Online supplemental appendix for *The versatility of SpAM: A fast, efficient spatial method of data collection for multidimensional scaling,* by Michael C. Hout, Stephen D. Goldinger, and Ryan W. Ferguson.

Table of Contents

Section *Page Number*

Section 1. *Stimuli used in Experiments 2 and 3.*

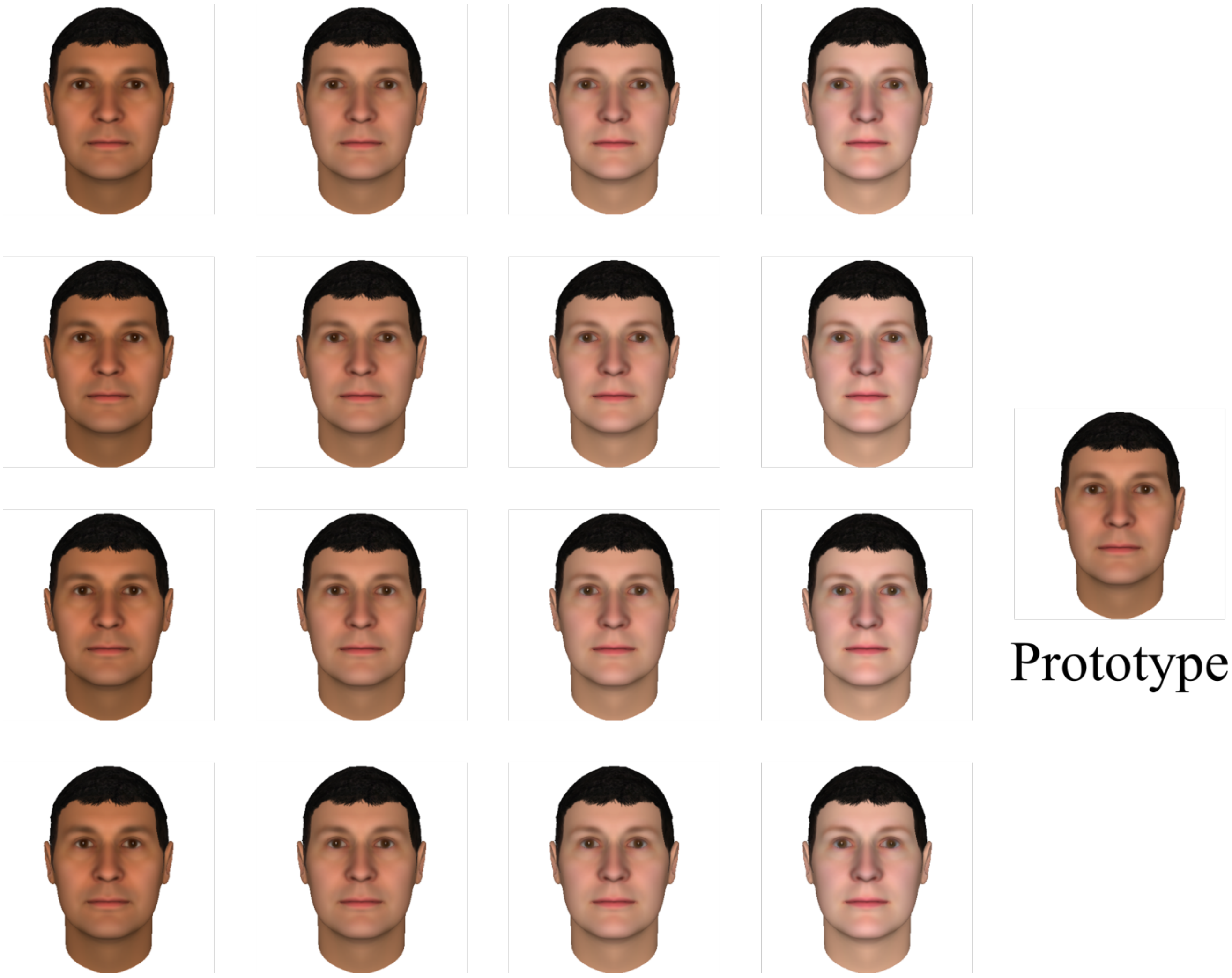
In Experiment 2, we used two sets of animal names (see Table A1). The first set, dubbed the “categorical” animals, were obtained from Hornberger et al. (2009), and were selected such that each animal could be easily categorized along two dimensions. Each was either a bird or four-legged animal (*avian* dimension), and was either a land or water dweller (*habitat* dimension). The second set of names, referred to as the “continuous” animals, were obtained from Henley (1969), and were selected with no obvious categorical classification nor pre-specified underlying structure.

In Experiment 3, we used two sets of stimuli. The first were a subset of the two-dimensional bugs, used in Experiment 1. The second set of stimuli (see Figure A1) were faces created using FaceGen Modeller (Singular Inversions, 2004) software. They were created by first generating a racially ambiguous, male face, for use as a prototype. We then systematically distorted the prototype along two dimensions: the skin shade, and separation of the eyes (varied in equal steps).

Table A1. *All stimuli used in Experiment 2. The left column shows the categorical animals, and the right shows the continuous items.*

Categorical	**Continuous**
Woodpecker	Dog
Pheasant	Cat
Turkey	Mouse
Ostrich	Cow
Pigeon	Horse
Chicken	Lion
Penguin	Bear
Duck	Deer
Swan	Sheep
Goose	Pig
Seagull	Gorilla
Pelican	Badger
Badger	Camel
Squirrel	Elephant
Hedgehog	Kangaroo
Elephant	Koala
Bear	Raccoon
Buffalo	Donkey
Wolf	Squirrel
Turtle	Zebra
Frog	Giraffe
Crocodile	Rhinoceros
Crab	Rabbit
Beaver	Tortoise
Dolphin	Buffalo

Figure A1. Face stimuli used in Experiment 3. The prototype was not used shown to participants, but was simply used to generate a range of distortions varying in skin shade and separation of the eyes.



Section 2. *Supplemental multidimensional scaling solutions.*

In Experiment 1, we tested the ability of four different techniques (pairwise, SpAM, total-set, and triad) to uncover the three-dimensions of our wheel and bug stimuli. Data obtained from each method was used to create three-dimensional multidimensional scaling (MDS) plots, shown in Figures A2-A5. The solutions are shown in side-by-side two-dimensional plots; the first showing dimension 1 against dimension 2, and the second showing dimension 1 against dimension 3.

In Experiment 3, we used the pairwise and SpAM techniques to obtain similarity estimates for a subset of the two-dimensional bugs, and new, computer-generated two-dimensional faces. Figure A6 shows the two-dimensional MDS solutions for each technique.

Figure A2. Three-dimensional MDS spaces generated by the pairwise method in Experiment 1.

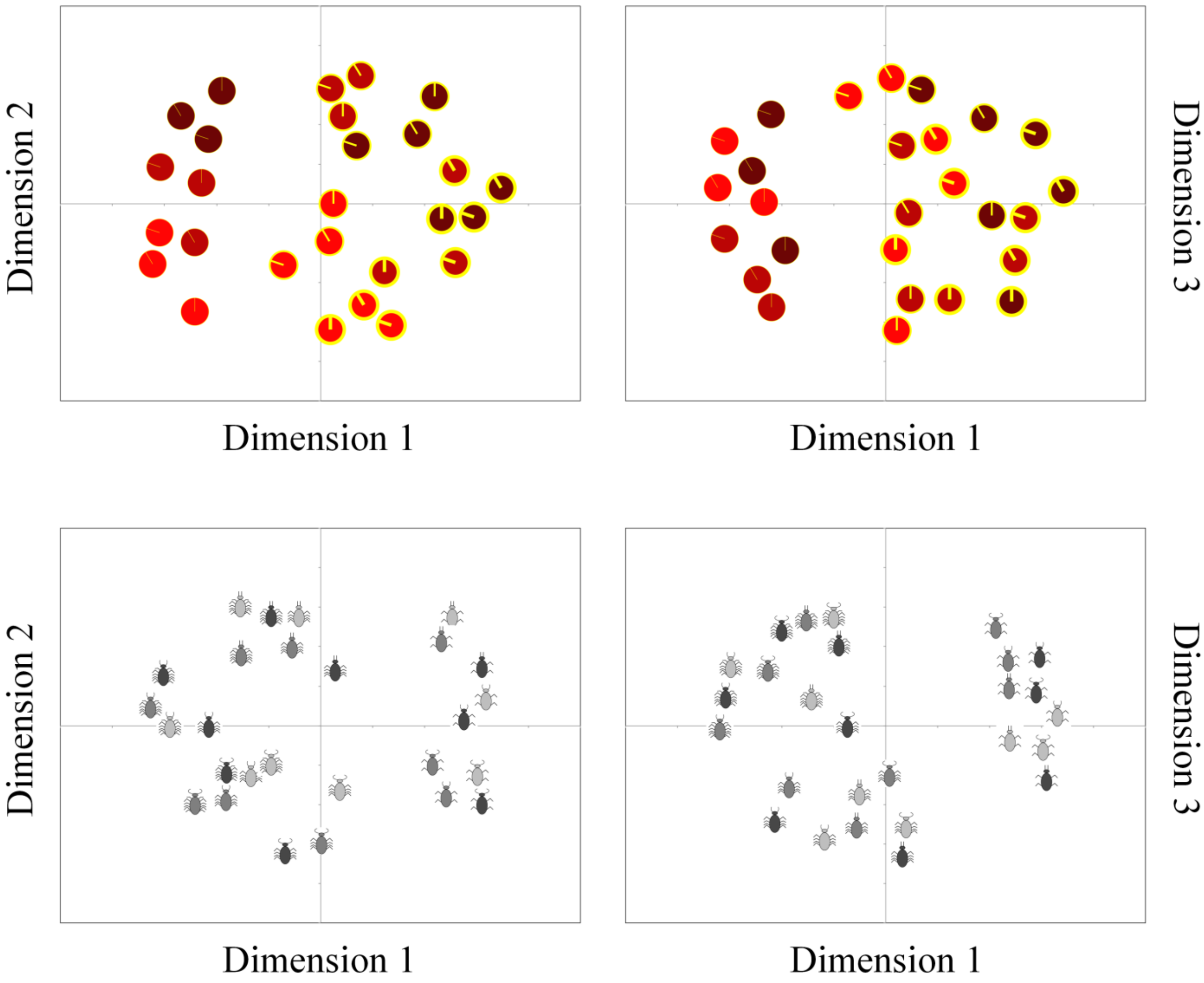


Figure A3. Three-dimensional MDS spaces generated by the spatial arrangement method in Experiment 1.

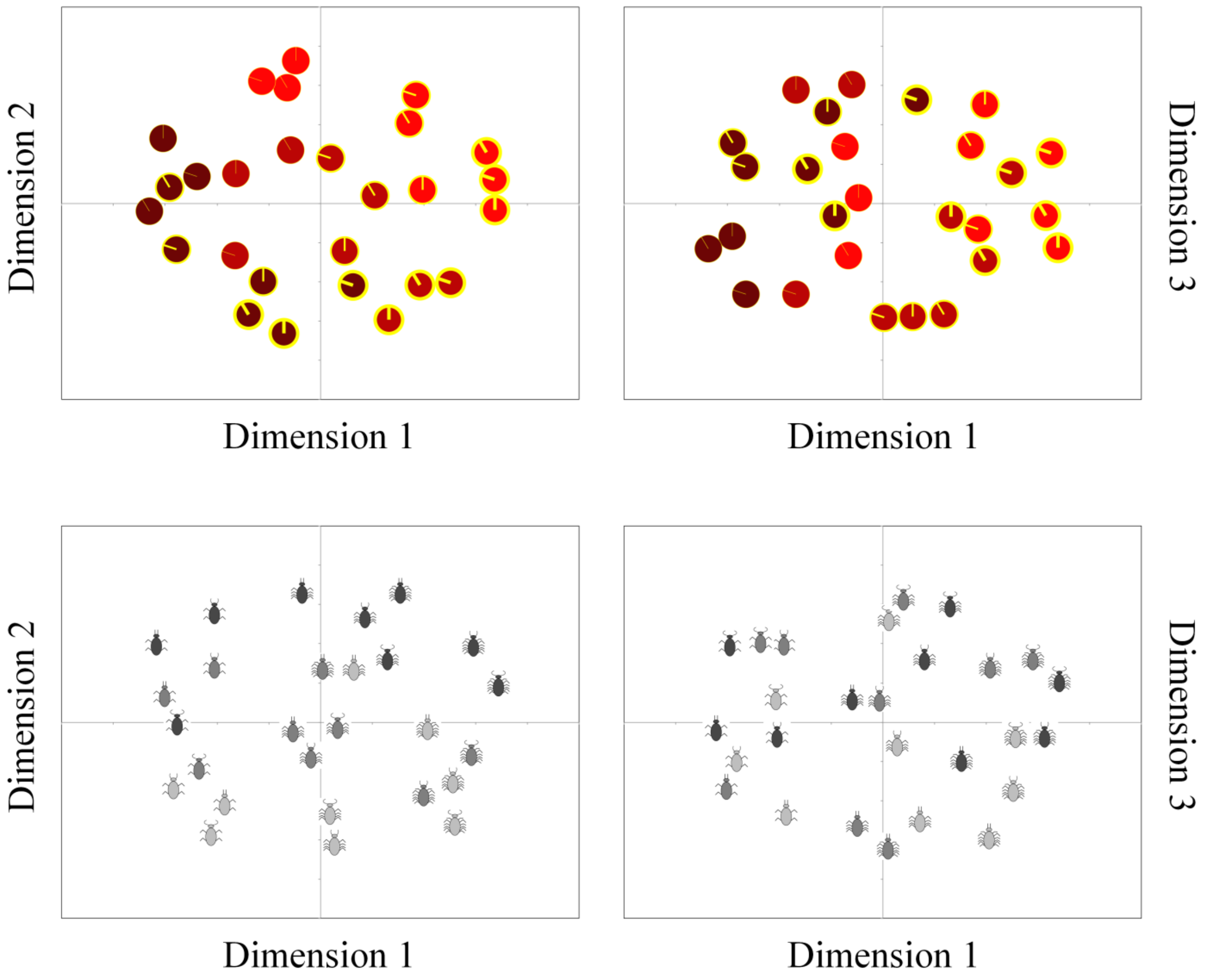


Figure A4. Three-dimensional MDS spaces generated by the total-set pairwise method in Experiment 1.

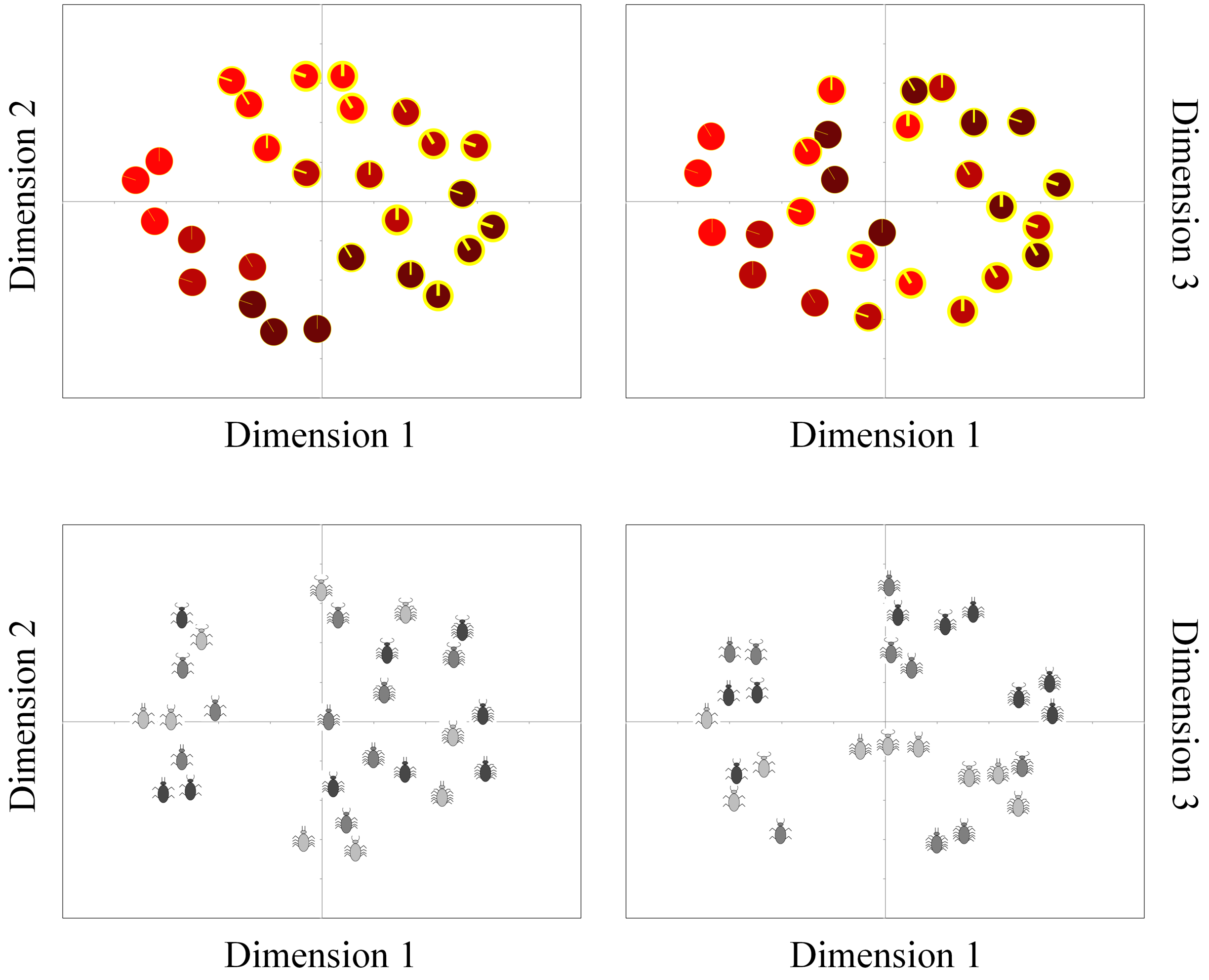


Figure A5. Three-dimensional MDS spaces generated by the triad method in Experiment 1.

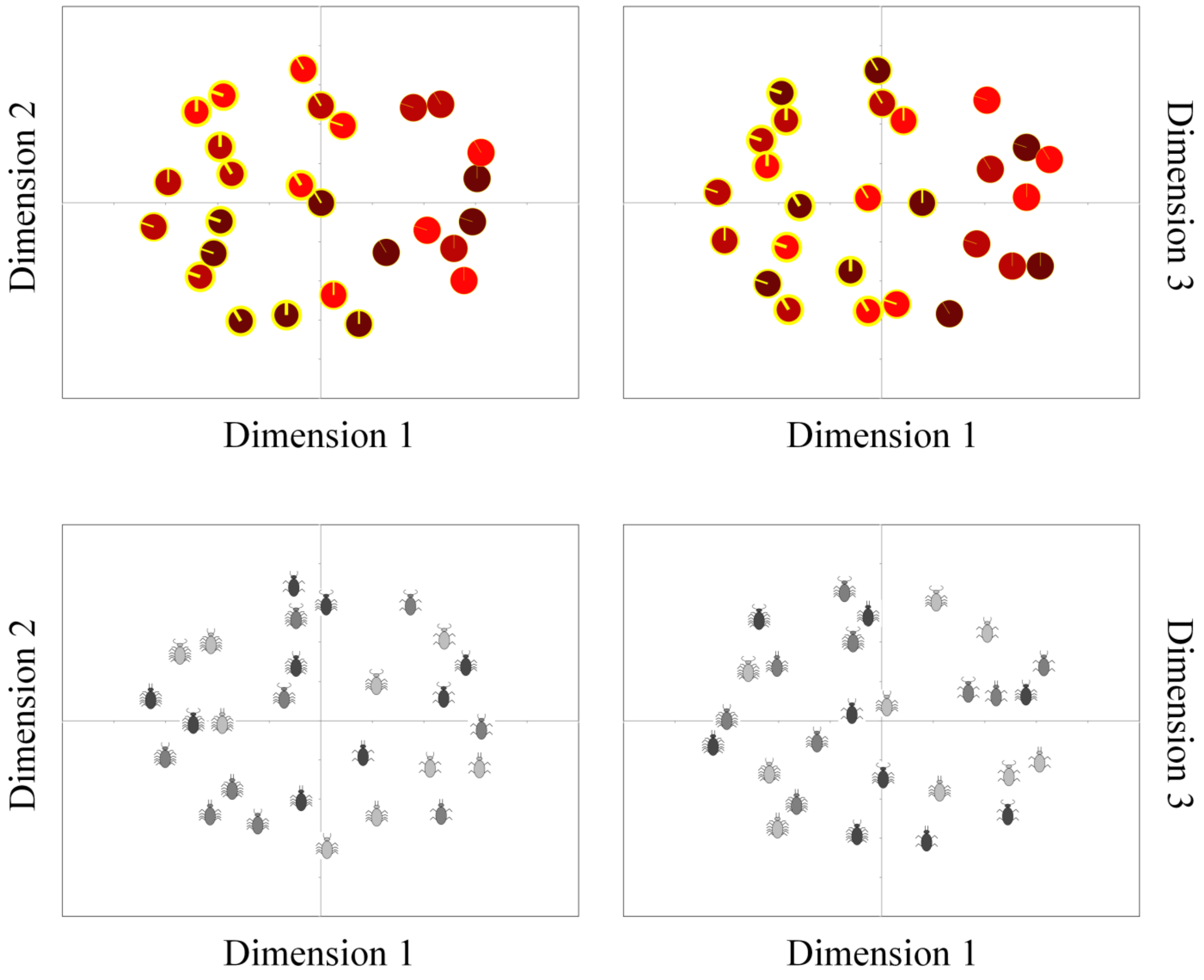
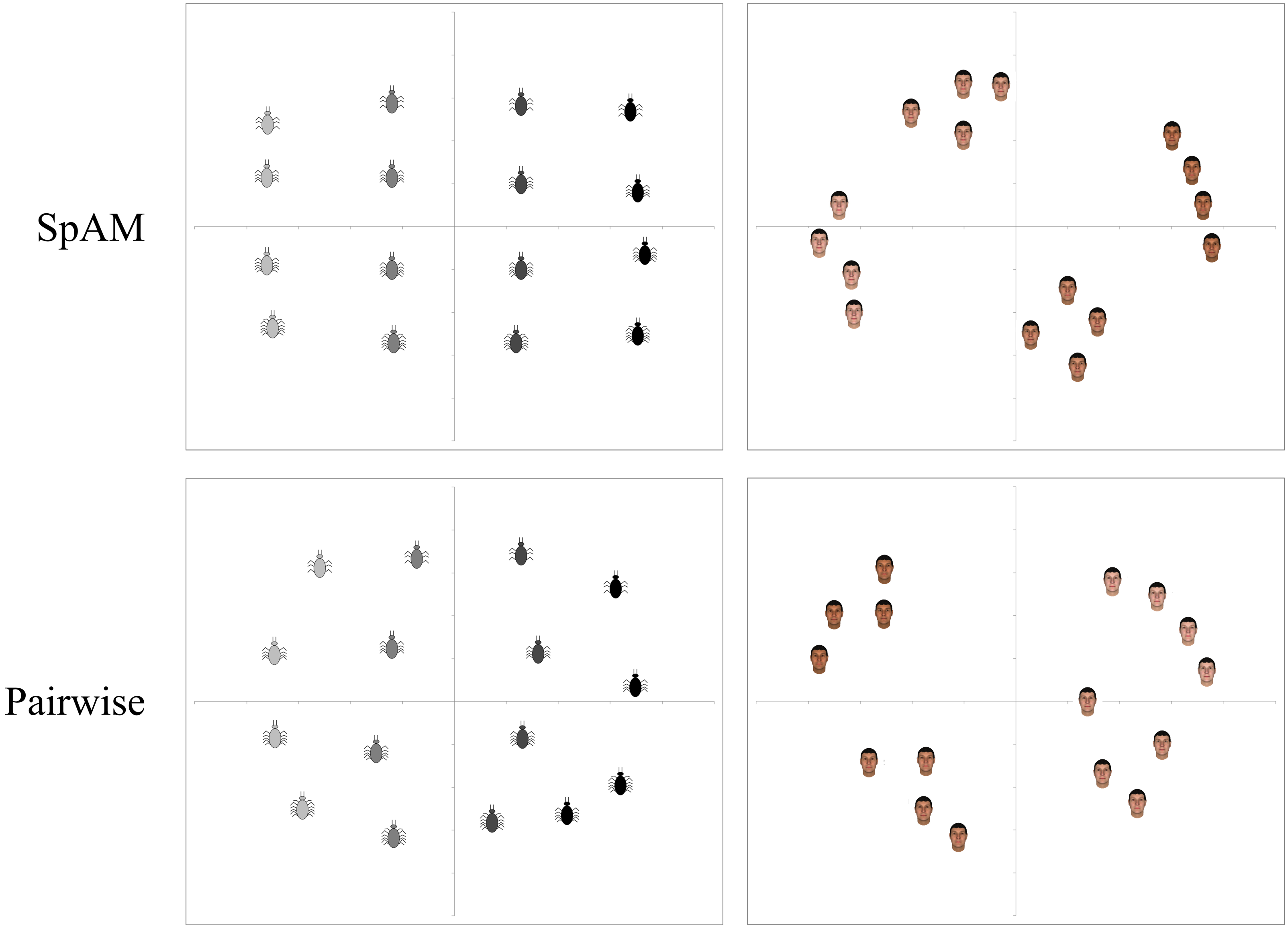


Figure A6. MDS spaces for two-dimensional bugs and faces, derived by SpAM (top panels) and pairwise (bottom panels) methods, from Experiment 3.



Section 3. *Monte Carlo simulations.*

In Experiments 1 and 2, we performed a series of simulations. Because each MDS solution is unique, the orderly solutions we obtained may simply have been fortuitous. In order to address this possibility, we performed Monte Carlo simulations on unmodified data from pairwise and SpAM techniques, and on modified SpAM data. In the latter case, we attempted to isolate characteristics of SpAM that elicit its high-quality solutions, by systematically stripping the SpAM data of its granularity and data mass. In the "reduced granularity" data, we converted the raw pixel values to rounded, single-digit integers (akin to scores obtained via Likert scales). For the "reduced subjects" data, we randomly sampled a subset of participants' data matrices to match the sample sizes obtained by pairwise methods. Finally, in "both reduced" data, both modifications were applied.

We applied scaling algorithms to the data 25 times each, and first examined the consistency of the solutions across iterations of the scaling algorithm. To do this, we calculated the inter-item distances from each solution, and correlated them within-methodologies. We asked to what degree the solutions generated by one data type were consistent with one another; high positive correlations indicate stability within a data set. We next correlated the inter-item distances from each simulation with those of the pairwise method simulations, essentially using the pairwise data as a baseline for comparison. This tells us how well the SpAM data correlates with pairwise solutions across multiple iterations, and how degradation of the SpAM data affects the agreement of the solutions. The tables below (see Tables A2-A3) first report within- and cross-method correlation coefficient frequencies and the percentage of significant correlations for each simulation and stimulus type, from Experiments 1 and 2. The adjacent histograms (see Figures A7-A10) provide a graphical representation of these findings.

Next, we used the Monte Carlo simulations to examine how closely and consistently the solutions conformed to our "ideal spaces" from Experiment 1. Because our wheel and bug stimuli were constructed with specific dimensions, it was possible to derive an ideal space for comparison to solutions derived from each method. These ideal spaces had perfect, orderly arrangements of stimulus items, with equal intervals between levels of each dimension (in essence, perfect squares or cubes). Deviation scores were calculated by taking the Euclidean distance from the *PROXSCAL* coordinates to the ideal location for each stimulus item. Low deviations indicate high conformity to the ideal solutions. Table A4 shows the binned deviation frequencies for each simulation technique and stimulus set. Figure A11 presents a graphical representation of the findings.

In Experiment 2, we examined how well each MDS solution uncovered the hypothesized underlying categorical structures in our animal stimuli. We calculated distance scores that measured the average item-to-item distance from each stimulus item to: 1) members of its own category; 2) items that matched on only the habitat dimension; 3) items that matched on only the avian dimension; and 4) items that were opposite on both dimensions. Solutions with consistent categorization should small within-category distances, large distances to items that are opposites on both dimensions, and intermediate values for items that share singular features. Table A5 presents the binned distance frequencies for each simulation, and Figure A12 presents a graphical representation of the findings.

Table A2. *Binned within- and cross-method correlation frequencies, and the percentage of significant correlations for each simulation, as a function of stimulus type, from Experiment 1.*

		Within-method correlation coefficients																					
		Two dimensional											**Three dimensional**										
Method	Stimuli	Percent Significant	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	Percent Significant	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Pairwise	Wheels	100%	0	0	0	0	50	162	61	22	5	0	100%	0	0	0	18	122	121	35	4	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	300	100%	0	0	0	0	0	6	96	166	32	0
SpAM	Wheels	100%	0	0	0	0	0	0	0	0	0	300	98%	12	35	74	67	61	39	10	2	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	300	100%	0	0	42	65	61	83	44	4	1	0
Reduced Granularity	Wheels	100%	0	0	0	0	0	0	0	0	0	300	71%	100	124	68	7	1	0	0	0	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	300	71%	101	121	60	15	3	0	0	0	0	0
Reduced Subjects	Wheels	100%	0	0	8	10	69	55	10	1	9	138	99%	8	37	39	33	55	53	58	17	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	300	96%	11	16	21	69	80	20	24	53	6	0
Both Reduced	Wheels	100%	0	0	0	22	87	10	0	0	16	165	85%	51	82	84	36	24	18	4	1	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	300	84%	51	44	61	61	40	20	13	9	1	0
		Cross-method correlation coefficients																					
		Two dimensional											**Three dimensional**										
Method	Stimuli	Percent Significant	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	Percent Significant	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
SpAM	Wheels	100%	0	0	0	8	297	170	4	62	59	25	100%	0	29	92	159	187	126	32	0	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	625	100%	0	0	9	300	209	89	16	2	0	0
Reduced Granularity	Wheels	100%	0	0	0	2	285	188	2	67	57	24	99%	7	132	304	141	34	7	0	0	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	625	94%	50	89	306	166	14	0	0	0	0	0
Reduced Subjects	Wheels	91%	57	41	65	49	174	128	18	33	43	17	99%	5	22	61	70	166	201	88	10	2	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	625	99%	4	22	33	70	235	152	67	33	9	0
Both Reduced	Wheels	100%	0	0	0	2	285	188	2	67	57	24	98%	19	89	171	128	124	83	10	1	0	0
	Bugs	100%	0	0	0	0	0	0	0	0	0	625	90%	65	48	70	195	157	48	21	20	1	0

Note: Alpha set at p < .05.

Figure A7. Histograms showing the binned within-method correlation frequencies for each simulation, as a function of stimulus type and dimensionality, from Experiment 1, Monte Carlo simulations.

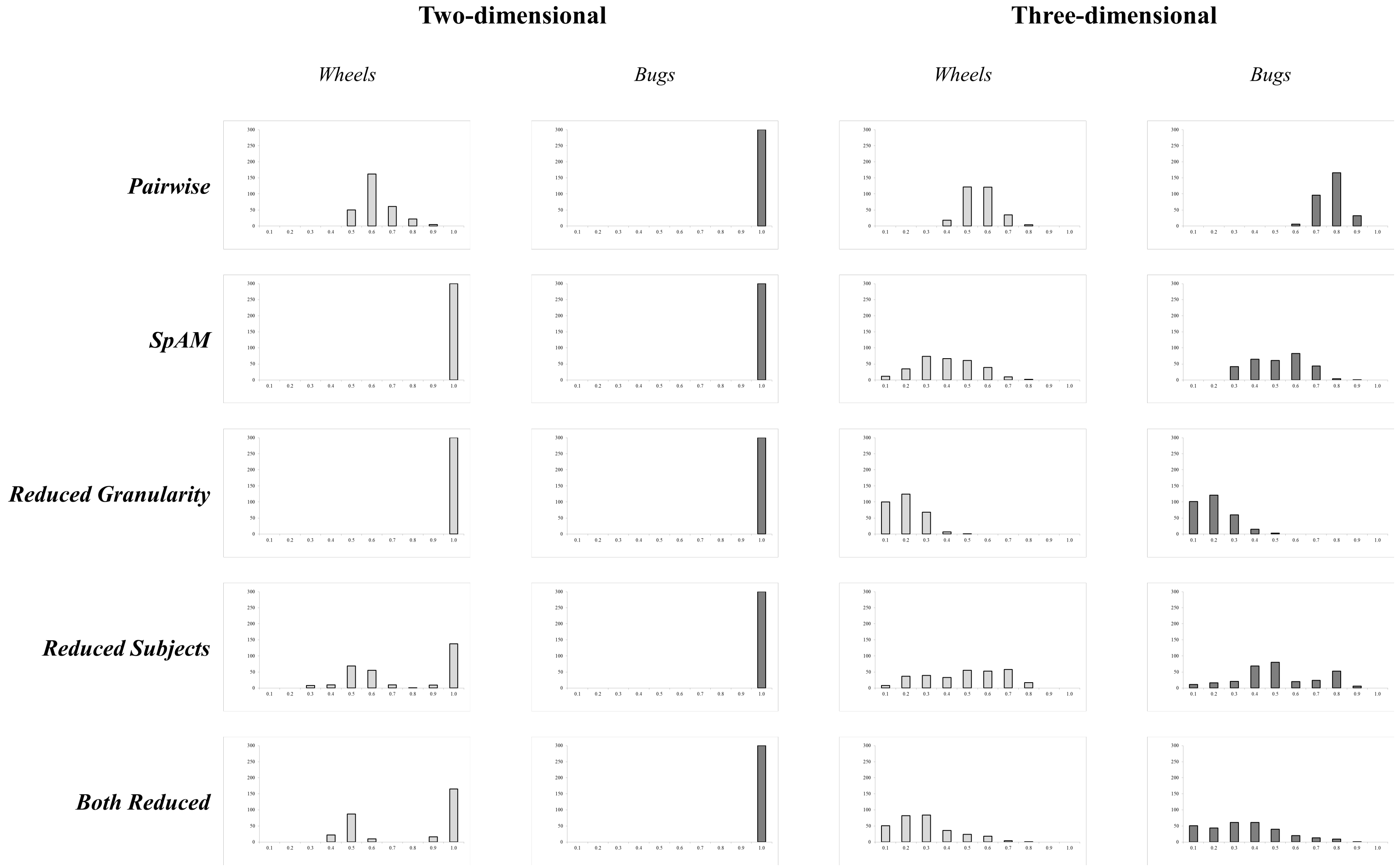


Figure A8. Histograms showing the binned cross-method correlation frequencies for each simulation, as a function of stimulus type and dimensionality, from Experiment 1, Monte Carlo simulations.

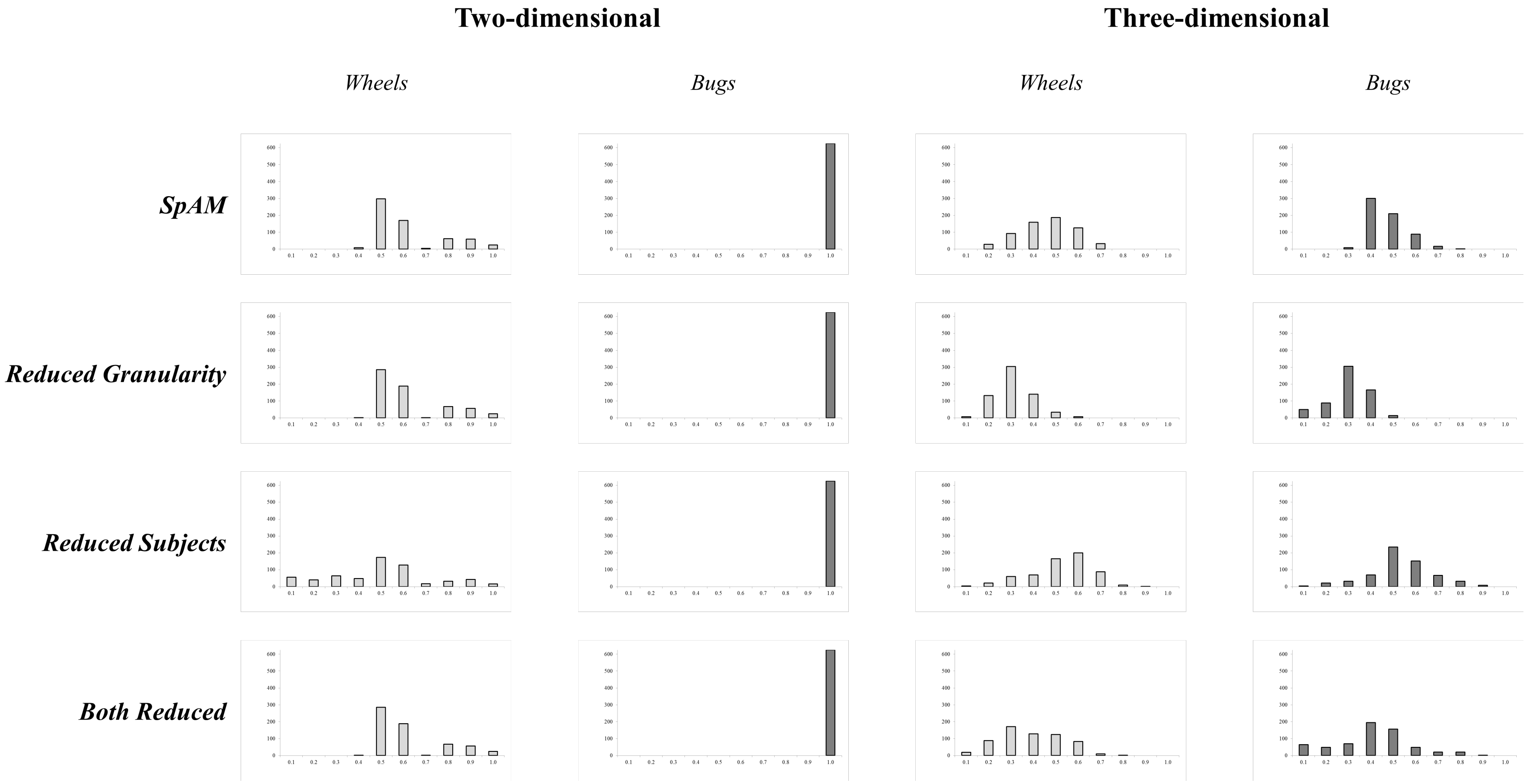


Table A3. *Binned within- and cross-method correlation frequencies, and the percentage of significant correlations for each simulation, as a function of stimulus set, from Experiment 2.*

			Within-method correlation coefficients									
Method	Stimuli	Percent Significant	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Pairwise	Categorical	100%	0	0	0	0	0	4	66	143	82	5
	Continuous	100%	0	0	0	1	25	168	82	21	3	0
Spatial	Categorical	100%	0	0	0	0	0	0	3	118	135	44
	Continuous	100%	0	0	4	58	125	89	17	5	2	0
Reduced Granularity	Categorical	100%	0	0	0	0	0	0	11	78	196	15
	Continuous	96%	14	86	106	66	27	1	0	0	0	0
Reduced Subjects	Categorical	100%	0	0	0	0	0	5	75	139	78	3
	Continuous	100%	0	8	34	70	95	54	34	5	0	0
Both Reduced	Categorical	100%	0	0	0	0	0	27	101	129	42	1
	Continuous	98%	8	36	112	99	31	11	2	1	0	0

			Cross-method correlation coefficients									
Method	Stimuli	Percent Significant	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Spatial	Categorical	100%	0	0	0	0	0	14	282	248	81	0
	Continuous	100%	0	1	21	131	288	158	24	2	0	0
Reduced Granularity	Categorical	100%	0	0	0	0	0	40	299	249	37	0
	Continuous	99%	1	44	160	279	119	17	5	0	0	0
Reduced Subjects	Categorical	100%	0	0	0	0	1	92	283	219	30	0
	Continuous	99%	1	10	72	186	214	110	32	0	0	0
Both Reduced	Categorical	100%	0	0	0	0	0	94	295	212	24	0
	Continuous	100%	0	12	210	235	112	46	9	1	0	0

Note: Alpha set at $p < .05$.

Figure A9. Histograms showing the binned within-method correlation frequencies for each simulation, as a function of stimulus set, from Experiment 2, Monte Carlo simulations.

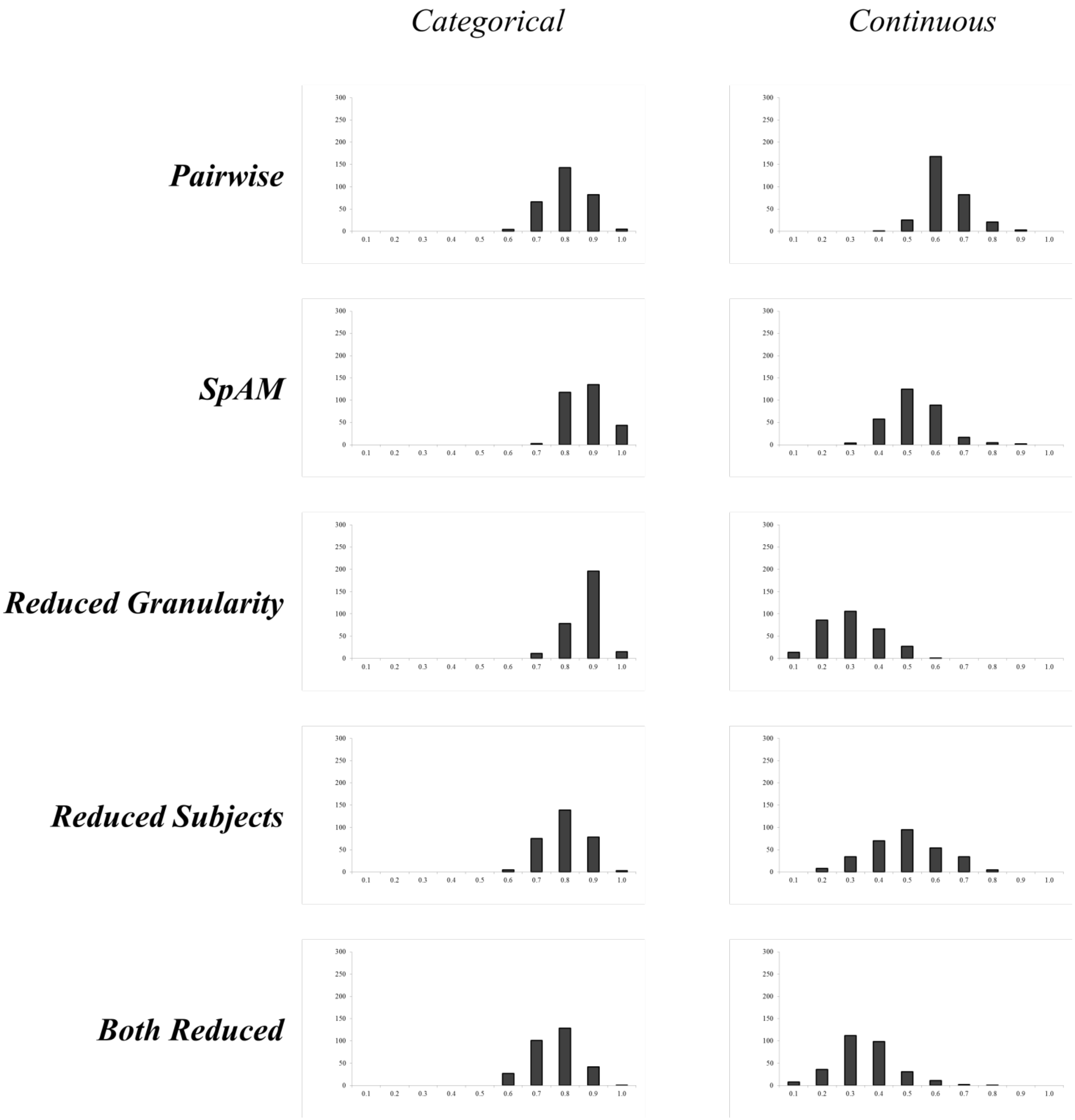


Figure A10. Histograms showing the binned cross-method correlation frequencies for each simulation, as a function of stimulus set, from Experiment 2, Monte Carlo simulations.

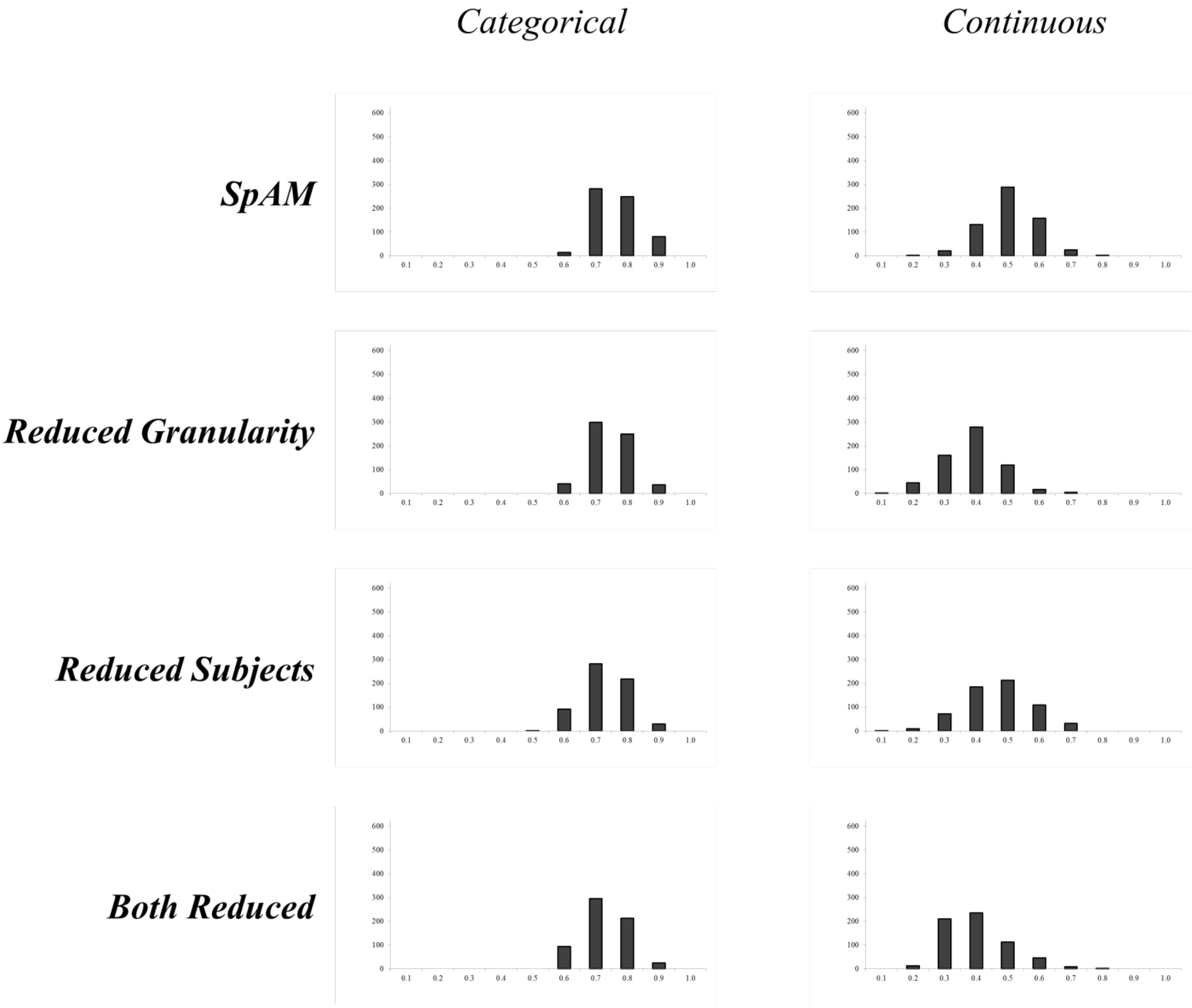


Table A4. *Binned deviation frequencies for each simulation, as a function of stimulus type and dimensionality, from Experiment 1, Monte Carlo simulations.*

		Two dimensional																			
Method	Stimuli	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
Pairwise	Wheels	55	91	126	93	49	41	46	23	23	30	16	11	8	7	5	1	0	0	0	0
	Bugs	103	167	202	103	39	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0
SpAM	Wheels	512	105	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Bugs	540	85	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Reduced Granularity	Wheels	459	119	34	7	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Bugs	525	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Reduced Subjects	Wheels	263	128	63	38	22	20	20	12	10	16	13	7	2	4	4	3	0	0	0	0
	Bugs	311	171	65	42	17	11	4	4	0	0	0	0	0	0	0	0	0	0	0	0
Both Reduced	Wheels	271	136	52	53	20	18	20	10	7	17	5	5	2	2	5	1	1	0	0	0
	Bugs	266	174	88	50	27	10	6	2	2	0	0	0	0	0	0	0	0	0	0	0

		Three dimensional																			
Method	Stimuli	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
Pairwise	Wheels	9	17	46	66	101	119	104	74	45	27	31	18	15	1	2	0	0	0	0	0
	Bugs	9	22	66	82	114	111	82	51	36	24	18	17	26	7	6	2	2	0	0	0
SpAM	Wheels	3	12	39	58	78	100	91	67	51	43	40	37	27	17	7	4	1	0	0	0
	Bugs	20	19	63	98	104	117	67	49	33	24	29	26	11	10	4	1	0	0	0	0
Reduced Granularity	Wheels	1	6	23	39	69	80	92	72	66	56	59	35	39	17	10	8	2	1	0	0
	Bugs	4	8	29	32	71	98	98	78	63	39	32	41	35	22	13	10	1	1	0	0
Reduced Subjects	Wheels	4	22	52	68	93	104	86	72	48	32	27	24	24	6	6	6	0	1	0	0
	Bugs	15	36	88	93	108	94	55	46	38	27	28	21	13	5	6	1	0	1	0	0
Both Reduced	Wheels	5	17	36	77	87	77	84	46	50	44	57	37	33	11	6	4	3	1	0	0
	Bugs	10	22	47	79	88	86	73	58	41	48	35	37	31	10	5	2	3	0	0	0

Figure A11. Histograms showing the binned deviation frequencies for each simulation, as a function of stimulus type and dimensionality, from Experiment 1, Monte Carlo simulations. Along the X-axis are the upper bounds of each “bin” (“0.1” indicates the frequency of deviations ranging from 0.0 – 0.1, and so on).

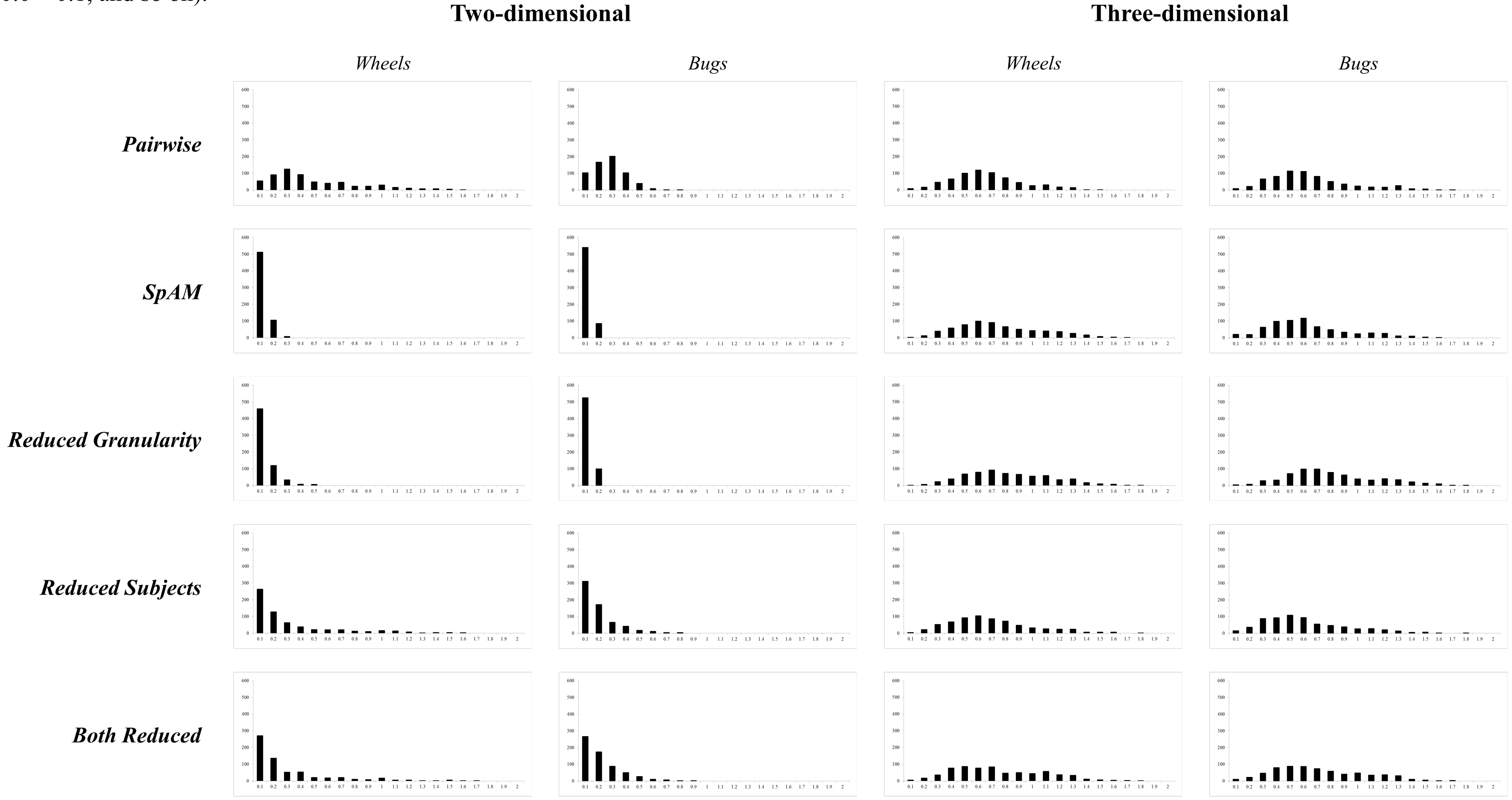
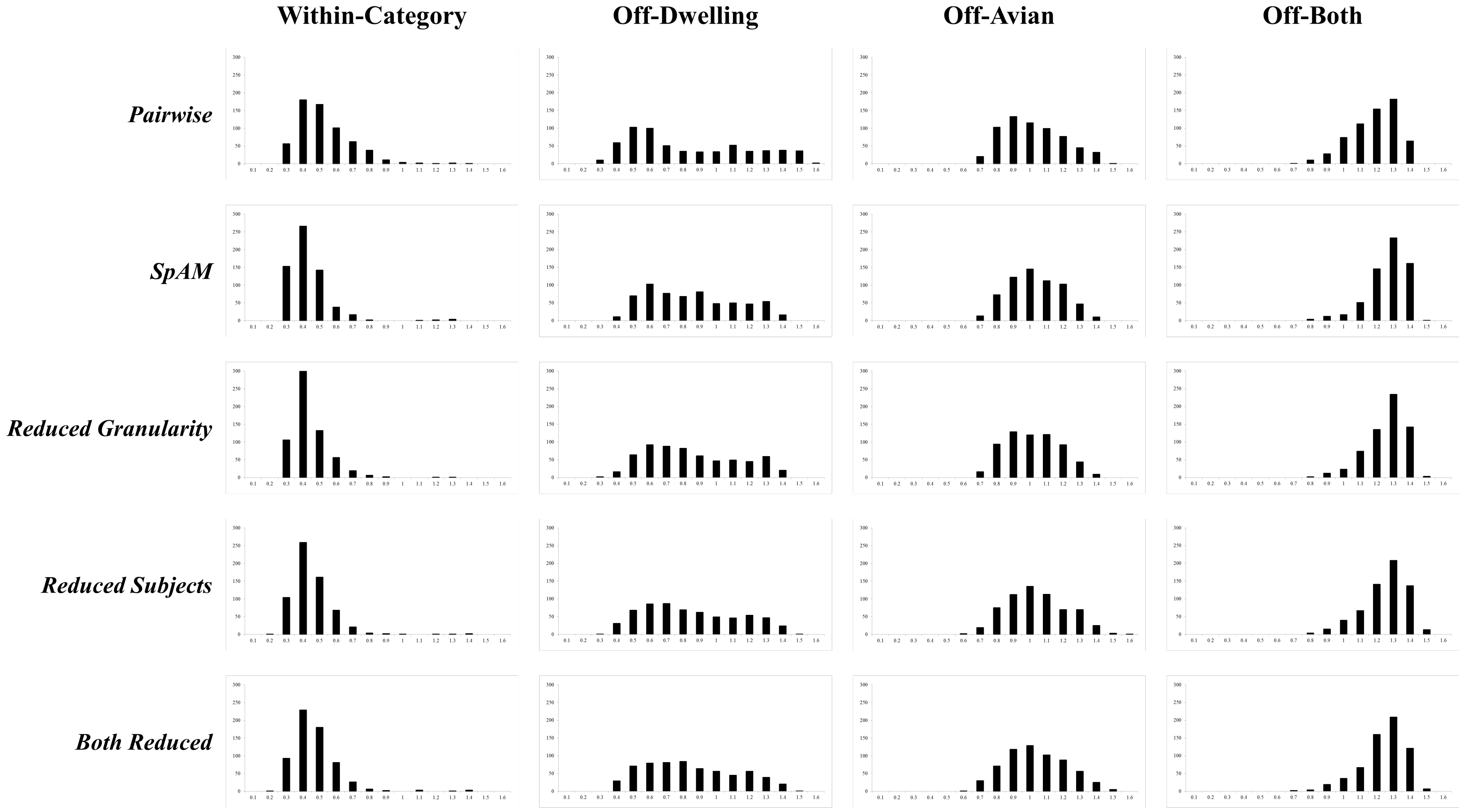


Table A5. *Binned distance frequencies (presented as a function of featural dissimilarity) for each simulation, from Experiment 2, Monte Carlo simulations (categorical animals).*

Method	Stimuli	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.3	1.4	1.5	1.6
Pairwise	Within-Category	0	0	56	180	167	101	62	38	11	4	2	1	2	1	0	0
	Off-Dwelling	0	0	10	59	103	100	51	35	33	34	52	35	37	38	36	2
	Off-Avian	0	0	0	0	0	0	20	103	133	115	99	77	45	32	1	0
	Off-Both	0	0	0	0	0	0	1	10	28	74	112	154	182	64	0	0
Spatial	Within-Category	0	0	153	266	142	38	17	2	0	0	1	2	4	0	0	0
	Off-Dwelling	0	0	0	11	70	103	77	68	81	48	50	47	54	16	0	0
	Off-Avian	0	0	0	0	0	0	13	73	122	145	112	103	47	10	0	0
	Off-Both	0	0	0	0	0	0	0	4	12	17	51	146	233	161	1	0
Reduced Granularity	Within-Category	0	0	106	302	132	56	19	6	2	0	0	1	1	0	0	0
	Off-Dwelling	0	0	2	16	64	92	88	82	61	47	49	45	59	20	0	0
	Off-Avian	0	0	0	0	0	0	16	94	129	120	121	92	44	9	0	0
	Off-Both	0	0	0	0	0	0	0	2	12	23	74	135	234	142	3	0
Reduced Subjects	Within-Category	0	1	104	259	161	68	21	4	2	1	0	1	1	2	0	0
	Off-Dwelling	0	0	1	31	68	86	87	69	62	49	46	54	47	24	1	0
	Off-Avian	0	0	0	0	0	2	19	75	112	135	113	70	70	25	3	1
	Off-Both	0	0	0	0	0	0	0	4	15	40	67	141	208	137	13	0
Both Reduced	Within-Category	0	1	93	229	180	81	26	6	2	0	3	0	1	3	0	0
	Off-Dwelling	0	0	0	29	71	79	81	84	64	56	45	56	39	20	1	0
	Off-Avian	0	0	0	0	0	1	30	71	118	129	102	88	56	25	5	0
	Off-Both	0	0	0	0	0	0	2	4	19	36	67	160	209	121	7	0

Figure A12. *Histograms showing the binned distance frequencies for each simulation, as a function of featural dissimilarity, from Experiment 2, Monte Carlo simulations. Along the X-axis are the upper bounds of each "bin" ("0.1" indicates the frequency of deviations ranging from 0.0 – 0.1, and so on).*



Section 4. Individual differences analyses.

In Experiments 1 and 2, we performed individual differences analyses that examined a potential shortcoming of the SpAM technique; namely, that participants may interpret the instructions differentially, or may implement different strategies in order to construct their arrangements. As such, we showed that these potential outliers are not particularly problematic for SpAM, and suggested a way in which to identify "irregular" participants.

Our general strategy was to identify outliers by analyzing the extent to which each participant's MDS space correlated with all others (we did this for SpAM and pairwise data). This entailed several steps: 1) We created individual MDS spaces for each participant, and derived vectors of inter-item distances from those spaces. 2) Next, we correlated the distance vectors across all participants (for each stimulus set and methodology, separately). 3) For each participant, we then calculated two scores: their average correlation coefficient, and the proportion of correlations that were statistically significant. 4) Finally, we rank-ordered the participants, and (in two separate analyses) identified those with the lowest average correlations or proportions of significant correlations. The bottom 25% of participants were identified as outliers.

Once we identified these irregular participants, we created two MDS spaces, one for the entire data set excluding the outliers and another for the outliers themselves. In order to gauge the extent to which these participants skewed the aggregate results, we then correlated the inter-item distances from these exclusionary solutions with the space that included every participant. Table A6 presents the findings, showing the correlation coefficients across aggregate data, and the "regular" and "irregular" solutions (for two- and three-dimensional bugs from Experiment 1, and the categorical animals from Experiment 2).

Table A6. Pearson product-moment correlation coefficients for the inter-item distances across solutions derived from "regular" and "irregular" participants, with respect to aggregate solutions (from Experiments 1 and 2).

Stimuli	Method	SpAM (aggregate)	Regulars (mean r)	Irregulars (mean r)	Regulars (% sign.)
Bugs (2D)	Regulars (mean *r*)	0.99 **			
	Irregulars (mean *r*)	-0.04	-0.05		
	Regulars (% sign.)	0.99 **	0.99 **	-0.04	
	Irregulars (% sign.)	0.18 **	0.17 **	0.02	0.15 **
Bugs (3D)	Regulars (mean *r*)	0.59 **			
	Irregulars (mean *r*)	0.32 **	0.19 **		
	Regulars (% sign.)	0.54 **	0.65 **	0.20 **	
	Irregulars (% sign.)	0.09	0.29 **	0.03	0.13 **
Animals (categorical)	Regulars (mean *r*)	0.91 **			
	Irregulars (mean *r*)	0.12 **	0.19 **		
	Regulars (% sign.)	--	--	--	
	Irregulars (% sign.)	--	--	--	--

Stimuli	Method	Pairwise (aggregate)	Regulars (mean r)	Irregulars (mean r)	Regulars (% sign.)
Bugs (2D)	Regulars (mean *r*)	0.97 **			
	Irregulars (mean *r*)	0.90 **	0.85 **		
	Regulars (% sign.)	0.99 **	0.96 **	0.89 **	
	Irregulars (% sign.)	0.44 **	0.48 **	0.30 **	0.43 **
Bugs (3D)	Regulars (mean *r*)	0.79 **			
	Irregulars (mean *r*)	0.19 **	0.11 *		
	Regulars (% sign.)	0.82 **	0.80 **	0.19 **	
	Irregulars (% sign.)	0.16 **	0.12 *	0.09 *	0.20 **
Animals (categorical)	Regulars (mean *r*)	0.85 **			
	Irregulars (mean *r*)	0.42 **	0.35 **		
	Regulars (% sign.)	--	--	--	
	Irregulars (% sign.)	--	--	--	--

Note: Participants identified as outliers for the categorical animal stimuli were identical across both measures, for both methodologies. Accordingly, correlations are reported only for the average correlation criterion.

* *$p < .05$;* ** *$p < .01$.*