Evaluations per product (P1–P6)	Rater share (%)	Product share (%)	Residual share (%)	Product variance profile 95% CI	Boundary LRT P	Singular	PRV rater (%) or family difference (pp)	Family difference 95% CI (pp)
A: variance decomposition	baseline_228	modernity_mean	228	456	98; 87; 59; 97; 55; 60	54.06	1.71	44.23	[0.0000, 0.1824]	0.033	FALSE	—	—
A: variance decomposition	baseline_228	refinement_mean	228	456	98; 87; 59; 97; 55; 60	67.23	1.83	30.94	[0.0034, 0.1449]	0.00296	FALSE	—	—
A: variance decomposition	baseline_228	dynamism_mean	228	456	98; 87; 59; 97; 55; 60	42.47	4.87	52.66	[0.0207, 0.4062]	2.44e-05	FALSE	—	—
A: variance decomposition	baseline_228	distinctiveness_mean	228	456	98; 87; 59; 97; 55; 60	58.66	0.00	41.34	[0.0000, 0.0322]	0.5	TRUE	—	—
A: variance decomposition	baseline_228	design_mean	228	456	98; 87; 59; 97; 55; 60	81.62	0.65	17.73	[0.0000, 0.0553]	0.0384	FALSE	—	—
A: variance decomposition	baseline_228	authenticity_mean	228	456	98; 87; 59; 97; 55; 60	81.83	0.35	17.83	[0.0000, 0.0357]	0.127	FALSE	—	—
A: variance decomposition	baseline_228	aesthetic_mean	228	456	98; 87; 59; 97; 55; 60	56.12	0.98	42.90	[0.0000, 0.1103]	0.0623	FALSE	—	—
A: variance decomposition	both_attention	modernity_mean	127	254	55; 47; 36; 50; 34; 32	54.72	3.13	42.16	[0.0000, 0.3248]	0.0189	FALSE	—	—
A: variance decomposition	both_attention	refinement_mean	127	254	55; 47; 36; 50; 34; 32	70.74	1.42	27.83	[0.0000, 0.1430]	0.0699	FALSE	—	—
A: variance decomposition	both_attention	dynamism_mean	127	254	55; 47; 36; 50; 34; 32	34.50	9.83	55.67	[0.0443, 0.8194]	1.65e-05	FALSE	—	—
A: variance decomposition	both_attention	distinctiveness_mean	127	254	55; 47; 36; 50; 34; 32	53.48	0.00	46.52	[0.0000, 0.0500]	0.5	TRUE	—	—
A: variance decomposition	both_attention	design_mean	127	254	55; 47; 36; 50; 34; 32	82.35	0.32	17.34	[0.0000, 0.0524]	0.278	FALSE	—	—
A: variance decomposition	both_attention	authenticity_mean	127	254	55; 47; 36; 50; 34; 32	84.58	0.09	15.33	[0.0000, 0.0298]	0.407	FALSE	—	—

Panel	Scenario or model	Construct	Raters	Evaluations	Evaluations per product (P1–P6)	Rater share (%)	Product share (%)	Residual share (%)	Product variance profile 95% CI	Boundary LRT P	Singular	PRV rater (%) or family difference (pp)	Family difference 95% CI (pp)
A: variance decomposition	both_attention	aesthetic_mean	127	254	55; 47; 36; 50; 34; 32	52.13	0.55	47.32	[0.0000, 0.1038]	0.282	FALSE	—	—
A: variance decomposition	direct_attention	modernity_mean	144	288	64; 54; 40; 57; 36; 37	53.99	2.21	43.81	[0.0000, 0.2678]	0.0466	FALSE	—	—
A: variance decomposition	direct_attention	refinement_mean	144	288	64; 54; 40; 57; 36; 37	74.20	1.11	24.68	[0.0000, 0.1257]	0.0737	FALSE	—	—
A: variance decomposition	direct_attention	dynamism_mean	144	288	64; 54; 40; 57; 36; 37	39.70	7.61	52.69	[0.0341, 0.6705]	3.48e-05	FALSE	—	—
A: variance decomposition	direct_attention	distinctiveness_mean	144	288	64; 54; 40; 57; 36; 37	54.16	0.00	45.84	[0.0000, 0.0626]	0.5	TRUE	—	—
A: variance decomposition	direct_attention	design_mean	144	288	64; 54; 40; 57; 36; 37	84.76	0.38	14.86	[0.0000, 0.0527]	0.202	FALSE	—	—
A: variance decomposition	direct_attention	authenticity_mean	144	288	64; 54; 40; 57; 36; 37	87.21	0.04	12.75	[0.0000, 0.0226]	0.446	FALSE	—	—
A: variance decomposition	direct_attention	aesthetic_mean	144	288	64; 54; 40; 57; 36; 37	58.91	0.85	40.24	[0.0000, 0.1180]	0.145	FALSE	—	—
A: variance decomposition	reverse_attention	modernity_mean	132	264	60; 51; 36; 50; 34; 33	53.98	2.70	43.33	[0.0000, 0.2884]	0.0329	FALSE	—	—
A: variance decomposition	reverse_attention	refinement_mean	132	264	60; 51; 36; 50; 34; 33	71.61	1.31	27.08	[0.0000, 0.1328]	0.0711	FALSE	—	—
A: variance decomposition	reverse_attention	dynamism_mean	132	264	60; 51; 36; 50; 34; 33	37.72	9.23	53.05	[0.0426, 0.7862]	1.51e-05	FALSE	—	—
A: variance decomposition	reverse_attention	distinctiveness_mean	132	264	60; 51; 36; 50; 34; 33	55.28	0.00	44.72	[0.0000, 0.0481]	0.5	TRUE	—	—
A: variance decomposition	reverse_attention	design_mean	132	264	60; 51; 36; 50; 34; 33	82.53	0.45	17.02	[0.0000, 0.0572]	0.191	FALSE	—	—
A: variance decomposition	reverse_attention	authenticity_mean	132	264	60; 51; 36; 50; 34; 33	84.53	0.16	15.30	[0.0000, 0.0338]	0.339	FALSE	—	—
A: variance decomposition	reverse_attention	aesthetic_mean	132	264	60; 51; 36; 50; 34; 33	53.10	0.72	46.18	[0.0000, 0.1097]	0.214	FALSE	—	—
A: variance decomposition	duration_ge_255	modernity_mean	206	412	91; 76; 55; 87; 51; 52	48.43	1.54	50.03	[0.0000, 0.1722]	0.0735	FALSE	—	—
A: variance decomposition	duration_ge_255	refinement_mean	206	412	91; 76; 55; 87; 51; 52	65.15	2.30	32.55	[0.0042, 0.1756]	0.00291	FALSE	—	—
A: variance decomposition	duration_ge_255	dynamism_mean	206	412	91; 76; 55; 87; 51; 52	41.62	5.10	53.28	[0.0200, 0.4103]	4.44e-05	FALSE	—	—
A: variance decomposition	duration_ge_255	distinctiveness_mean	206	412	91; 76; 55; 87; 51; 52	55.50	0.14	44.36	[0.0000, 0.0455]	0.416	FALSE	—	—
A: variance decomposition	duration_ge_255	design_mean	206	412	91; 76; 55; 87; 51; 52	80.73	0.57	18.71	[0.0000, 0.0529]	0.084	FALSE	—	—
A: variance decomposition	duration_ge_255	authenticity_mean	206	412	91; 76; 55; 87; 51; 52	81.10	0.32	18.59	[0.0000, 0.0367]	0.189	FALSE	—	—
A: variance decomposition	duration_ge_255	aesthetic_mean	206	412	91; 76; 55; 87; 51; 52	58.85	0.85	40.30	[0.0000, 0.0998]	0.0895	FALSE	—	—
A: variance decomposition	both_attention_duration_ge_255	modernity_mean	118	236	53; 42; 33; 47; 34; 27	49.04	2.83	48.13	[0.0000, 0.2883]	0.0468	FALSE	—	—
A: variance decomposition	both_attention_duration_ge_255	refinement_mean	118	236	53; 42; 33; 47; 34; 27	71.24	2.20	26.56	[0.0000, 0.1985]	0.0357	FALSE	—	—
A: variance decomposition	both_attention_duration_ge_255	dynamism_mean	118	236	53; 42; 33; 47; 34; 27	36.14	9.50	54.36	[0.0382, 0.7602]	4.36e-05	FALSE	—	—
A: variance decomposition	both_attention_duration_ge_255	distinctiveness_mean	118	236	53; 42; 33; 47; 34; 27	48.78	0.00	51.22	[0.0000, 0.0584]	0.5	TRUE	—	—
A: variance decomposition	both_attention_duration_ge_255	design_mean	118	236	53; 42; 33; 47; 34; 27	80.37	0.36	19.27	[0.0000, 0.0605]	0.295	FALSE	—	—
A: variance decomposition	both_attention_duration_ge_255	authenticity_mean	118	236	53; 42; 33; 47; 34; 27	83.56	0.18	16.26	[0.0000, 0.0361]	0.348	FALSE	—	—
A: variance decomposition	both_attention_duration_ge_255	aesthetic_mean	118	236	53; 42; 33; 47; 34; 27	53.21	0.05	46.73	[0.0000, 0.0758]	0.478	FALSE	—	—
A: variance decomposition	relaxed_duration_ge_177	modernity_mean	307	614	131; 118; 82; 132; 73; 78	55.99	2.53	41.47	[0.0069, 0.2170]	0.00101	FALSE	—	—
A: variance decomposition	relaxed_duration_ge_177	refinement_mean	307	614	131; 118; 82; 132; 73; 78	69.82	2.40	27.77	[0.0085, 0.1734]	3.88e-05	FALSE	—	—

Panel	Scenario or model	Construct	Raters	Evaluations	Evaluations per product (P1–P6)	Rater share (%)	Product share (%)	Residual share (%)	Product variance profile 95% CI	Boundary LRT P	Singular	PRV rater (%) or family difference (pp)	Family difference 95% CI (pp)
A: variance decomposition	relaxed_duration_ge_177	dynamism_mean	307	614	131; 118; 82; 132; 73; 78	51.55	4.25	44.20	[0.0196, 0.3265]	8.57e-07	FALSE	—	—
A: variance decomposition	relaxed_duration_ge_177	distinctiveness_mean	307	614	131; 118; 82; 132; 73; 78	63.03	0.44	36.53	[0.0000, 0.0467]	0.182	FALSE	—	—
A: variance decomposition	relaxed_duration_ge_177	design_mean	307	614	131; 118; 82; 132; 73; 78	82.73	0.48	16.79	[0.0000, 0.0425]	0.0484	FALSE	—	—
A: variance decomposition	relaxed_duration_ge_177	authenticity_mean	307	614	131; 118; 82; 132; 73; 78	79.41	0.30	20.29	[0.0000, 0.0313]	0.113	FALSE	—	—
A: variance decomposition	relaxed_duration_ge_177	aesthetic_mean	307	614	131; 118; 82; 132; 73; 78	61.19	0.64	38.17	[0.0000, 0.0738]	0.0836	FALSE	—	—
B: both attention checks	family difference	descriptive – evaluative	127	254	—	—	—	—	—	—	—	3.27	[1.15, 6.40]
B: both attention checks	M1	design_mean	127	254	—	—	—	—	—	0.0706	FALSE	1.96	—
B: both attention checks	M2	design_mean	127	254	—	—	—	—	—	0.381	FALSE	2.00	—
B: both attention checks	M1	authenticity_mean	127	254	—	—	—	—	—	6.7e-05	FALSE	12.07	—
B: both attention checks	M2	authenticity_mean	127	254	—	—	—	—	—	0.656	FALSE	10.94	—
B: both attention checks	M1	aesthetic_mean	127	254	—	—	—	—	—	1.28e-05	FALSE	19.71	—
B: both attention checks	M2	aesthetic_mean	127	254	—	—	—	—	—	0.335	FALSE	19.66	—

Supplementary Table 3. Cultural familiarity components and product interactions

Note. The composite is the mean of three familiarity items; knowledge is the mean of items 1–2, and interest is item 3. Coefficients and intervals come from crossed random-intercept models. PRV denotes proportional reduction in respondent-level variance. LRT compares each familiarity model with M0. Interaction tests have five degrees of freedom. Predicted means are at composite familiarity ± 1 SD, with rank 1 highest. A dash means not applicable.

Section	Outcome	Model / level	Term / product	β (unstandardised)	SE	95% CI	Standardised β	PRV rater (%)	LRT χ^2	df	LRT P	Predicted mean	Rank	Rank order / Spearman ρ	Knowledge-item Pearson r	Spearman–Brown reliability	Direction
A: familiarity components	design_mean	composite	familiarity_c	0.235	0.058	[0.122, 0.349]	0.249	7.2	16.128	1	5.92e-05	—	—	—	—	—	positive
A: familiarity components	design_mean	knowledge	knowledge_c	0.167	0.054	[0.061, 0.272]	0.192	4.1	9.514	1	0.00204	—	—	—	—	—	positive
A: familiarity components	design_mean	interest	interest_c	0.205	0.047	[0.113, 0.297]	0.265	8.2	18.467	1	1.73e-05	—	—	—	—	—	positive
A: familiarity components	design_mean	both	knowledge_c	0.067	0.061	[–0.052, 0.187]	0.078	8.3	19.704	2	5.26e-05	—	—	—	—	—	positive
A: familiarity components	design_mean	both	interest_c	0.174	0.054	[0.068, 0.281]	0.226	8.3	19.704	2	5.26e-05	—	—	—	—	—	positive
A: familiarity components	authenticity_mean	composite	familiarity_c	0.322	0.054	[0.216, 0.429]	0.350	14.5	32.963	1	9.39e-09	—	—	—	—	—	positive
A: familiarity components	authenticity_mean	knowledge	knowledge_c	0.242	0.051	[0.142, 0.341]	0.287	9.6	21.649	1	3.27e-06	—	—	—	—	—	positive
A: familiarity components	authenticity_mean	interest	interest_c	0.259	0.044	[0.172, 0.346]	0.344	14.1	31.838	1	1.68e-08	—	—	—	—	—	positive
A: familiarity components	authenticity_mean	both	knowledge_c	0.128	0.057	[0.015, 0.240]	0.152	15.7	36.814	2	1.01e-08	—	—	—	—	—	positive
A: familiarity components	authenticity_mean	both	interest_c	0.201	0.051	[0.101, 0.301]	0.268	15.7	36.814	2	1.01e-08	—	—	—	—	—	positive
A: familiarity components	aesthetic_mean	composite	familiarity_c	0.394	0.058	[0.280, 0.509]	0.362	22.8	41.766	1	1.03e-10	—	—	—	—	—	positive
A: familiarity components	aesthetic_mean	knowledge	knowledge_c	0.309	0.055	[0.201, 0.416]	0.310	16.6	29.870	1	4.62e-08	—	—	—	—	—	positive
A: familiarity components	aesthetic_mean	interest	interest_c	0.296	0.048	[0.201, 0.391]	0.333	19.1	34.788	1	3.68e-09	—	—	—	—	—	positive
A: familiarity components	aesthetic_mean	both	knowledge_c	0.189	0.062	[0.068, 0.310]	0.190	23.5	44.104	2	2.65e-10	—	—	—	—	—	positive

Section	Outcome	Model / level	Term / product	β (unstandardised)	SE	95% CI	Standardised β	PRV rater (%)	LRT χ^2	df	LRT P	Predicted mean	Rank	Rank order / Spearman ρ	Knowledge-item Pearson r	Spearman-Brown reliability	Direction
A: familiarity components	aesthetic_mean	both	interest_c	0.210	0.055	[0.102, 0.318]	0.236	23.5	44.104	2	2.65e-10	—	—	—	—	—	positive
B: knowledge-item reliability	Knowledge items 1–2	—	—	—	—	—	—	—	—	—	—	—	—	—	0.652	0.790	—
C: familiarity $\times$ product	design_mean	omnibus interaction	six products	—	—	—	—	—	5.537	5	0.354	—	—	—	—	—	—
C: familiarity $\times$ product	authenticity_mean	omnibus interaction	six products	—	—	—	—	—	17.605	5	0.00348	—	—	—	—	—	—
C: familiarity $\times$ product	aesthetic_mean	omnibus interaction	six products	—	—	—	—	—	1.505	5	0.912	—	—	—	—	—	—
D: predicted product means	design_mean	minus_1_SD	P1	—	—	—	—	—	—	—	—	4.770	4	—	—	—	—
D: predicted product means	design_mean	minus_1_SD	P2	—	—	—	—	—	—	—	—	4.895	3	—	—	—	—
D: predicted product means	design_mean	minus_1_SD	P3	—	—	—	—	—	—	—	—	4.712	5	—	—	—	—
D: predicted product means	design_mean	minus_1_SD	P4	—	—	—	—	—	—	—	—	4.624	6	—	—	—	—
D: predicted product means	design_mean	minus_1_SD	P5	—	—	—	—	—	—	—	—	5.065	1	—	—	—	—
D: predicted product means	design_mean	minus_1_SD	P6	—	—	—	—	—	—	—	—	5.024	2	—	—	—	—
D: predicted product means	design_mean	plus_1_SD	P1	—	—	—	—	—	—	—	—	5.389	3	—	—	—	—
D: predicted product means	design_mean	plus_1_SD	P2	—	—	—	—	—	—	—	—	5.356	6	—	—	—	—
D: predicted product means	design_mean	plus_1_SD	P3	—	—	—	—	—	—	—	—	5.367	4	—	—	—	—
D: predicted product means	design_mean	plus_1_SD	P4	—	—	—	—	—	—	—	—	5.404	2	—	—	—	—
D: predicted product means	design_mean	plus_1_SD	P5	—	—	—	—	—	—	—	—	5.364	5	—	—	—	—
D: predicted product means	design_mean	plus_1_SD	P6	—	—	—	—	—	—	—	—	5.667	1	—	—	—	—
D: predicted product means	authenticity_mean	minus_1_SD	P1	—	—	—	—	—	—	—	—	4.686	5	—	—	—	—
D: predicted product means	authenticity_mean	minus_1_SD	P2	—	—	—	—	—	—	—	—	4.757	3	—	—	—	—
D: predicted product means	authenticity_mean	minus_1_SD	P3	—	—	—	—	—	—	—	—	4.702	4	—	—	—	—
D: predicted product means	authenticity_mean	minus_1_SD	P4	—	—	—	—	—	—	—	—	4.359	6	—	—	—	—
D: predicted product means	authenticity_mean	minus_1_SD	P5	—	—	—	—	—	—	—	—	4.913	1	—	—	—	—
D: predicted product means	authenticity_mean	minus_1_SD	P6	—	—	—	—	—	—	—	—	4.831	2	—	—	—	—
D: predicted product means	authenticity_mean	plus_1_SD	P1	—	—	—	—	—	—	—	—	5.431	4	—	—	—	—
D: predicted product means	authenticity_mean	plus_1_SD	P2	—	—	—	—	—	—	—	—	5.308	6	—	—	—	—
D: predicted product means	authenticity_mean	plus_1_SD	P3	—	—	—	—	—	—	—	—	5.596	2	—	—	—	—

Section	Outcome	Model / level	Term / product	β (unstandardised)	SE	95% CI	Standardised β	PRV rater (%)	LRT χ^2	df	LRT P	Predicted mean	Rank	Rank order / Spearman ρ	Knowledge-item Pearson r	Spearman-Brown reliability	Direction
D: predicted product means	authenticity_mean	plus_1_SD	P4	—	—	—	—	—	—	—	—	5.576	3	—	—	—	—
D: predicted product means	authenticity_mean	plus_1_SD	P5	—	—	—	—	—	—	—	—	5.415	5	—	—	—	—
D: predicted product means	authenticity_mean	plus_1_SD	P6	—	—	—	—	—	—	—	—	5.669	1	—	—	—	—
D: predicted product means	aesthetic_mean	minus_1_SD	P1	—	—	—	—	—	—	—	—	4.239	6	—	—	—	—
D: predicted product means	aesthetic_mean	minus_1_SD	P2	—	—	—	—	—	—	—	—	4.625	2	—	—	—	—
D: predicted product means	aesthetic_mean	minus_1_SD	P3	—	—	—	—	—	—	—	—	4.415	3	—	—	—	—
D: predicted product means	aesthetic_mean	minus_1_SD	P4	—	—	—	—	—	—	—	—	4.354	5	—	—	—	—
D: predicted product means	aesthetic_mean	minus_1_SD	P5	—	—	—	—	—	—	—	—	4.385	4	—	—	—	—
D: predicted product means	aesthetic_mean	minus_1_SD	P6	—	—	—	—	—	—	—	—	4.848	1	—	—	—	—
D: predicted product means	aesthetic_mean	plus_1_SD	P1	—	—	—	—	—	—	—	—	5.214	6	—	—	—	—
D: predicted product means	aesthetic_mean	plus_1_SD	P2	—	—	—	—	—	—	—	—	5.530	3	—	—	—	—
D: predicted product means	aesthetic_mean	plus_1_SD	P3	—	—	—	—	—	—	—	—	5.446	4	—	—	—	—
D: predicted product means	aesthetic_mean	plus_1_SD	P4	—	—	—	—	—	—	—	—	5.554	2	—	—	—	—
D: predicted product means	aesthetic_mean	plus_1_SD	P5	—	—	—	—	—	—	—	—	5.429	5	—	—	—	—
D: predicted product means	aesthetic_mean	plus_1_SD	P6	—	—	—	—	—	—	—	—	5.679	1	—	—	—	—
E: rank concordance	design_mean	-1 vs +1 SD	P1-P6	—	—	—	—	—	—	—	—	—	—	$\rho=-0.257$; -1 SD: P5>P6>P2>P1>P3>P4; +1 SD: P6>P4>P1>P3>P5>P2	—	—	—
E: rank concordance	authenticity_mean	-1 vs +1 SD	P1-P6	—	—	—	—	—	—	—	—	—	—	$\rho=-0.143$; -1 SD: P5>P6>P2>P3>P1>P4; +1 SD: P6>P3>P4>P1>P5>P2	—	—	—
E: rank concordance	aesthetic_mean	-1 vs +1 SD	P1-P6	—	—	—	—	—	—	—	—	—	—	$\rho=0.657$; -1 SD: P6>P2>P3>P5>P4>P1; +1 SD: P6>P4>P2>P3>P5>P1	—	—	—

Panel: Conditional product differences

Outcome	Range at -1 SD	Range at +1 SD	Change: min	Change: max	Change: mean	Slope SD	Spearman ρ	Interaction LRT P
design_mean	0.441	0.310	0.298	0.780	0.576	0.004	-0.257	0.354
authenticity_mean	0.554	0.360	0.502	1.217	0.791	0.085	-0.143	0.003

aesthetic_mean	0.608	0.465	0.831	1.200	0.997	0.000	0.657	0.912
Outcome	Product		Predicted mean at -1 SD		Predicted mean at +1 SD		Change (+1 minus -1 SD)	
design_mean	P1		4.770		5.389		0.619	
design_mean	P2		4.895		5.356		0.461	
design_mean	P3		4.712		5.367		0.655	
design_mean	P4		4.624		5.404		0.780	
design_mean	P5		5.065		5.364		0.298	
design_mean	P6		5.024		5.667		0.643	
authenticity_mean	P1		4.686		5.431		0.745	
authenticity_mean	P2		4.757		5.308		0.552	
authenticity_mean	P3		4.702		5.596		0.894	
authenticity_mean	P4		4.359		5.576		1.217	
authenticity_mean	P5		4.913		5.415		0.502	
authenticity_mean	P6		4.831		5.669		0.838	
aesthetic_mean	P1		4.239		5.214		0.974	
aesthetic_mean	P2		4.625		5.530		0.904	
aesthetic_mean	P3		4.415		5.446		1.031	
aesthetic_mean	P4		4.354		5.554		1.200	
aesthetic_mean	P5		4.385		5.429		1.044	
aesthetic_mean	P6		4.848		5.679		0.831	

Supplementary Table 4. Descriptive and evaluative family robustness

Note. Panel A uses 1,000 respondent-cluster bootstrap replicates. Panel B uses REML point estimates and 1,000 parametric-bootstrap percentile intervals for product share. Standard profile intervals and the available boundary-adjusted interval are on the product-variance scale, so their units differ from share intervals. The boundary-adjusted interval exists only for design perception; dashes for other constructs indicate unavailable values. Very small positive bootstrap limits retain scientific notation.

Panel	Variant / construct	Descriptive mean product share (%)	Evaluative mean product share (%)	Difference (pp)	Respondent-cluster bootstrap 95% CI (pp)	Product share (%)	Parametric bootstrap share 95% CI (%)	Standard profile product-variance 95% CI	Boundary-adjusted product-variance 95% CI	Standard profile excludes zero	Singular point fit
A: leave-one-construct-out	drop_distinctiveness	2.81	0.66	2.15	[0.47, 5.17]	—	—	—	—	—	—
A: leave-one-construct-out	drop_dynamism	1.18	0.66	0.52	[-0.85, 3.24]	—	—	—	—	—	—
B: seven constructs	modernity_mean	—	—	—	—	1.71	[0.00, 5.23]	[0.00000, 0.18238]	—	FALSE	FALSE
B: seven constructs	refinement_mean	—	—	—	—	1.83	[2.69e-14, 5.40]	[0.00343, 0.14487]	—	TRUE	FALSE
B: seven constructs	dynamism_mean	—	—	—	—	4.87	[4.23e-12, 12.91]	[0.02069, 0.40619]	—	TRUE	FALSE
B: seven constructs	distinctiveness_mean	—	—	—	—	0.00	[0.00, 1.24]	[0.00000, 0.03223]	—	FALSE	TRUE
B: seven constructs	design_mean	—	—	—	—	0.65	[0.00, 2.33]	[0.00000, 0.05528]	[0.00009, 0.04067]	FALSE	FALSE
B: seven constructs	authenticity_mean	—	—	—	—	0.35	[0.00, 1.49]	[0.00000, 0.03567]	—	FALSE	FALSE
B: seven constructs	aesthetic_mean	—	—	—	—	0.98	[0.00, 3.76]	[0.00000, 0.11029]	—	FALSE	FALSE

Supplementary Table 5. Pooled regressions and indirect association

Note. Ordinary least squares models include product fixed effects and control for cultural familiarity and presentation order. Standard errors are clustered by respondent with a small-sample correction; n = 456 evaluations from 228 respondents. P values for direct associations were derived from the saved coefficient and standard error using a t distribution with 227 degrees of freedom. The indirect-association standard error is the standard deviation of 2,000 respondent-cluster bootstrap estimates; its 95% interval is percentile based. The empirical two-sided P value is twice the smaller proportion of bootstrap estimates at or below versus at or above zero; no replicate crossed zero, so $P < 0.001$ at the bootstrap resolution. The indirect association is standardised by $SD(\text{design perception}) / SD(\text{aesthetic evaluation})$.

Association	Unstandardised b	SE	95% CI	Standardised β [95% CI]	P
Design perception → perceived cultural authenticity	0.740	0.040	[0.661, 0.819]	0.761 [0.680, 0.842]	<0.001
Perceived cultural authenticity → aesthetic evaluation (design perception controlled)	0.405	0.087	[0.235, 0.574]	0.342 [0.198, 0.485]	<0.001
Design perception → aesthetic evaluation (authenticity controlled)	0.369	0.086	[0.201, 0.537]	0.321 [0.175, 0.467]	<0.001
Design perception → aesthetic evaluation (total)	0.669	0.056	[0.559, 0.779]	0.581 [0.485, 0.677]	<0.001
Indirect association	0.300	0.065	[0.174, 0.425]	0.260 [0.151, 0.369]	<0.001

Supplementary Table 6. Within-rater and between-rater associations: measurement error and screening sensitivity

Panel A. Multilevel reliability and item ICC

Construct	Level	α	ω	Converged	Heywood	Robust χ^2	df	P	Robust CFI	Robust RMSEA	SRMR
DP	within	0.669	0.711	TRUE	FALSE	58.049	18	<0.001	0.952	0.087	0.071
DP	between	0.936	0.968	TRUE	FALSE	58.049	18	<0.001	0.952	0.087	0.055
CA	within	0.458	0.559	TRUE	FALSE	0.000	0			0.000	0.001
CA	between	0.891	0.944	TRUE	TRUE	0.000	0			0.000	0.001
AE	within	0.843	0.844	TRUE	FALSE	0.000	0			0.000	0.001
AE	between	0.929	0.966	TRUE	FALSE	0.000	0			0.000	0.002

Construct	Item	Respondent ICC
DP	dp_color	0.651
DP	dp_form	0.654
DP	dp_pattern	0.647
DP	dp_craft	0.655
DP	dp_symbol	0.695
DP	dp_coherence	0.687
CA	ca_essence	0.721
CA	ca_atmosphere	0.731
CA	ca_credible	0.671
AE	ae_appealing	0.521
AE	ae_recommend	0.517
AE	ae_overall like	0.476

Panel B. Observed, latent and attenuation-corrected associations

Path	Level	Observed β [95% CI]	Latent β [95% CI]	α -corrected β [95% CI]	α valid n	ω -corrected β [95% CI]	ω valid n
DP → CA	within	0.103 [0.022, 0.184]	0.368 [-0.119, 0.854]	0.140 [0.027, 0.238]	1,770	0.131 [0.027, 0.235]	1,278
DP → CA	between	0.767 [0.688, 0.846]	0.930 [0.858, 1.002]	0.865 [0.800, 0.924]	1,770	0.820 [0.754, 0.873]	1,278
CA → AE (DP controlled)	within	0.005 [-0.079, 0.090]	-0.079 [-0.262, 0.104]	-0.021 [-0.725, 0.272]	1,770	-0.010 [-0.288, 0.156]	1,278
CA → AE (DP controlled)	between	0.391 [0.224, 0.557]	1.069 [-0.650, 2.789]	0.652 [0.218, 1.545]	1,770	0.474 [0.226, 0.763]	1,278
DP → AE (CA controlled)	within	0.060 [-0.053, 0.173]	0.133 [-0.165, 0.431]	0.095 [-0.160, 0.736]	1,770	0.083 [-0.084, 0.388]	1,278
DP → AE (CA controlled)	between	0.277 [0.112, 0.441]	-0.120 [-1.803, 1.564]	0.067 [-0.803, 0.472]	1,770	0.222 [-0.054, 0.457]	1,278
DP → AE (total)	within	0.061 [-0.060, 0.182]	—	0.080 [-0.086, 0.234]	1,770	0.076 [-0.072, 0.231]	1,278
DP → AE (total)	between	0.591 [0.500, 0.682]	—	0.645 [0.543, 0.738]	1,770	0.622 [0.521, 0.711]	1,278
DP → CA → AE (indirect)	within	0.002 [-0.034, 0.039]	-0.029 [-0.107, 0.048]	-0.013 [-0.689, 0.218]	1,770	-0.005 [-0.247, 0.106]	1,278
DP → CA → AE (indirect)	between	0.314 [0.178, 0.442]	0.994 [-0.650, 2.639]	0.578 [0.192, 1.452]	1,770	0.400 [0.190, 0.648]	1,278

Latent SEM fit metric	Value
chisq.scaled	321.114
df.scaled	165.000
pvalue.scaled	<0.001
cfi.robust	0.945
tli.robust	0.928
rmsea.robust	0.049
srmr_within	0.061
srmr_between	0.059

Level	Indirect (unstandardised)	20,000-draw Monte Carlo 95% CI
w	-0.029	-0.144 to 0.056
b	14.706	-171.369 to 325.828

Panel C. Screening subsamples

Screen	Respondents	Model	Path	Level	Standardised β	95% CI
both_attention	127	C1	DP	within	0.054	-0.042 to 0.149
both_attention	127	C1	DP	between	0.754	0.644 to 0.863
both_attention	127	C2	DP	within	0.160	0.003 to 0.317
both_attention	127	C2	DP	between	0.139	-0.064 to 0.343
both_attention	127	C2	CA	within	0.008	-0.080 to 0.097
both_attention	127	C2	CA	between	0.485	0.270 to 0.699
both_attention	127	C3	DP	within	0.162	0.001 to 0.322
both_attention	127	C3	DP	between	0.520	0.383 to 0.656
duration_ge_255	206	C1	DP	within	0.108	0.021 to 0.196
duration_ge_255	206	C1	DP	between	0.753	0.666 to 0.840
duration_ge_255	206	C2	DP	within	0.106	-0.006 to 0.218
duration_ge_255	206	C2	DP	between	0.280	0.109 to 0.451
duration_ge_255	206	C2	CA	within	0.014	-0.064 to 0.093
duration_ge_255	206	C2	CA	between	0.393	0.222 to 0.564
duration_ge_255	206	C3	DP	within	0.110	-0.006 to 0.225
duration_ge_255	206	C3	DP	between	0.591	0.494 to 0.687
relaxed_duration_ge_177	307	C1	DP	within	0.127	0.032 to 0.221
relaxed_duration_ge_177	307	C1	DP	between	0.789	0.727 to 0.852
relaxed_duration_ge_177	307	C2	DP	within	0.043	-0.050 to 0.136
relaxed_duration_ge_177	307	C2	DP	between	0.317	0.160 to 0.473
relaxed_duration_ge_177	307	C2	CA	within	0.006	-0.056 to 0.069
relaxed_duration_ge_177	307	C2	CA	between	0.384	0.230 to 0.537
relaxed_duration_ge_177	307	C3	DP	within	0.045	-0.051 to 0.141
relaxed_duration_ge_177	307	C3	DP	between	0.636	0.561 to 0.711

Screen	Model	Path	Between – within (unstd.)	95% CI	P
both_attention	C1	DP	0.594	0.267 to 0.922	<0.001
both_attention	C2	DP	-0.440	-1.051 to 0.171	0.158
both_attention	C2	CA	0.546	0.112 to 0.980	0.014
both_attention	C3	DP	0.002	-0.615 to 0.618	0.996
duration_ge_255	C1	DP	0.430	0.148 to 0.711	0.003
duration_ge_255	C2	DP	-0.053	-0.512 to 0.405	0.820
duration_ge_255	C2	CA	0.435	0.059 to 0.811	0.023
duration_ge_255	C3	DP	0.308	-0.135 to 0.752	0.173
relaxed_duration_ge_177	C1	DP	0.390	0.067 to 0.712	0.018
relaxed_duration_ge_177	C2	DP	0.209	-0.181 to 0.598	0.294
relaxed_duration_ge_177	C2	CA	0.438	0.133 to 0.742	0.005
relaxed_duration_ge_177	C3	DP	0.576	0.206 to 0.946	0.002

Screen	Level	Standardised indirect	95% bootstrap CI	Valid / requested
both_attention	within	0.002	-0.040 to 0.024	2000 / 2000
both_attention	between	0.380	0.217 to 0.540	2000 / 2000
both_attention	pooled	0.270	0.126 to 0.418	2000 / 2000
duration_ge_255	within	0.005	-0.029 to 0.042	2000 / 2000
duration_ge_255	between	0.311	0.176 to 0.453	2000 / 2000
duration_ge_255	pooled	0.252	0.138 to 0.366	2000 / 2000
relaxed_duration_ge_177	within	0.003	-0.027 to 0.029	2000 / 2000
relaxed_duration_ge_177	between	0.320	0.199 to 0.445	2000 / 2000
relaxed_duration_ge_177	pooled	0.261	0.170 to 0.356	2000 / 2000

Panel D. Both-attention-check subsample: attenuation correction

Method	Model	Path	Level	Corrected β	95% bootstrap CI	Valid / 2,000
alpha	C1	DP	within	0.074	-0.087 to 0.189	1564
alpha	C1	DP	between	0.865	0.772 to 0.945	1564
alpha	C2	DP	within	0.220	0.002 to 1.660	1564
alpha	C2	DP	between	-0.407	-5.947 to 0.251	1564
alpha	C2	CA	within	-0.023	-1.468 to 0.220	1564
alpha	C2	CA	between	1.103	0.381 to 7.173	1564
alpha	C3	DP	within	0.211	0.034 to 0.427	1564
alpha	C3	DP	between	0.566	0.427 to 0.701	1564
alpha	indirect	DP_CA_AE	within	-0.009	-1.476 to 0.102	1564
alpha	indirect	DP_CA_AE	between	0.973	0.345 to 6.697	1564
omega	C1	DP	within	0.069	-0.085 to 0.167	1045
omega	C1	DP	between	0.820	0.725 to 0.902	1045
omega	C2	DP	within	0.204	0.034 to 0.547	1045
omega	C2	DP	between	-0.044	-0.663 to 0.328	1045
omega	C2	CA	within	-0.010	-0.305 to 0.157	1045
omega	C2	CA	between	0.704	0.301 to 1.400	1045
omega	C3	DP	within	0.201	0.033 to 0.389	1045
omega	C3	DP	between	0.550	0.418 to 0.671	1045
omega	indirect	DP_CA_AE	within	-0.003	-0.218 to 0.048	1045
omega	indirect	DP_CA_AE	between	0.593	0.262 to 1.193	1045

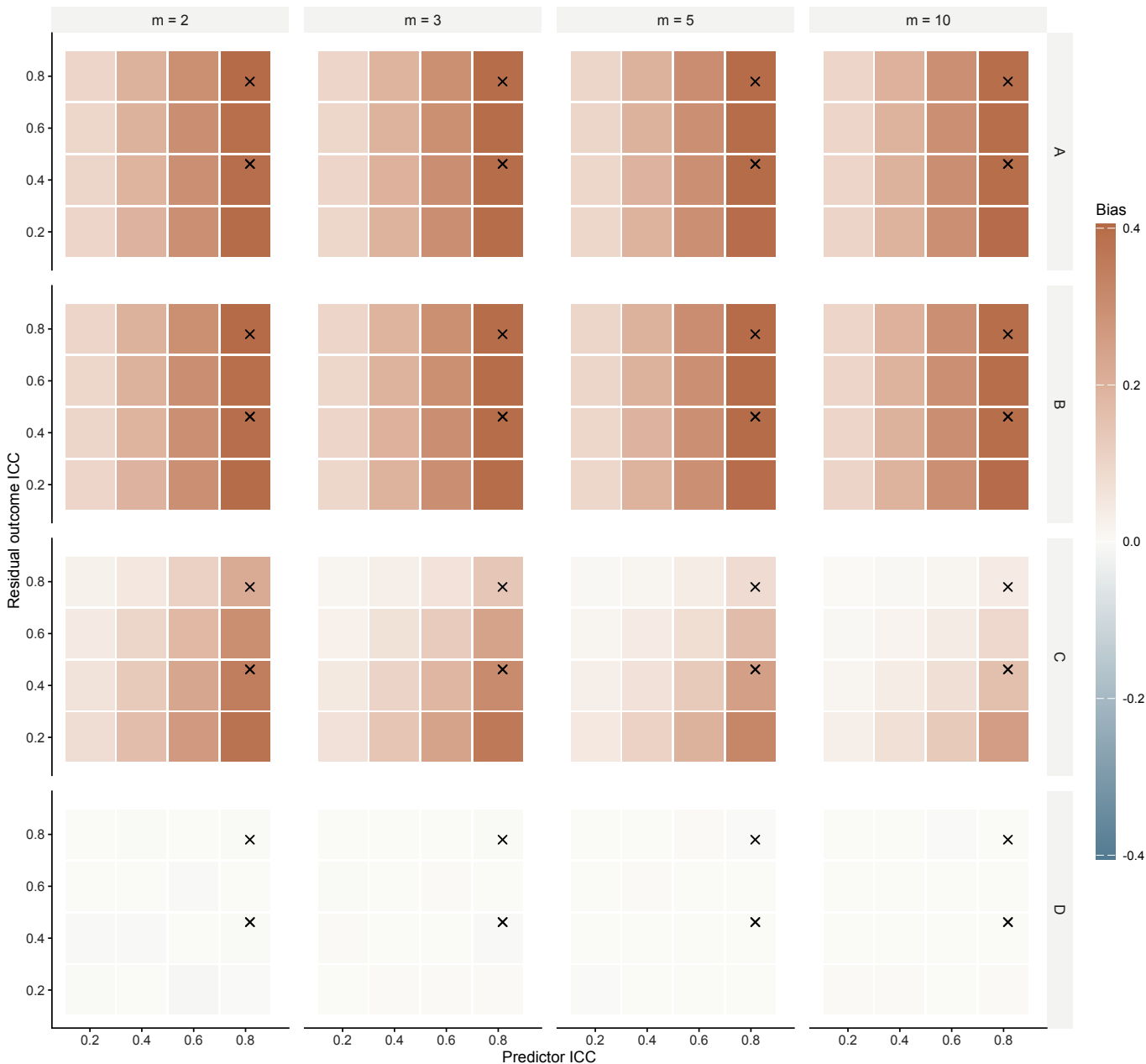
Note. DP = design perception; CA = perceived cultural authenticity; AE = aesthetic evaluation. Intervals for attenuation correction are percentile intervals from 2,000 respondent-cluster bootstrap replicates; invalid reliability estimates and non-positive-definite corrected cross-product matrices were excluded. Valid replicates were 1,770 (α) and 1,278 (ω) for the full sample, and 1,564 (α) and 1,045 (ω) for the both-attention-check sample. Failures by invalid reliability / non-positive-definite matrix were respectively 6 / 224, 714 / 8, 35 / 401 and 947 / 8 in that order. The between-level CA CFA had a Heywood residual variance. The two-level SEM converged with an approximately non-positive-definite parameter covariance matrix warning; between-level latent DP and CA were nearly collinear, so their separate paths to AE were not identified. The single-construct CA and AE CFAs have $df = 0$ at each level, so their overall fit indices do not test model fit. Product indicators were included at the within level of the SEM.

Supplementary Data 1

Bias and 95% confidence-interval coverage surfaces for estimators A–D (A, ordinary least squares; B, ordinary least squares with cluster-robust standard errors; C, random-intercept model; D, within–between decomposition) across all simulation cells. See Supplementary Note 3.

Bias surface | N = 100; true within slope = 0.000

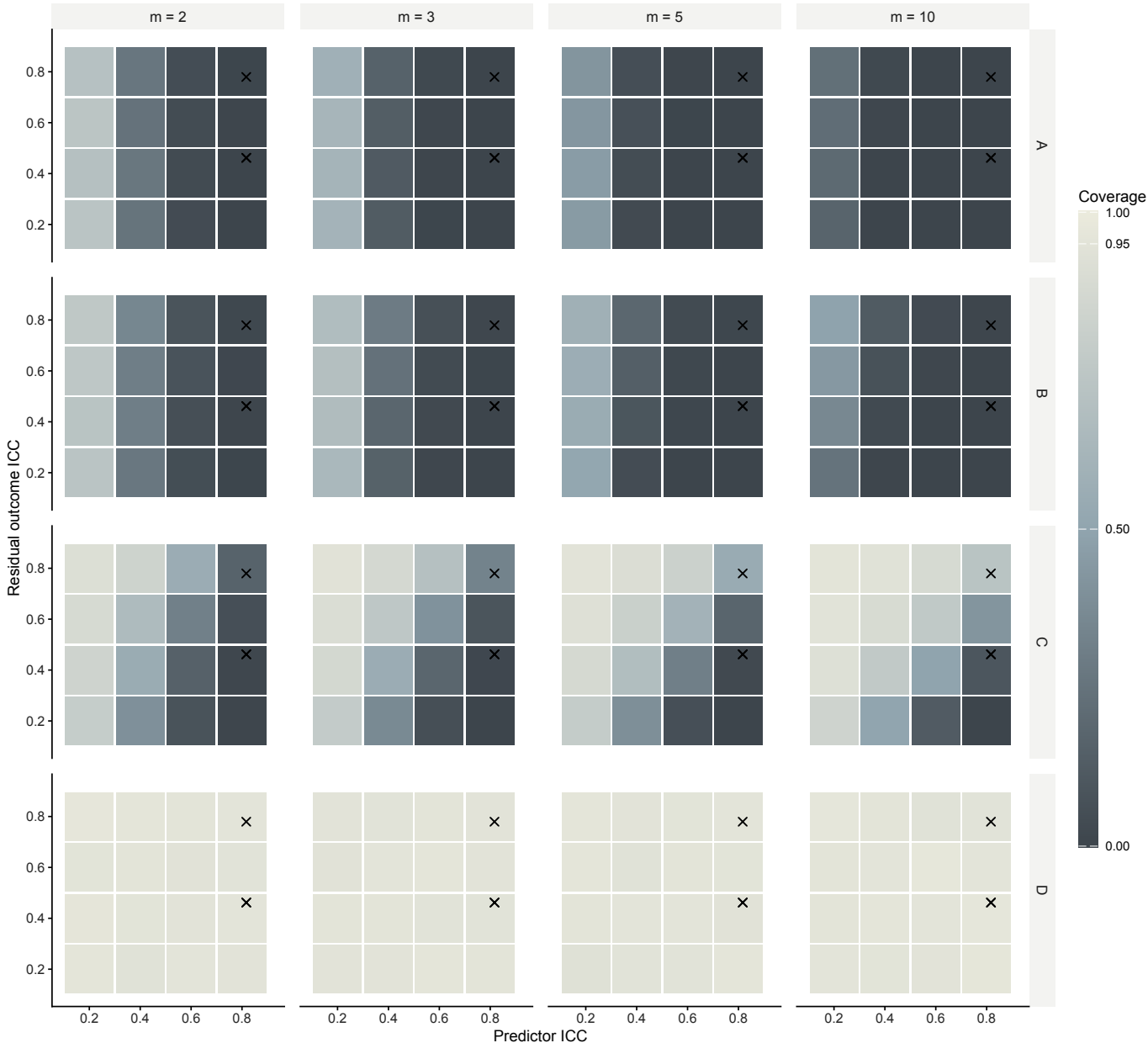
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 1 of 24

95% CI coverage surface | N = 100; true within slope = 0.000

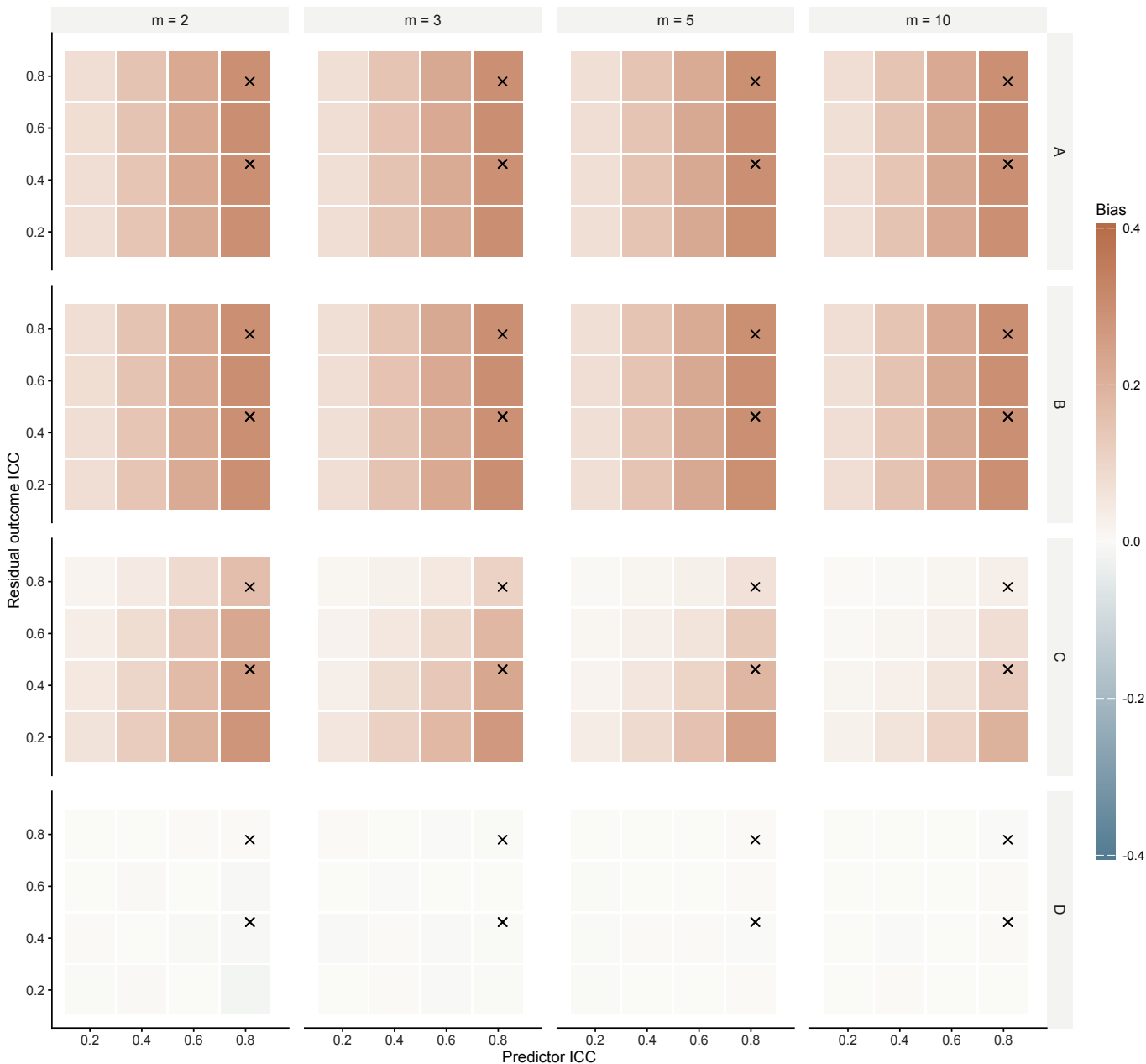
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 2 of 24

Bias surface | N = 100; true within slope = 0.125

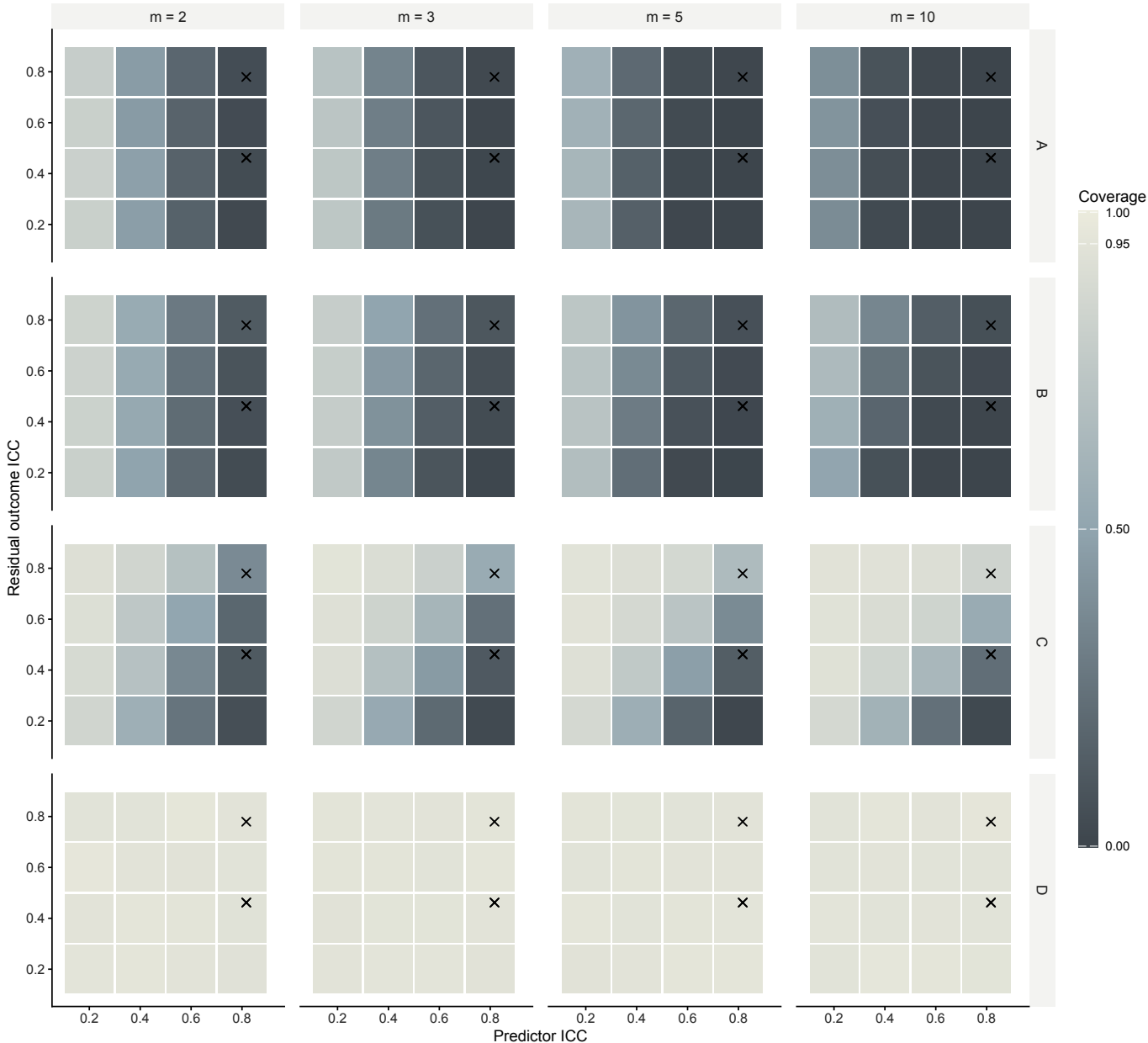
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 3 of 24

95% CI coverage surface | N = 100; true within slope = 0.125

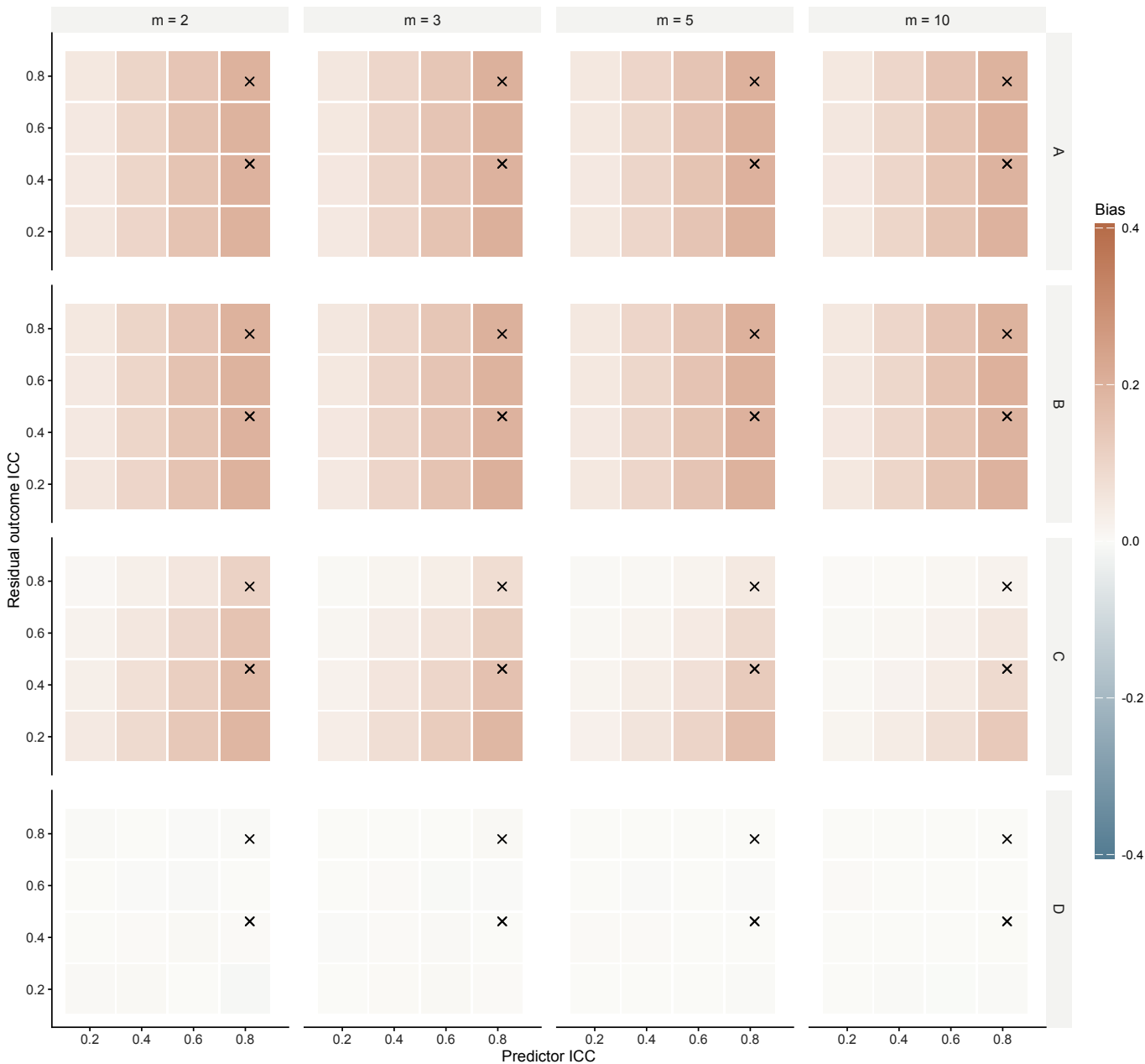
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 4 of 24

Bias surface | N = 100; true within slope = 0.250

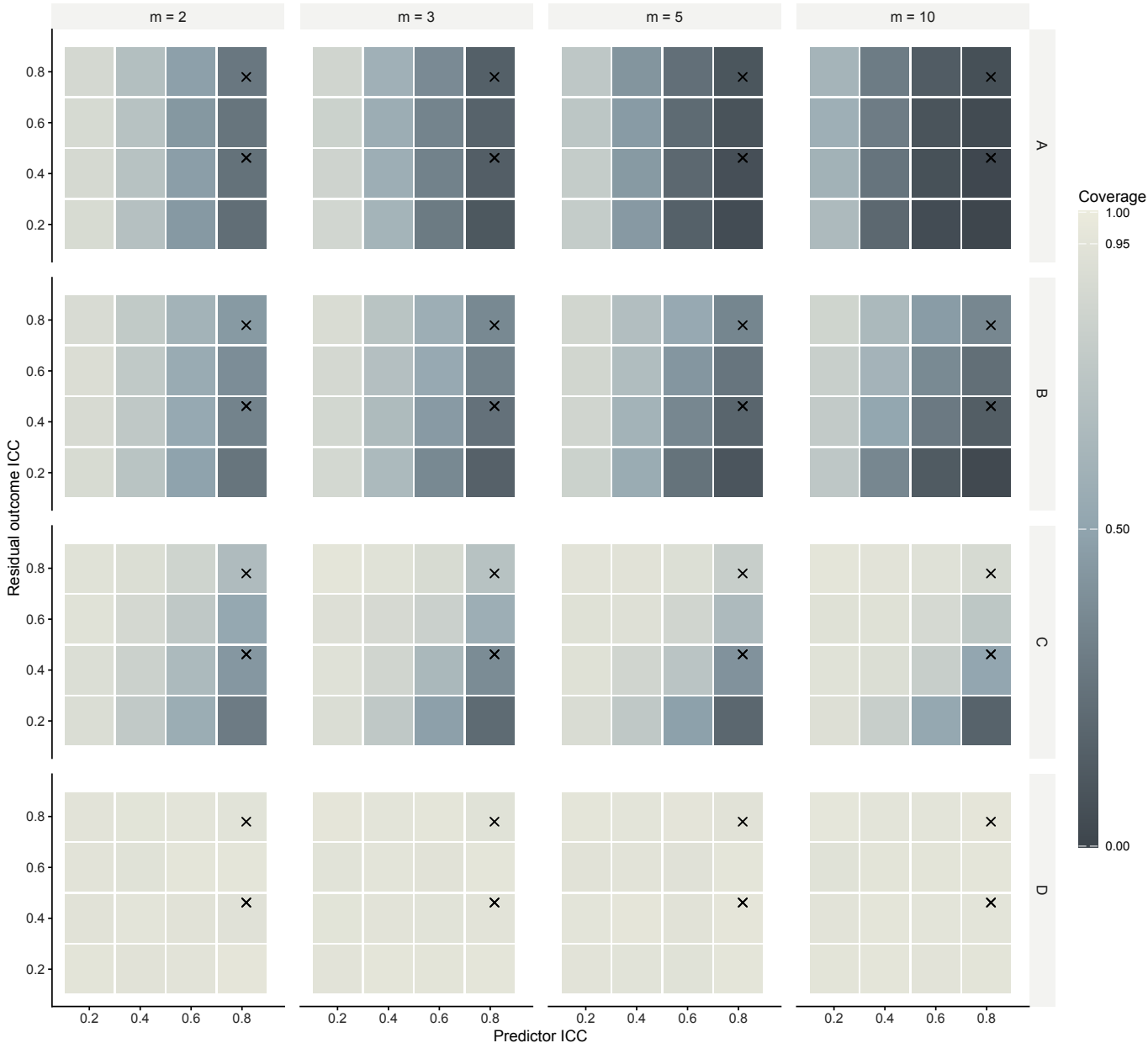
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 5 of 24

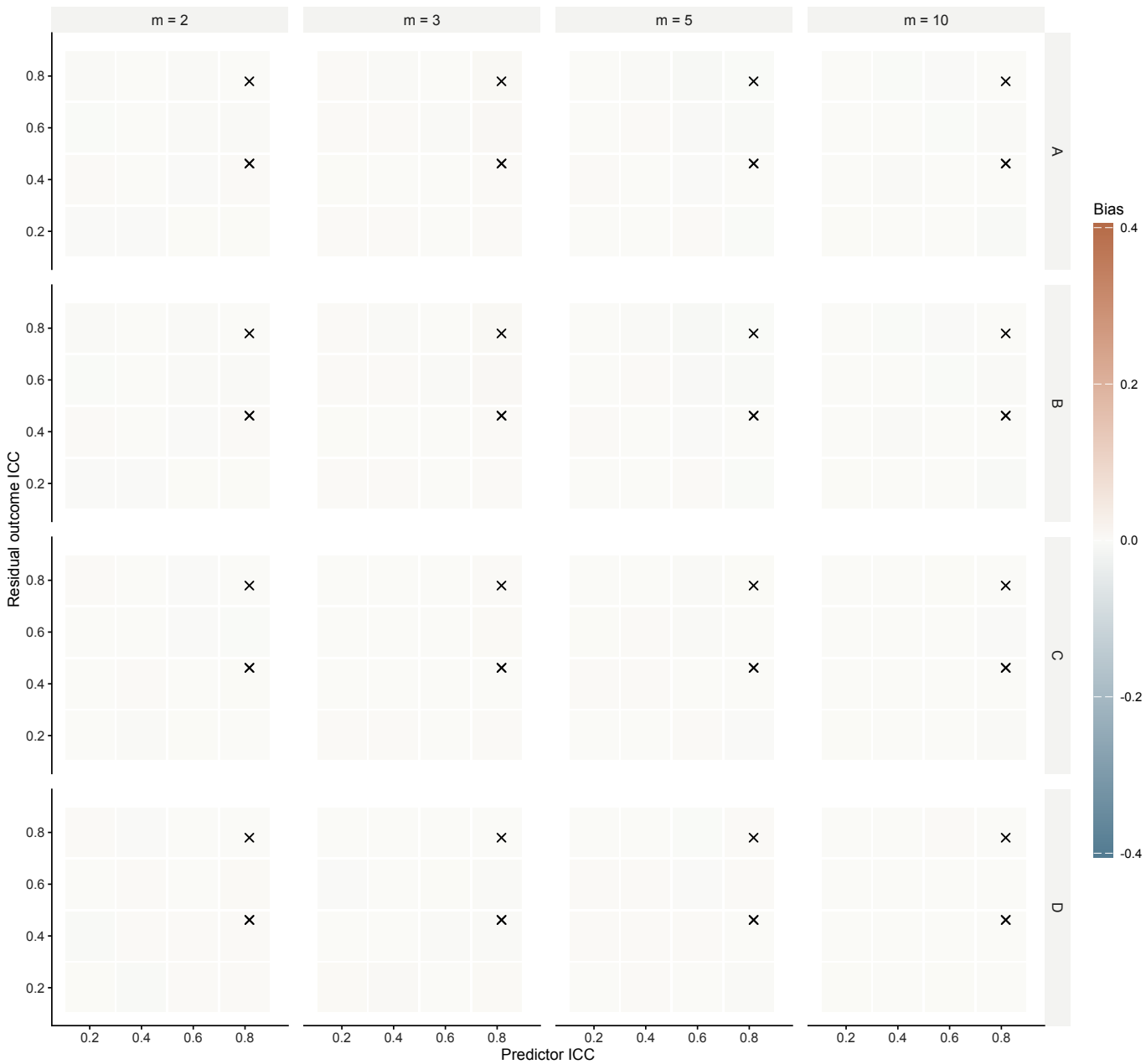
95% CI coverage surface | N = 100; true within slope = 0.250

A: OLS | B: clustered OLS | C: random intercept | D: within-between



Bias surface | N = 100; true within slope = 0.500

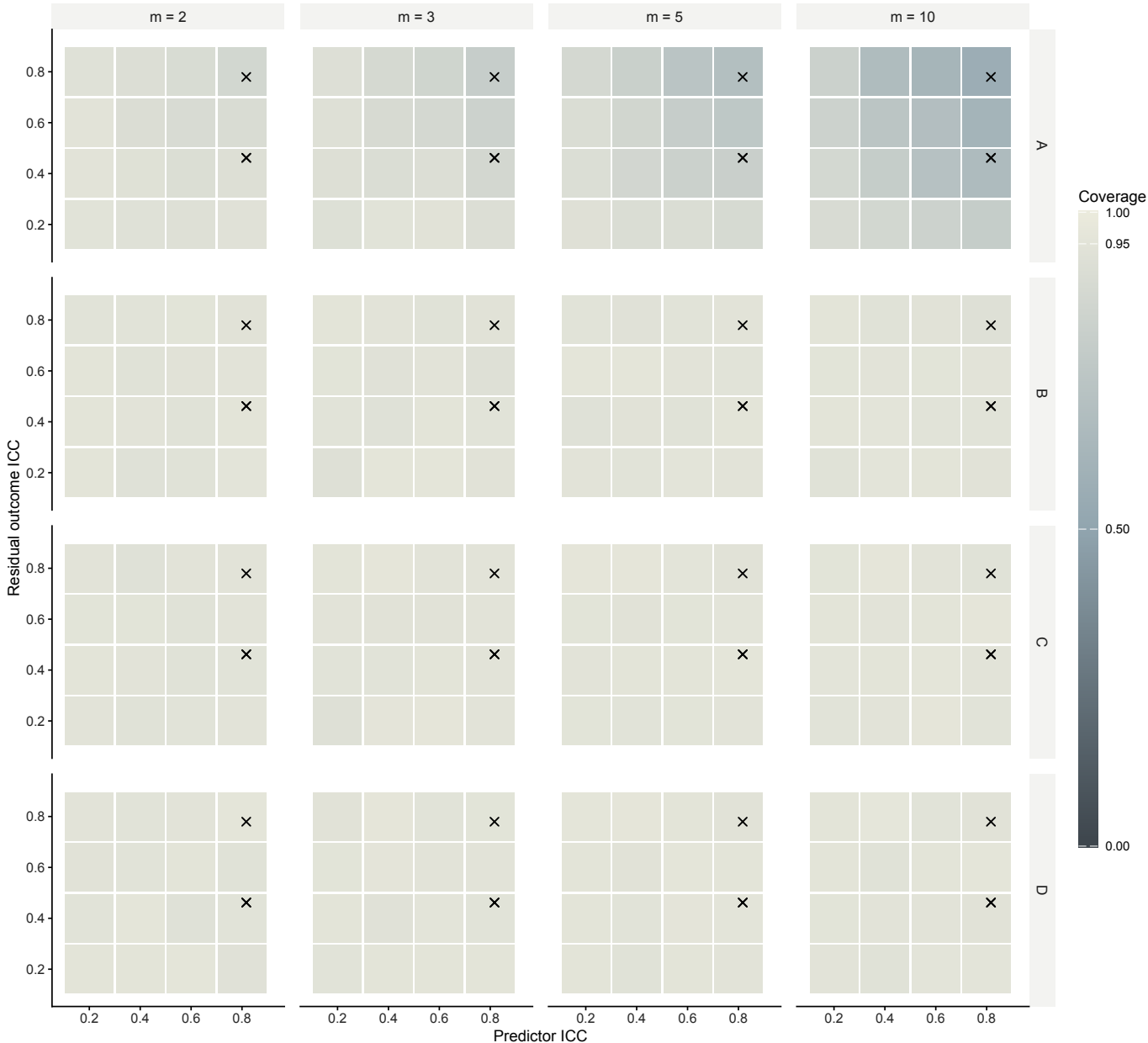
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 7 of 24

95% CI coverage surface | N = 100; true within slope = 0.500

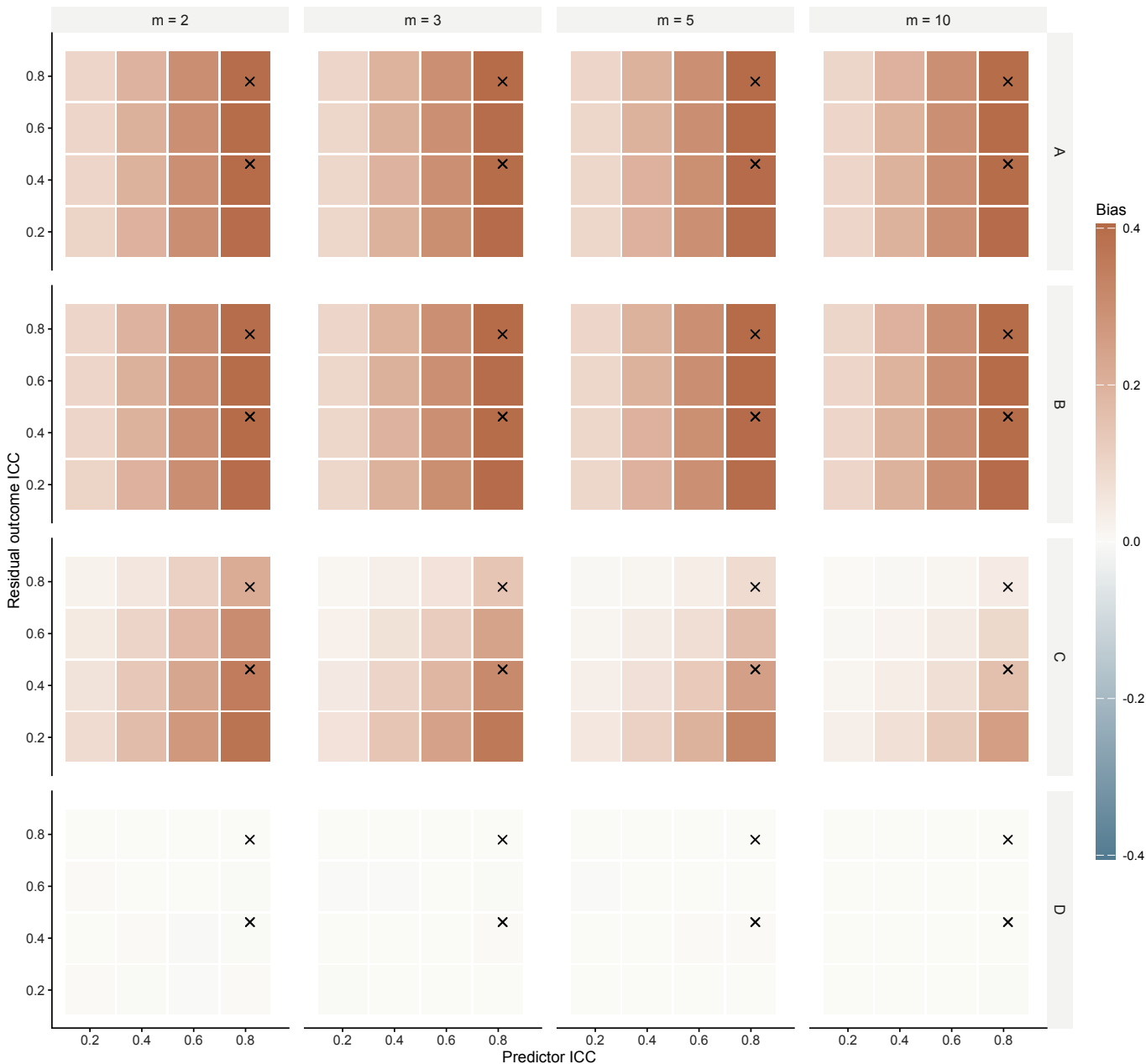
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 8 of 24

Bias surface | N = 228; true within slope = 0.000

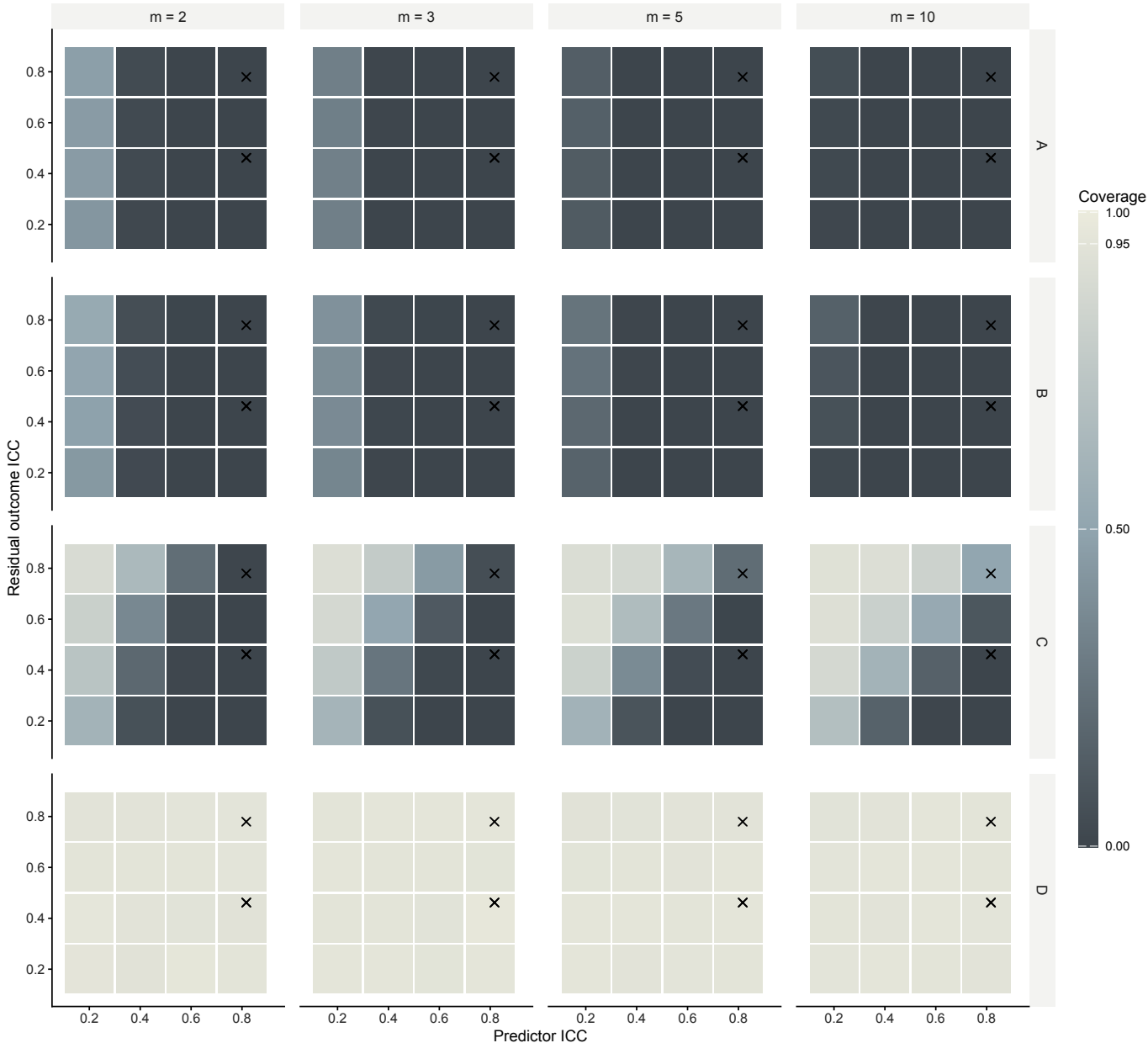
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 9 of 24

95% CI coverage surface | N = 228; true within slope = 0.000

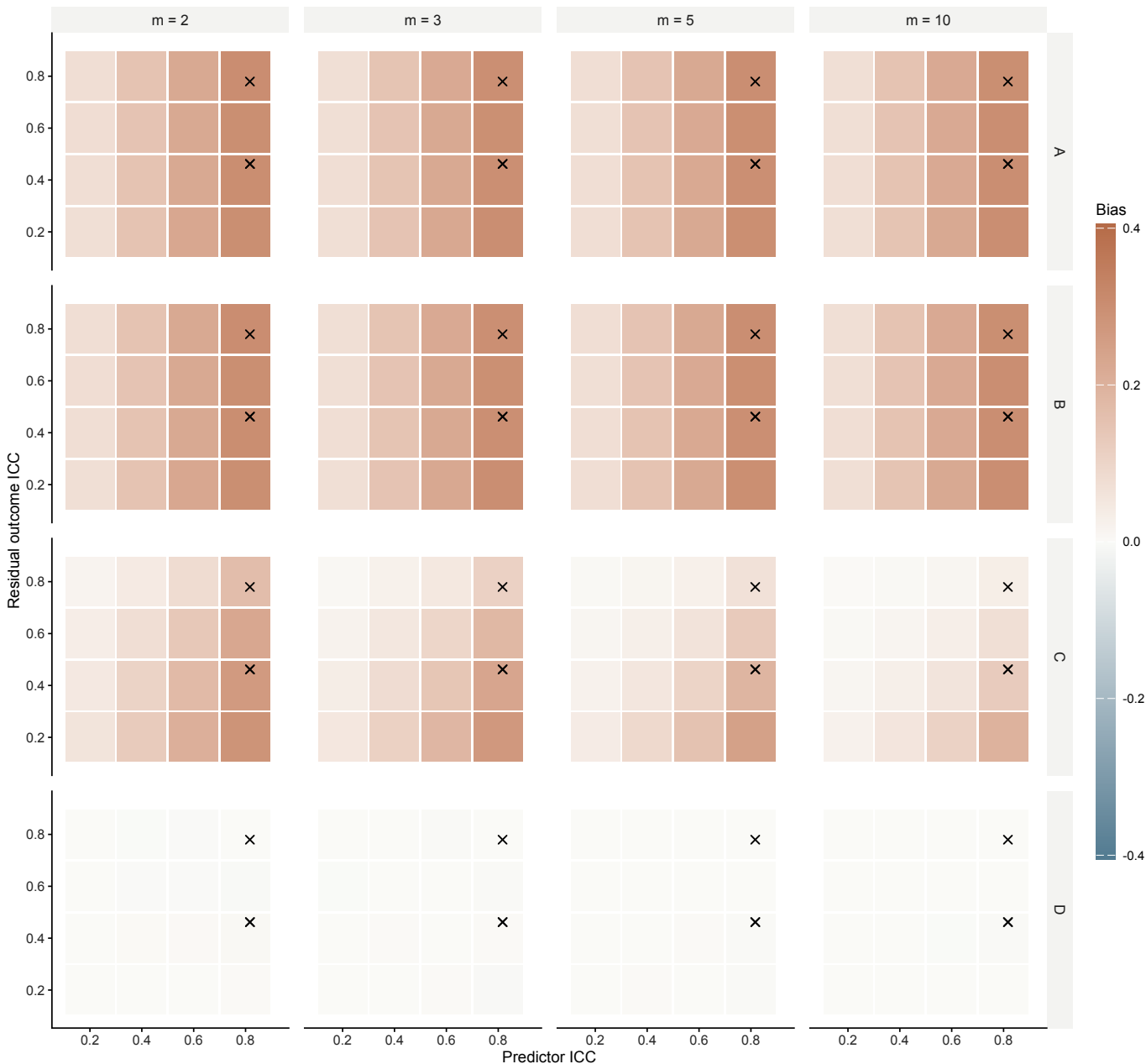
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 10 of 24

Bias surface | N = 228; true within slope = 0.125

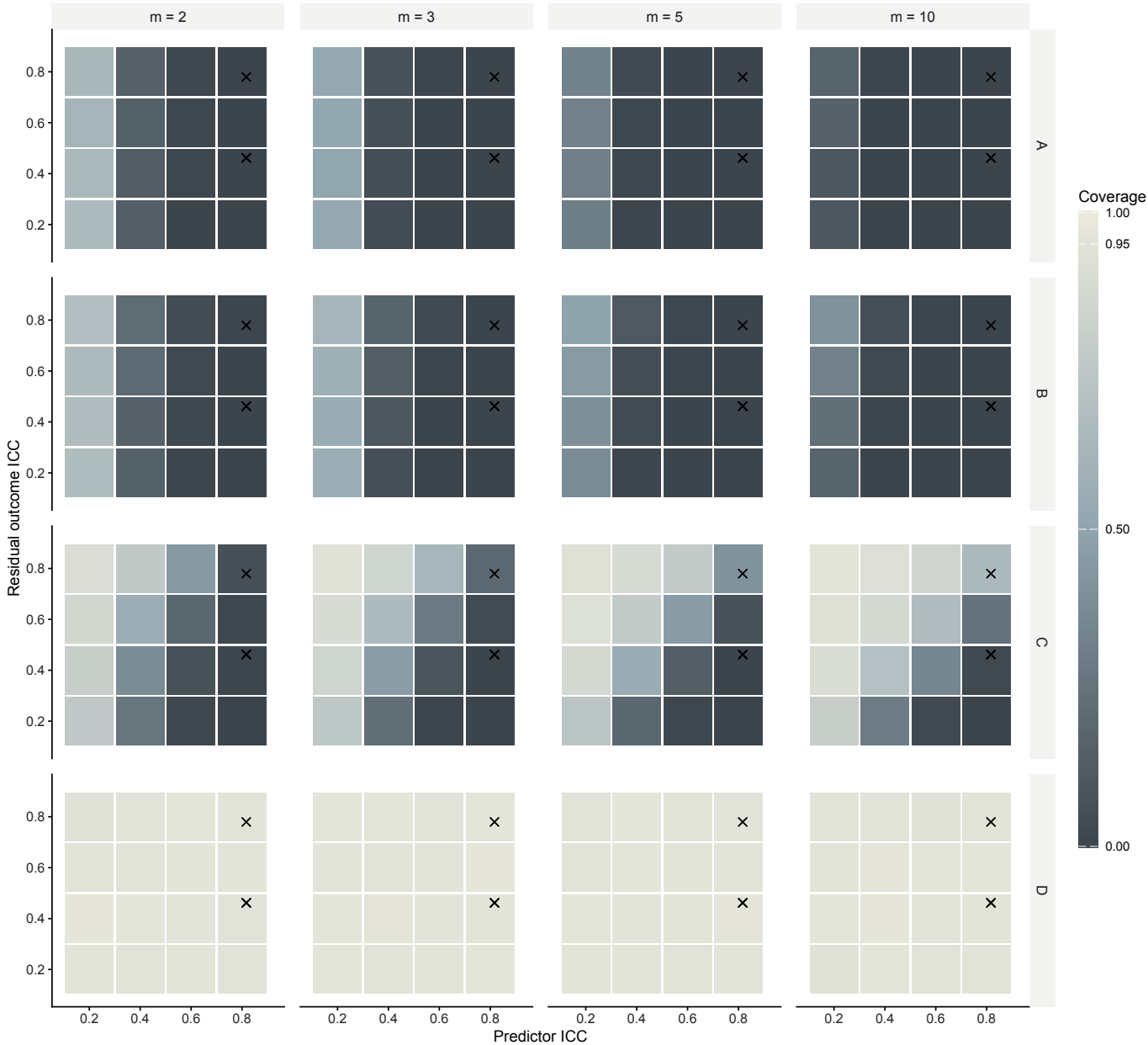
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 11 of 24

95% CI coverage surface | N = 228; true within slope = 0.125

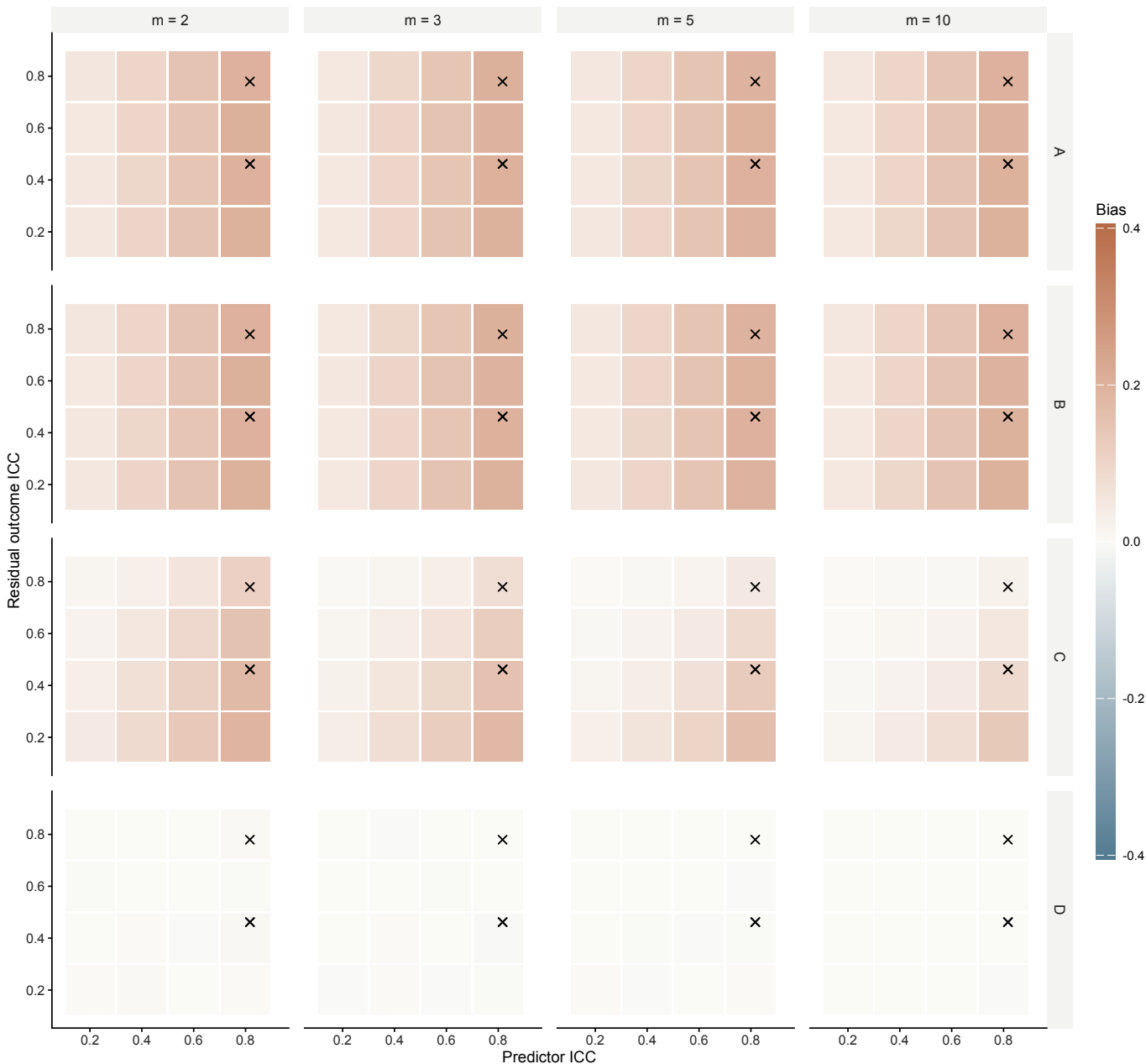
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 12 of 24

Bias surface | N = 228; true within slope = 0.250

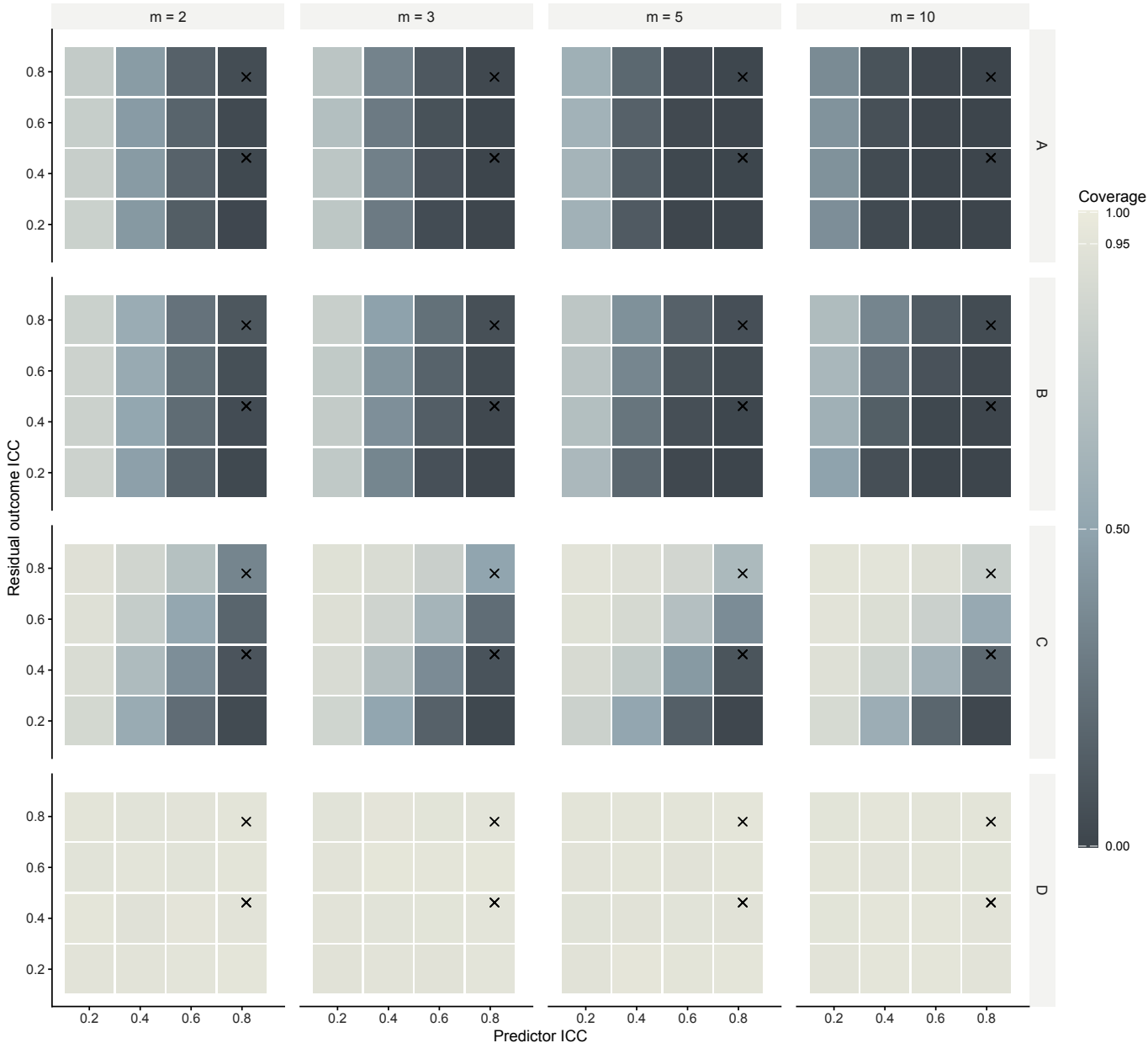
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 13 of 24

95% CI coverage surface | N = 228; true within slope = 0.250

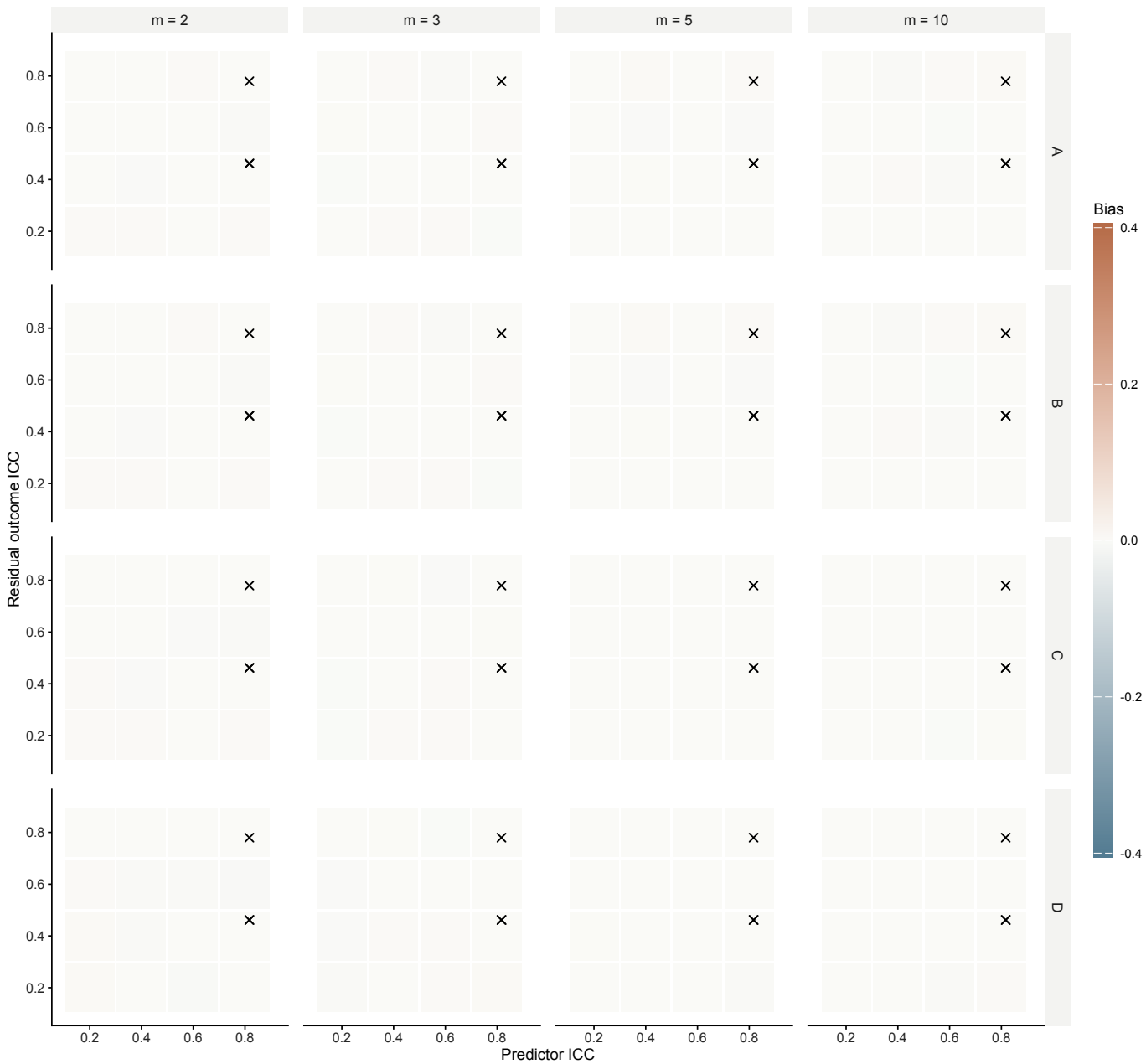
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 14 of 24

Bias surface | N = 228; true within slope = 0.500

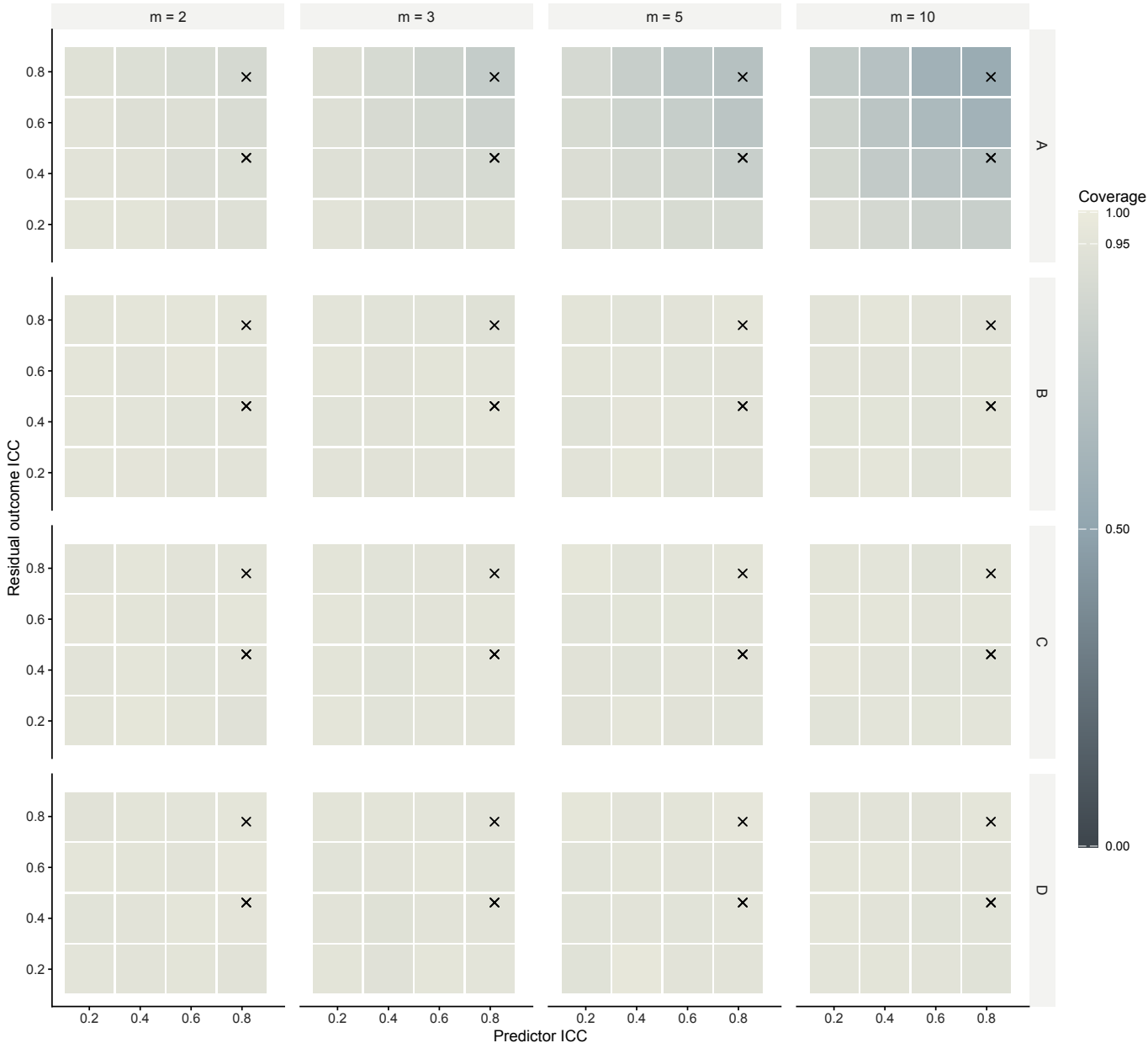
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 15 of 24

95% CI coverage surface | N = 228; true within slope = 0.500

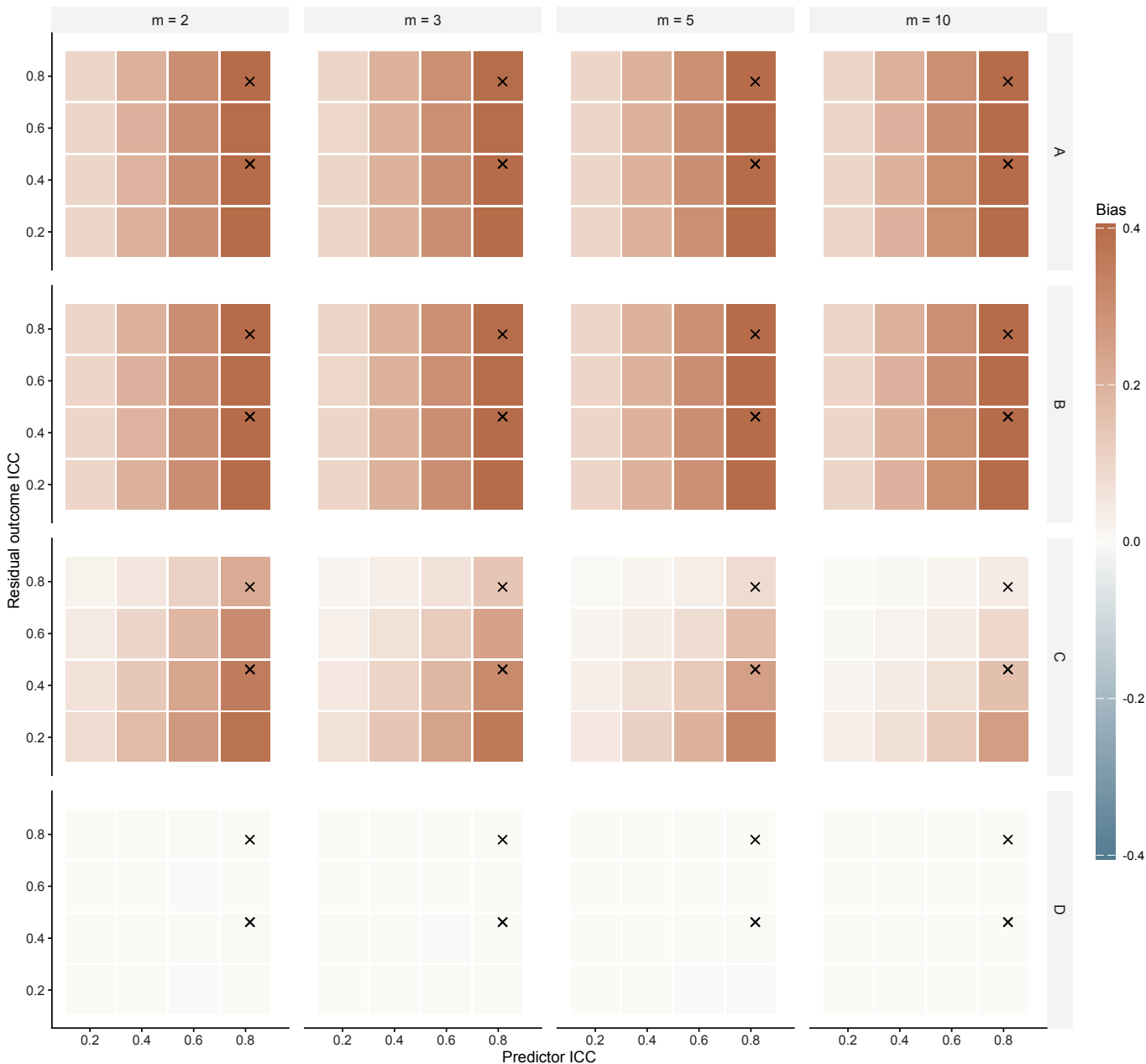
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 16 of 24

Bias surface | N = 500; true within slope = 0.000

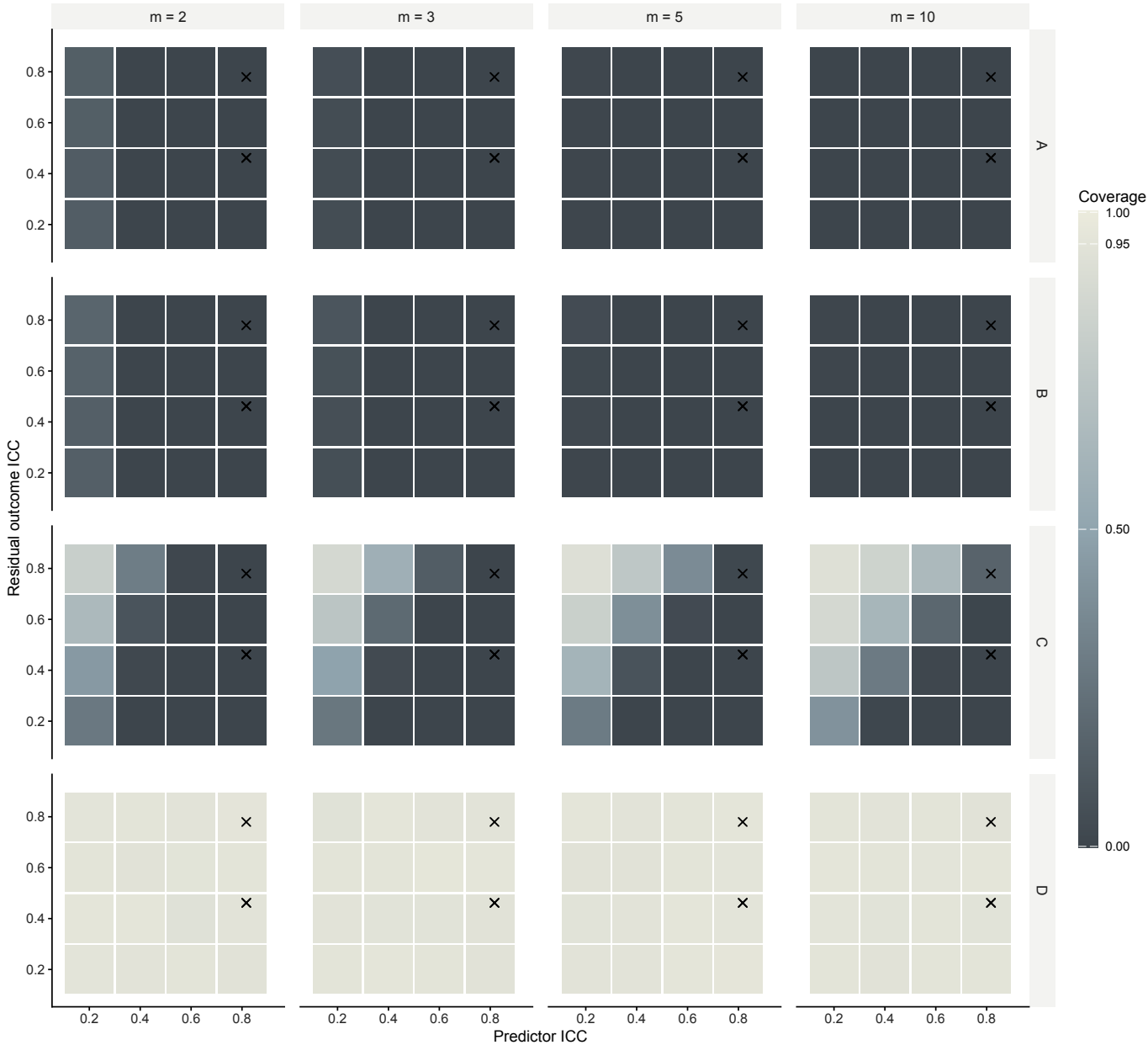
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 17 of 24

95% CI coverage surface | N = 500; true within slope = 0.000

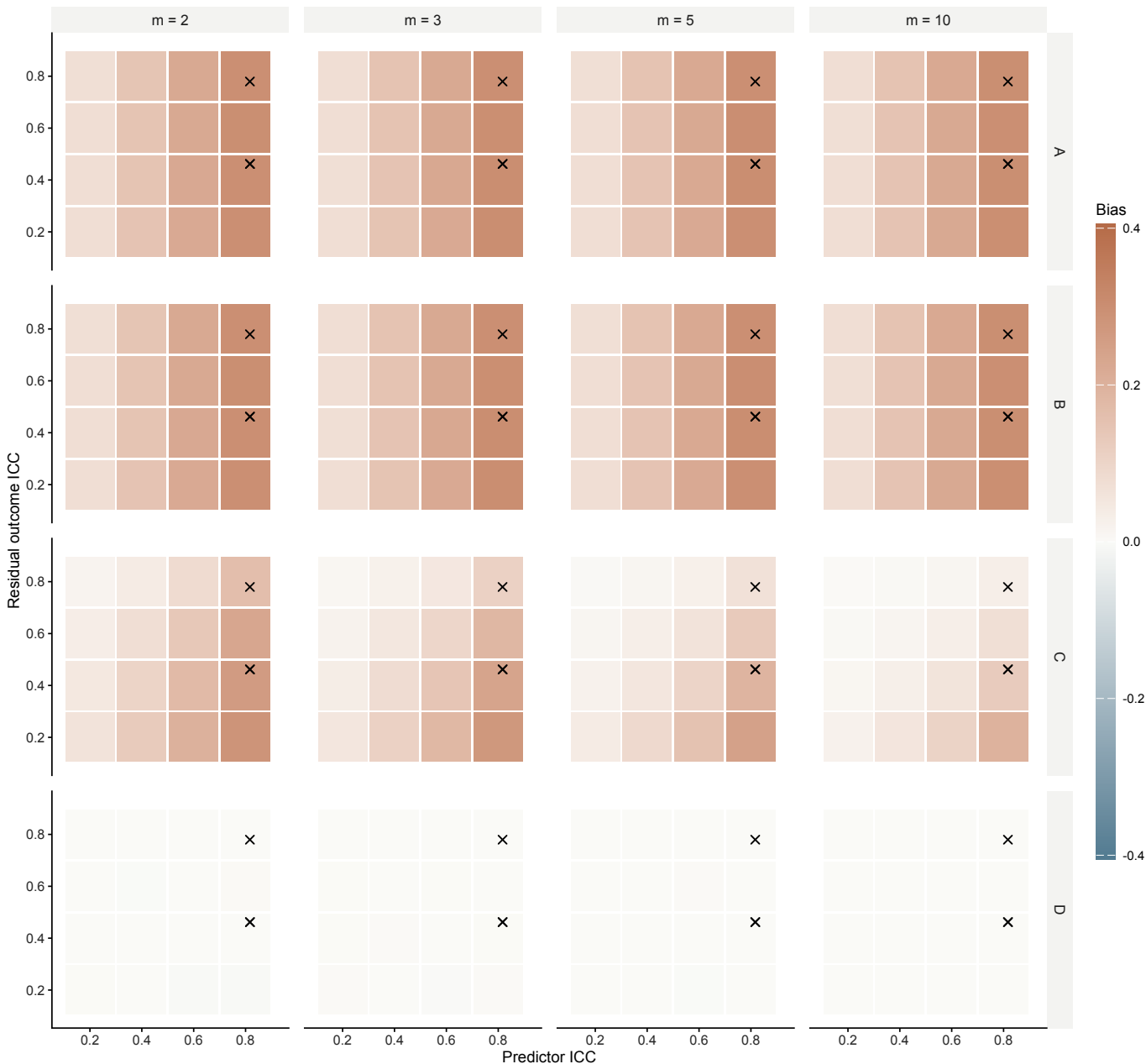
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 18 of 24

Bias surface | N = 500; true within slope = 0.125

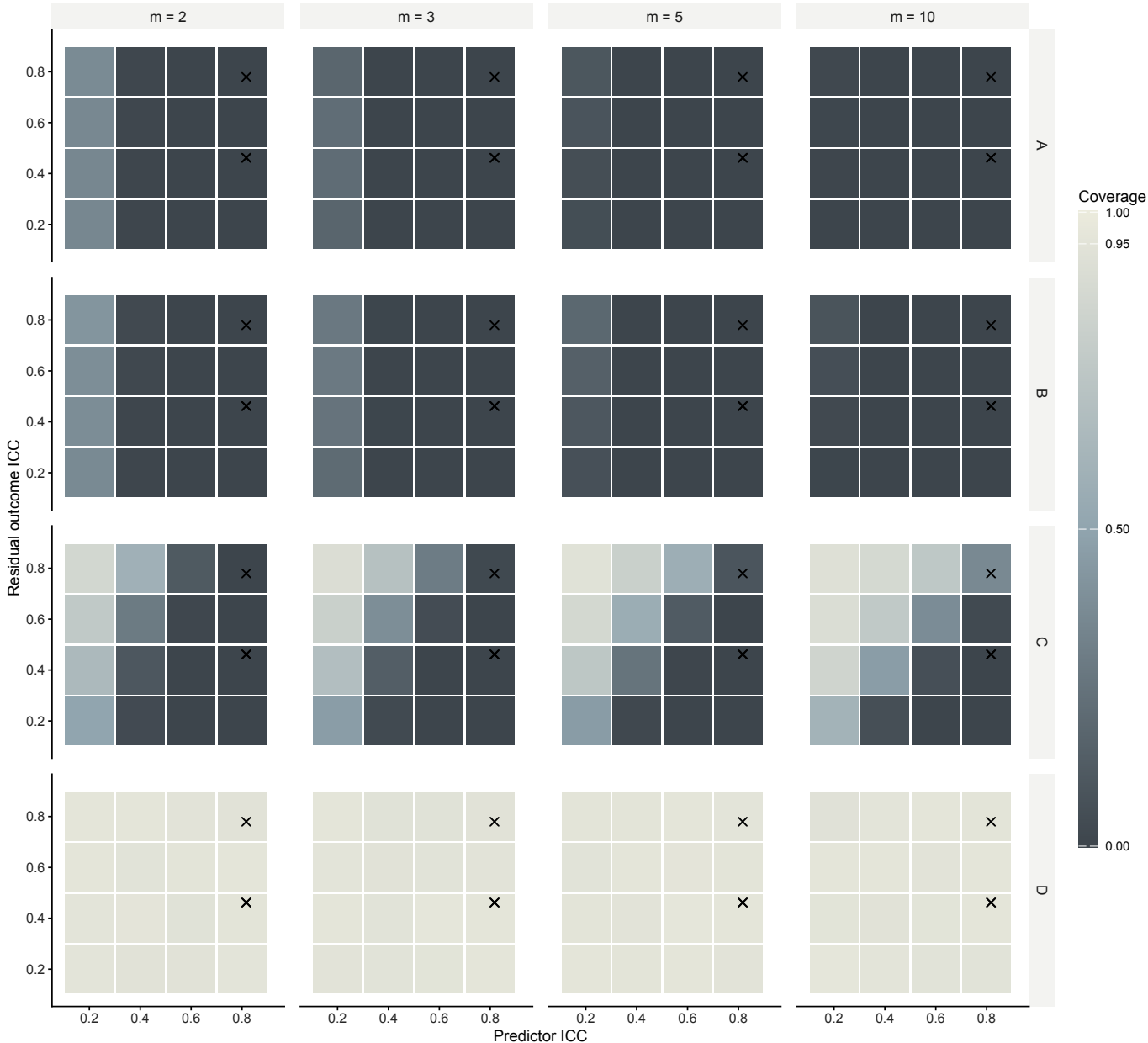
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 19 of 24

95% CI coverage surface | N = 500; true within slope = 0.125

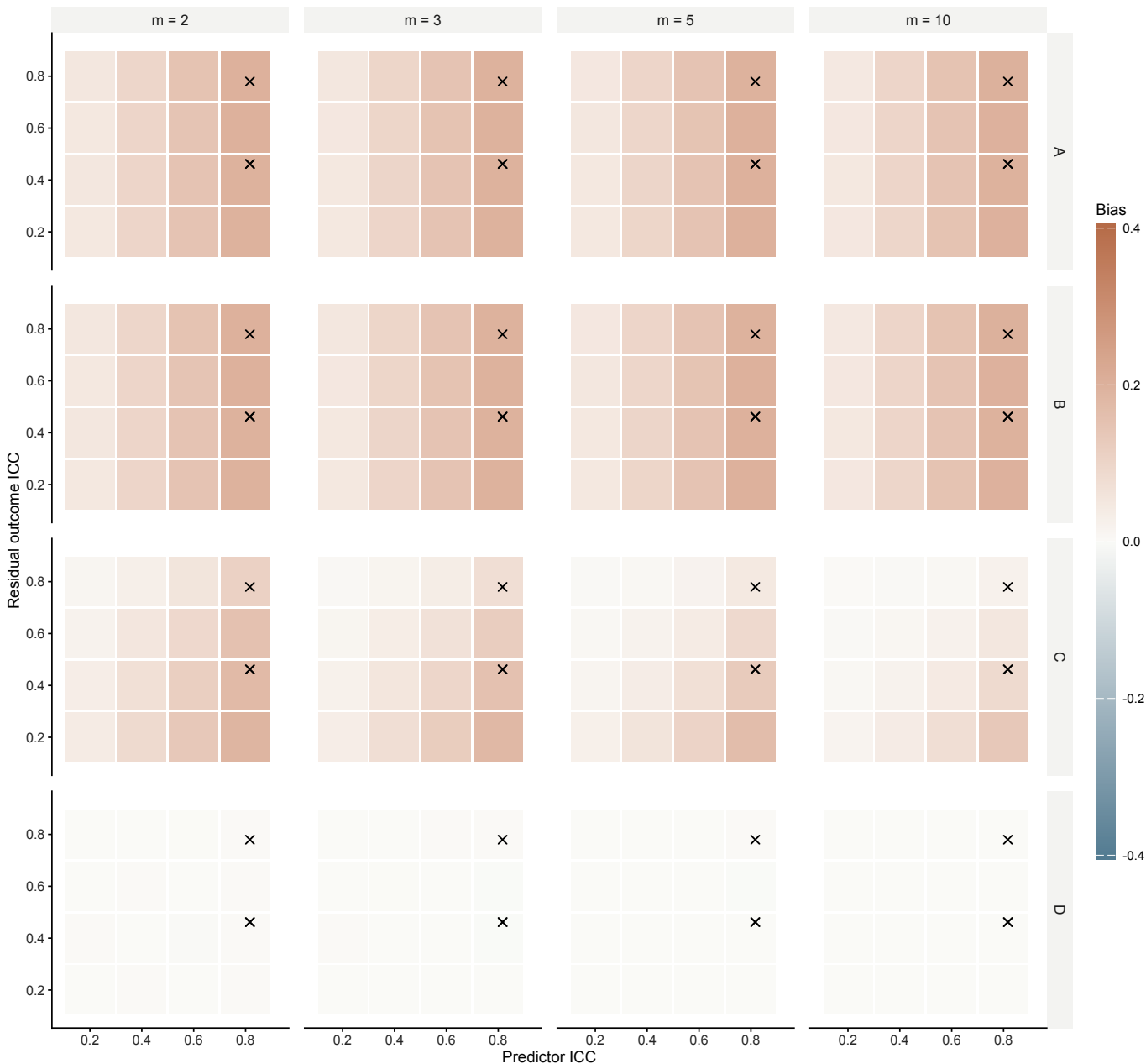
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 20 of 24

Bias surface | N = 500; true within slope = 0.250

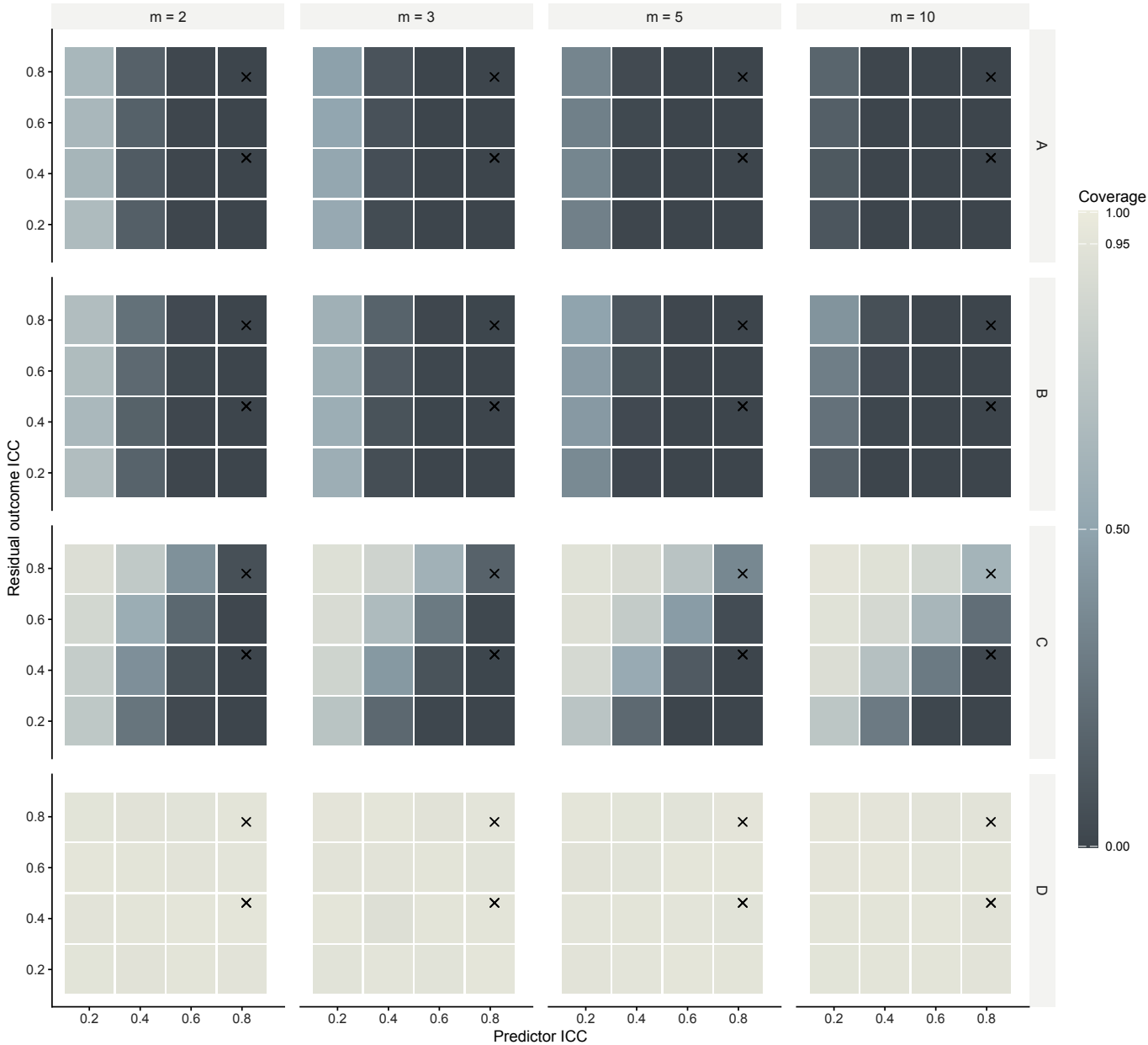
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 21 of 24

95% CI coverage surface | N = 500; true within slope = 0.250

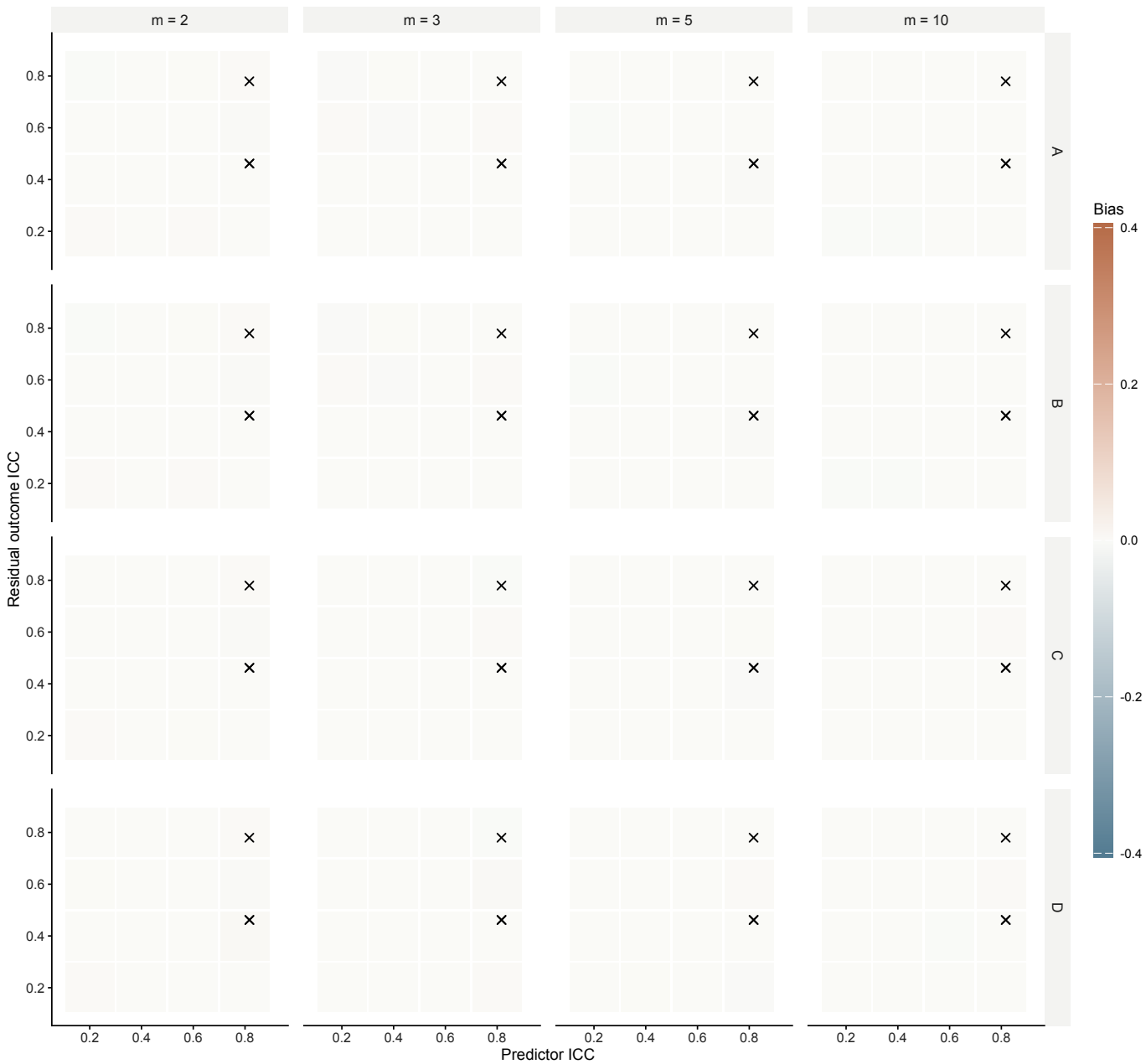
A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 22 of 24

Bias surface | N = 500; true within slope = 0.500

A: OLS | B: clustered OLS | C: random intercept | D: within-between



Cells show simulated grid values; no interpolation. Crosses: empirical ICC context (null-effect calibration).
Empirical design/authenticity markers nearly coincide. Markers are outside the ICC[x] grid; exact anchors are separate.
Supplementary page 23 of 24

95% CI coverage surface | N = 500; true within slope = 0.500

A: OLS | B: clustered OLS | C: random intercept | D: within-between

