

1 Supplementary Material: Entropy trade-offs in
2 artistic design: A case study of Tamil *kolam*

3 N.-Han Tran, Timothy Waring, Silke Atmaca, Bret A. Beheim

4 **Contents**

5	S1 Approximation of entropy using richness and the Gini index	S2
6	S2 <i>Kolam</i> Data	S5
7	S2.0.1 Lexicon of Gestures	S5
8	S2.1 Database	S7
9	S2.1.1 Survey Information	S7
10	S2.1.2 GPS	S7
11	S2.1.3 <i>Kolam</i> Drawings	S7
12	S3 Statistical Analyses	S8
13	S3.1 Random Intercept Models	S8
14	S3.1.1 Statistical Model	S8
15	S3.1.2 Visual MCMC Diagnostics	S11
16	S3.1.3 Intraclass Correlation (ICC)	S19
17	S3.1.4 Illustration of Random Effect Estimates for Caste	S21
18	S3.2 Gaussian Process Models	S22
19	S3.2.1 Statistical Model	S22
20	S3.2.2 Estimation of Variation	S28

21 **S1** Approximation of entropy using richness and
22 the Gini index

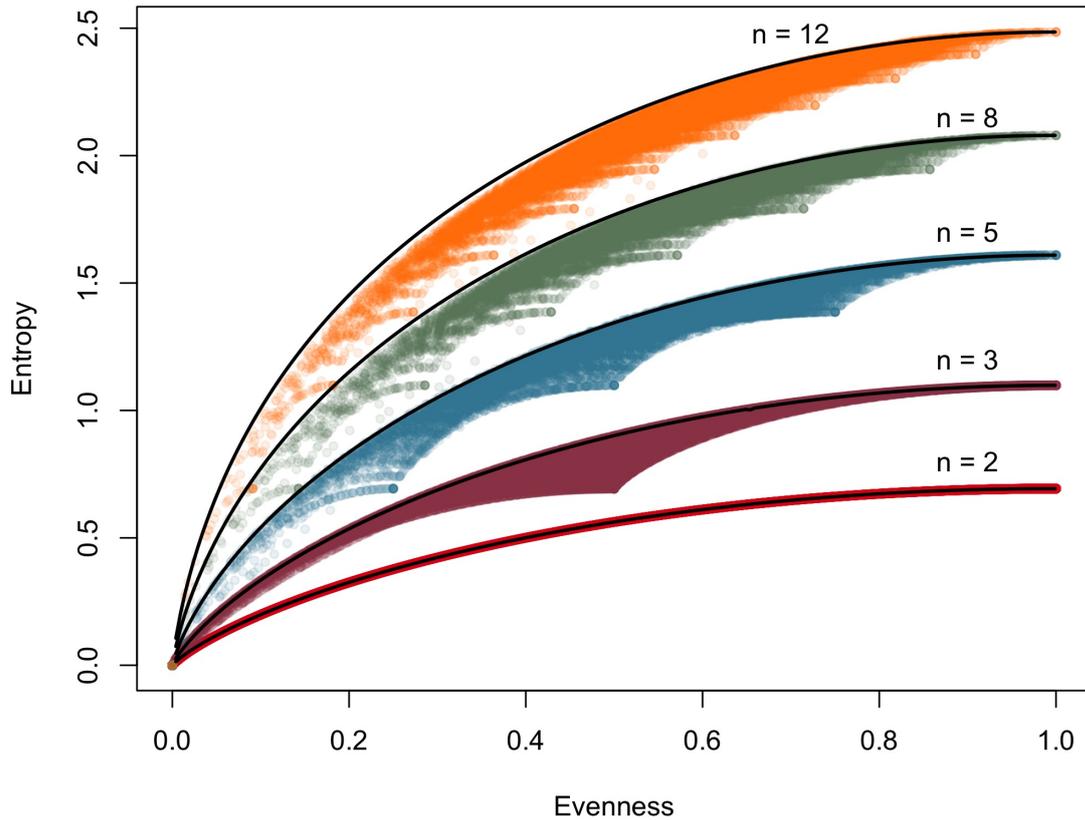


Figure S1: High-resolution simulations showing the entropy distribution of a given richness and evenness. The black lines show the maximum entropy for a given number of variant. The differently coloured points represent the entropy distribution corresponding to the different number of variants. Equation 3 defines the curve for $n = 2$.

23 For any discrete probability distribution with n possible outcomes, each i of which

24 occurs with probability p_i , we can calculate a number of information measures. In
 25 ecology, Shannon information entropy is a popular measure of biological diversity
 26 because it contains two different aspects: richness and evenness. The Shannon infor-
 27 mation entropy is a measure of the expected “surprise” or uncertainty, given in the
 28 discrete case by

$$H(p) = - \sum_{i=1}^n p_i \log(p_i) \quad (1)$$

29 for each outcome i . In economics, the Gini index (Zoli, 1999; Ravallion, 2014) is
 30 used to describe relative inequality in the probability distribution, calculated as the
 31 mean absolute difference between all pair of variants,

$$g(n) = \frac{\sum_{i=1}^n \sum_{j=1}^n |p_i - p_j|}{2(n-1)} \quad (2)$$

32 where n is the richness (the number of unique variants) and p the frequency of
 33 specific variants.

34 Entropy and the Gini index capture variation in the relative abundance of each
 35 outcome following the focal probability distribution. That is, when one outcome
 36 $p_j \rightarrow 1$ and all $p_j \rightarrow 0$, the Shannon entropy goes to its lower asymptote of 0, and
 37 the Gini index goes to its upper asymptote of 1. Likewise, when $p_j = \frac{1}{n}$ for all j , the
 38 entropy is maximized at $\log(n)$ and the Gini index is minimized at 0.

39 In the special case of $n = 2$ with two variant frequencies p and $q = 1 - p$, such
 40 that $p > q$, the Gini index simplifies to $g_2 = p - q$. Using the fact that $p + q = 1$, we
 41 can rewrite each as: $p = \frac{1+g_2}{2}$, $q = \frac{1-g_2}{2}$. Thus, there is an exact relationship between
 42 H_2 and g_2 , such that entropy is maximized when the Gini index is minimized, and
 43 vice versa.

$$H_2 = \left(\frac{1+g}{2}\right) \log\left(\frac{2}{1+g}\right) + \left(\frac{1-g}{2}\right) \log\left(\frac{2}{1-g}\right) \quad (3)$$

44 .

45 Equation 3 defines the curve for $n = 2$ in Figure S1.

46 However, the relationship between a Gini index and an entropy is indefinite if
 47 $n > 2$ because multiple distributions with the same n with different entropy could
 48 take the same Gini index value. Figure S1 illustrates the relationship between the
 49 entropy and Gini index calculated for 100,000 simulated probability distributions.

50 To understand the relationship between entropy, the Gini index and the richness
 51 further, the location of the minimum and maximum entropy within this wing-shaped
 52 “envelope” are important. Given any particular value of g , and number of variants n ,
 53 we can describe the range of possible distributions between a maximum and minimum
 54 entropy. Although analytic solutions for the minimum or maximum entropy are
 55 elusive for $n > 3$, numerical solutions are readily available using simulation and non-
 56 linear optimization algorithms. In the supplementary codebase, we use the `Rsolnp`
 57 package to run the non-linear optimization algorithm and find maximum-entropy
 58 solutions for the cases in Figure S1.

59 The lower “tips” of each distribution in Figure S1 represent the minimum-entropy
 60 limits at which the least-common non-zero variant tends to a zero frequency. In
 61 the case of $n = 3$, the minimum-entropy boundary can be approximated by noting
 62 that the minimum entropy (i.e., the “tip”) occurs when $g = \frac{1}{2}$. At a closer look at
 63 Figure S1, the minima of the entropy distribution lie at $\{\frac{0}{n-1}, \frac{1}{n-1}, \dots, \frac{n-1}{n-1}\}$, following
 64 a limiting boundary described by the equation:

$$H_{min} = \log(n - (n - 1)g). \quad (4)$$

65 More precisely, the minimum entropy boundaries can be defined by generalizing the
 66 entropy equation for $n = 2$ in equation 3 to: $p = \frac{y-a}{b-a}, q = \frac{b-y}{b-a}$, where $v = 1 - g$ in the
 67 special case of $n = 2, a = 0$ and $b = 2$. Empirical exploration of the entropy envelope
 68 by the simulation of 100,000 gesture distributions indicates that no distribution can
 69 occupy a lower entropy than defined by this boundary. This can be used to decompose
 70 entropy using g and n because if g is known, $H_{n,g}$ varies between boundaries defined
 71 by the equation:

$$\exp(\widehat{H}) \approx n - (n - 1)v^{1+b} = n - (n - 1)v^{1+\frac{2}{2+a}+\frac{a}{a+n}}, \quad (5)$$

72 where evenness is $v = 1 - g$. If $b = 0$, this equation gives the lower theoretical
73 boundary on entropy for any given g and n described by equation 4 and the “tips” in
74 Figure S1. The numerically-calculated maximum entropy values are described by the
75 equation $b = \frac{2}{2+a} + \frac{a}{a+n}$ with $a \approx \exp(0.5139)$. Thus, if we assume that a distribution
76 tends towards maximum entropy (Frank, 2009), we can calculate a distribution’s
77 entropy knowing only its richness n and the evenness $v = 1 - g$.

78 As evenness and richness are now related to entropy by a mathematical identity
79 (and so have no single causal direction to their relationship), we can aspire to un-
80 derstand what actually could explain why the observed *kolam* patterns follow the
81 entropy isoclines.

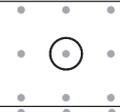
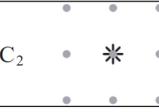
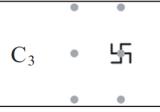
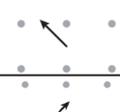
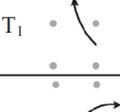
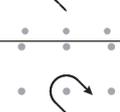
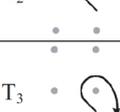
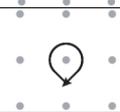
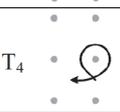
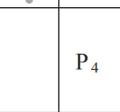
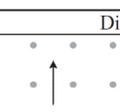
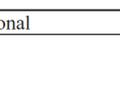
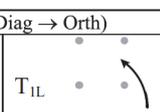
82 S2 *Kolam* Data

83 S2.0.1 Lexicon of Gestures

84 One specific class of *kolam* are those with loop patterns, called square loop *kolam*
85 drawings (i.e., the *ner pulli nelevu* or *sikku kolam* family) (Waring, 2012). These
86 *kolam* drawings are composed of an initial grid of dots (*pulli*) that reflect the canvas
87 size. Lines consisting of multiple gestures are sequentially drawn around the dots to
88 form loops. We only focus on these square loop *kolam* drawings because the patterns
89 can be mapped onto a small identifiable set of gestures which is suitable for analyses.

90 The geometry of the *kolam* can be divided in two fundamental geometric spaces
91 and a transitional geometric space with their specific corresponding positions, orienta-
92 tions and gestures. All gestures are always located in relation to the neighboring dots,
93 called *pulli*. For a detailed description of the sequential encoding of *kolam* drawings
94 with the gestural lexicon, please consult Waring (2012).

(a)

Orthogonal	Transitional (Orth → Diag)	Stylistic variations	
C ₁ 		C ₂ 	C ₃ 
O ₁ 	T ₁ 		
O ₂ 	T ₂ 		
O ₃ 	T ₃ 	H _{3R} 	H _{3L} 
O ₄ 	T ₄ 	P ₄ 	

(b)

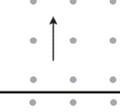
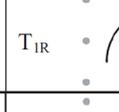
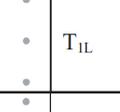
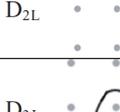
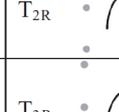
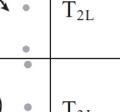
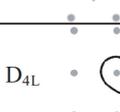
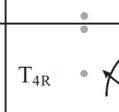
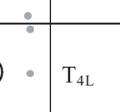
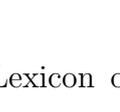
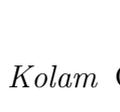
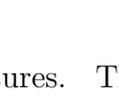
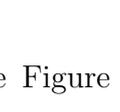
Diagonal		Transitional (Diag → Orth)	
D ₁ 		T _{1R} 	T _{1L} 
D _{2R} 	D _{2L} 	T _{2R} 	T _{2L} 
D _{3R} 	D _{3L} 	T _{3R} 	T _{3L} 
D _{4R} 	D _{4L} 	T _{4R} 	T _{4L} 

Figure S2: The Lexicon of *Kolam* Gestures. The Figure illustrates the gestures and the corresponding code to encode *kolam* drawings. Taken and adapted with permission from [Waring \(2012\)](#)

95 **S2.1 Database**

96 **S2.1.1 Survey Information**

97 Information on individuals' *kolam* drawing abilities and behaviour were gathered as
98 well as demographic information. Demographic information entailed data such as
99 GPS data of the current residency, marriage status, native places and measures of
100 socio-economic status. To investigate individual's *kolam* practice, survey questions
101 included information on individual's frequency of drawing *kolam* drawings in front of
102 their doorstep or in their practice book and the age of initial learning.

103 **S2.1.2 GPS**

104 The geographical position of each individual's current residency was measured. A
105 GPS tracker of type Garmin GPSmap 60 CSx was used. The interviewer only asked
106 for the name of the native place and during data pre-processing the name of the native
107 place and if needed other demographic information was then used to manually map
108 the name of the native place to a GPS position. All GPS positions of the current
109 residency and the native place were recorded allowing distances between points to
110 be calculated by applying the distance formula to the x-y-z coordinates of the two
111 points.

112 **S2.1.3 *Kolam* Drawings**

113 A corpus of *kolam* drawings was compiled by soliciting individual's to draw *kolam*
114 drawings as part of the survey. Each individual was asked to draw a minimum of a
115 total of 20 *kolam* drawings.

116 S3 Statistical Analyses

117 S3.1 Random Intercept Models

118 S3.1.1 Statistical Model

$$\begin{aligned} \text{Density} &\sim \text{Log-Normal}(\mu_i, \sigma) \\ \mu_i &= \alpha_{\text{density}} + \alpha_j + \beta_{\text{caste}} + \beta_{\text{age}} \times \text{age} + \beta_{\text{practice duration}} \times \text{practice duration} \\ \beta_{\text{caste}} &= \sigma_{\text{caste}} \times z_{\text{caste}} \\ \alpha_j &= \sigma_{\text{artist}} \times z_{\text{artist}} \\ \alpha_{\text{density}} &\sim \text{Normal}(1, 2) \\ \sigma_{\text{caste}} &\sim \text{Normal}(0.5, 1) \\ \sigma_{\text{artist}} &\sim \text{Normal}(0, 0.5) \\ \sigma &\sim \text{Normal}(0.5, 0.5) \\ \beta_{\text{caste}} &\sim \text{Normal}(0, 1) \\ \beta_{\text{age}} &\sim \text{Normal}(0, 1) \\ z_{\text{caste}} &\sim \text{Normal}(0, 1) \\ z_{\text{artist}} &\sim \text{Normal}(0, 1) \end{aligned} \tag{6}$$

Evenness \sim Truncated Normal(μ_i, σ)[0, 1]

$$\mu_i = \alpha_{evenness} + \alpha_j + \beta_{caste} + \beta_{age} \times \text{age} + \beta_{\text{practice duration}} \times \text{practice duration}$$

$$\beta_{caste} = \sigma_{caste} \times z_{caste}$$

$$\alpha_j = \sigma_{artist} \times z_{artist}$$

$$\alpha_{evenness} \sim \text{Normal}(1, 2)$$

$$\sigma_{caste} \sim \text{Normal}(0.5, 1)$$

$$\sigma_{artist} \sim \text{Normal}(0, 0.5)$$

$$\sigma \sim \text{Normal}(0.5, 1)$$

$$\beta_{caste} \sim \text{Normal}(0, 1)$$

$$\beta_{age} \sim \text{Normal}(0, 1)$$

$$z_{caste} \sim \text{Normal}(0, 1)$$

$$z_{artist} \sim \text{Normal}(0, 1)$$

(7)

Richness \sim Poisson(λ_i)

$$\log(\lambda_i) = \alpha_{richness} + \alpha_j + \beta_{caste} + \beta_{age} \times \text{age} + \beta_{\text{practice duration}} \times \text{practice duration}$$

$$\beta_{caste} = \sigma_{caste} \times z_{caste}$$

$$\alpha_j = \sigma_{artist} \times z_{artist}$$

$$\alpha_{richness} \sim \text{Normal}(1, 2)$$

$$\sigma_{caste} \sim \text{Normal}(0.5, 1)$$

$$\sigma_{artist} \sim \text{Normal}(0, 0.5)$$

$$\beta_{caste} \sim \text{Normal}(0, 1)$$

$$\beta_{age} \sim \text{Normal}(0, 1)$$

$$z_{caste} \sim \text{Normal}(0, 1)$$

$$z_{artist} \sim \text{Normal}(0, 1)$$

(8)

$$\begin{aligned}
& \text{Canvas Size}_i \sim \text{NegBinom}(\mu_i, \phi) \\
\log(\mu_i) &= \alpha_{size} + \alpha_j + \beta_{caste} + \beta_{age} \times \text{age} + \beta_{\text{practice duration}} \times \text{practice duration} \\
& \beta_{caste} = \sigma_{caste} \times z_{caste} \\
& \alpha_j = \sigma_{artist} \times z_{artist} \\
& \alpha_{size} \sim \text{Normal}(1, 2) \\
& \sigma_{caste} \sim \text{Normal}(0.5, 1) \\
& \sigma_{artist} \sim \text{Normal}(0, 0.5) \\
& \phi \sim \text{Normal}(1.5, 3) \\
& \beta_{caste} \sim \text{Normal}(0, 1) \\
& \beta_{age} \sim \text{Normal}(0, 1) \\
& z_{caste} \sim \text{Normal}(0, 1) \\
& z_{artist} \sim \text{Normal}(0, 1)
\end{aligned} \tag{9}$$

$$\begin{aligned}
& \text{Entropy}_i \sim \text{Truncated Normal}(\mu_i, \sigma)[0, 1] \\
\mu_i &= \alpha_{entropy} + \alpha_j + \beta_{caste} + \beta_{age} \times \text{age} + \beta_{\text{practice duration}} \times \text{practice duration} \\
& \beta_{caste} = \sigma_{caste} \times z_{caste} \\
& \alpha_j = \sigma_{artist} \times z_{artist} \\
& \alpha_{entropy} \sim \text{Normal}(1, 2) \\
& \sigma_{caste} \sim \text{Normal}(0.5, 1) \\
& \sigma_{artist} \sim \text{Normal}(0, 0.5) \\
& \phi \sim \text{Normal}(1.5, 3) \\
& \beta_{caste} \sim \text{Normal}(0, 1) \\
& \beta_{age} \sim \text{Normal}(0, 1) \\
& z_{caste} \sim \text{Normal}(0, 1) \\
& z_{artist} \sim \text{Normal}(0, 1)
\end{aligned} \tag{10}$$

119 S3.1.2 Visual MCMC Diagnostics

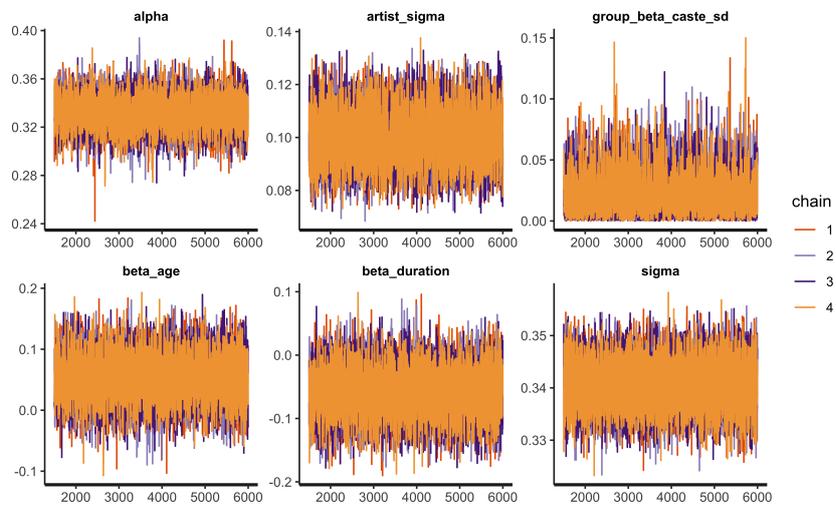


Figure S3: Traceplot for the random intercept model on density showing mixing across chains and convergence.

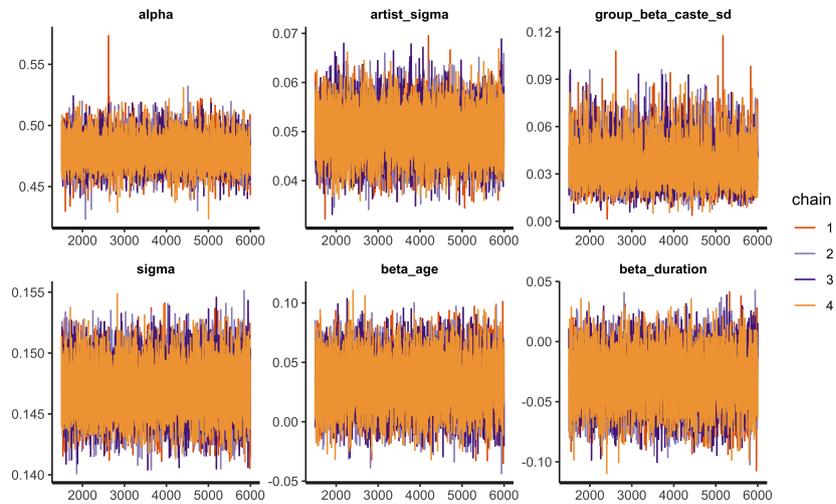


Figure S4: Traceplot for the random intercept model on evenness showing mixing across chains and convergence.

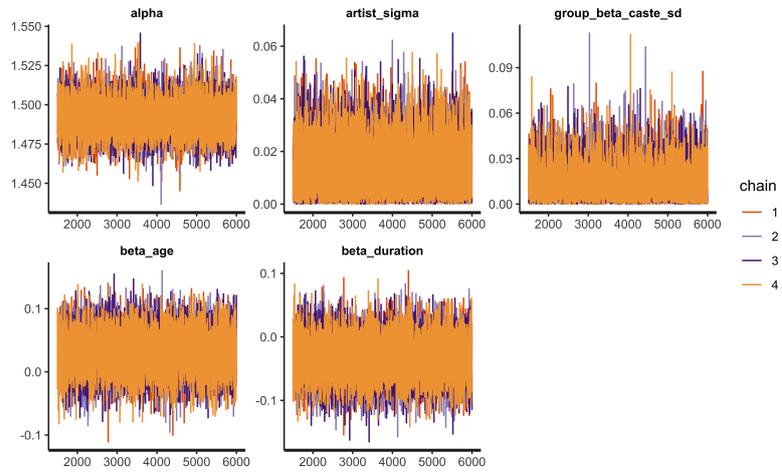


Figure S5: Traceplot for the random intercept model on richness showing mixing across chains and convergence.

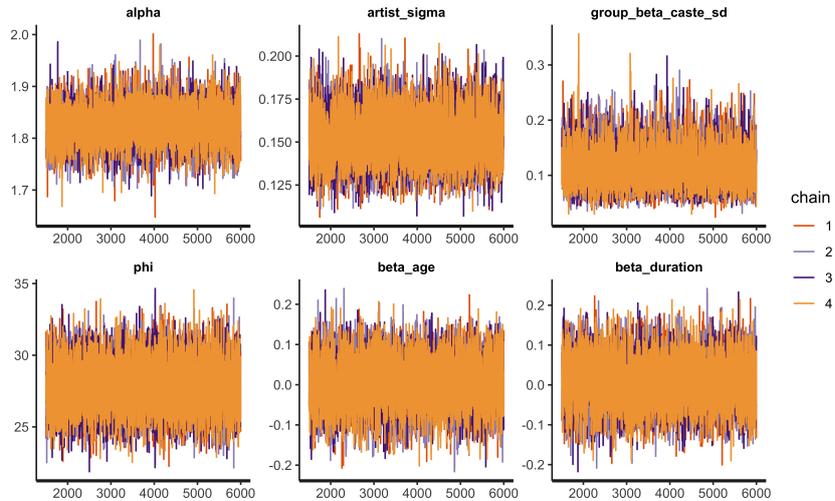


Figure S6: Traceplot for the random intercept model on canvas size showing mixing across chains and convergence.

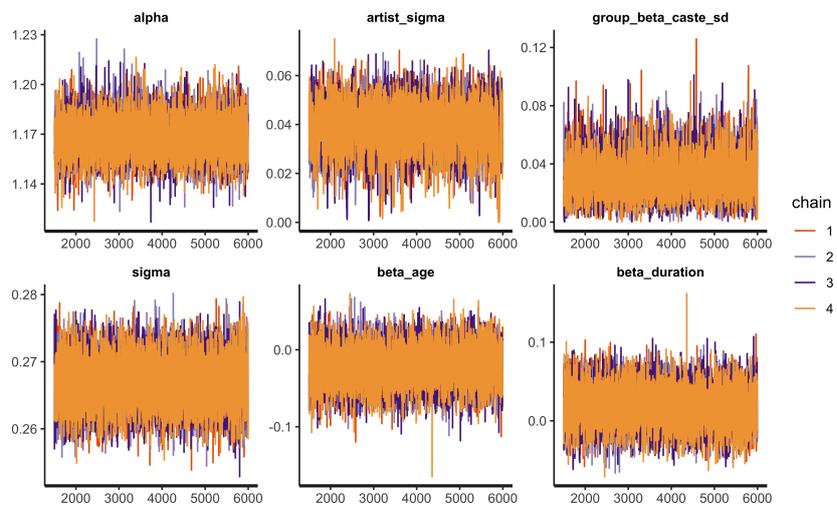


Figure S7: Traceplot for the random intercept model on entropy showing mixing across chains and convergence.

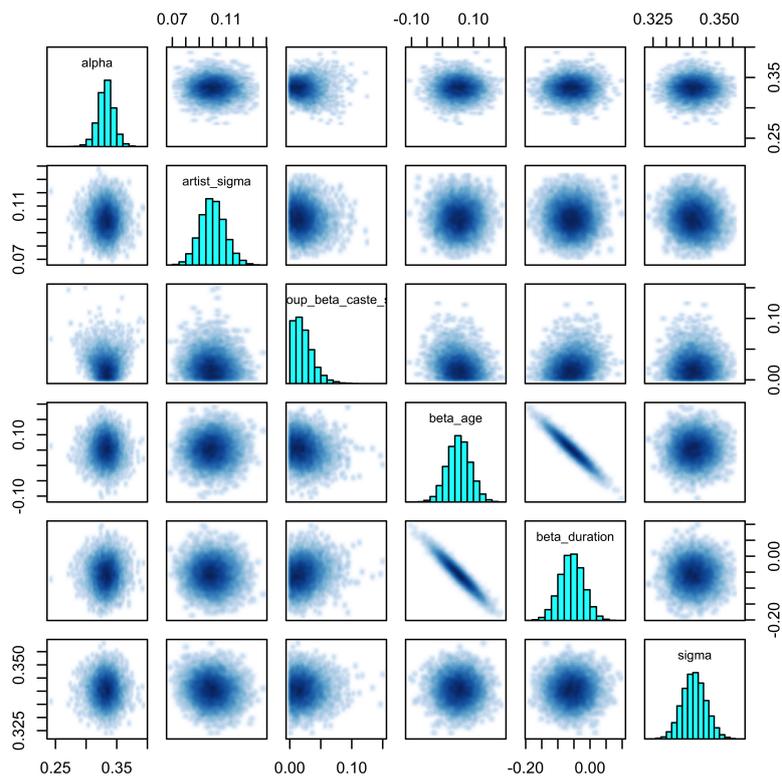


Figure S8: Pairs plot for the random intercept model on density showing correlation among parameters and no sampling problems.

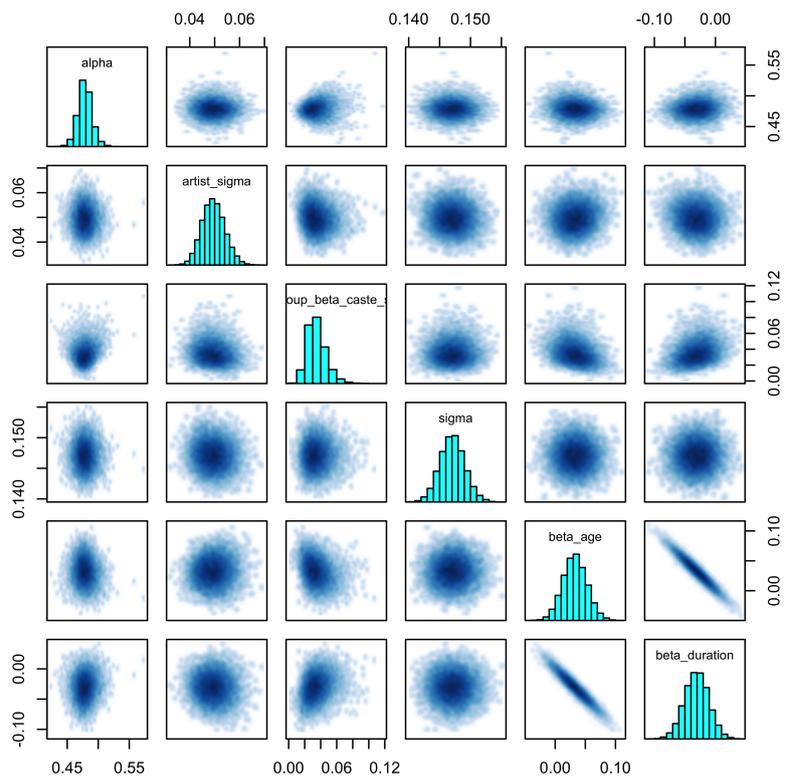


Figure S9: Pairs plot for the random intercept model on gini showing correlation among parameters and no sampling problems.

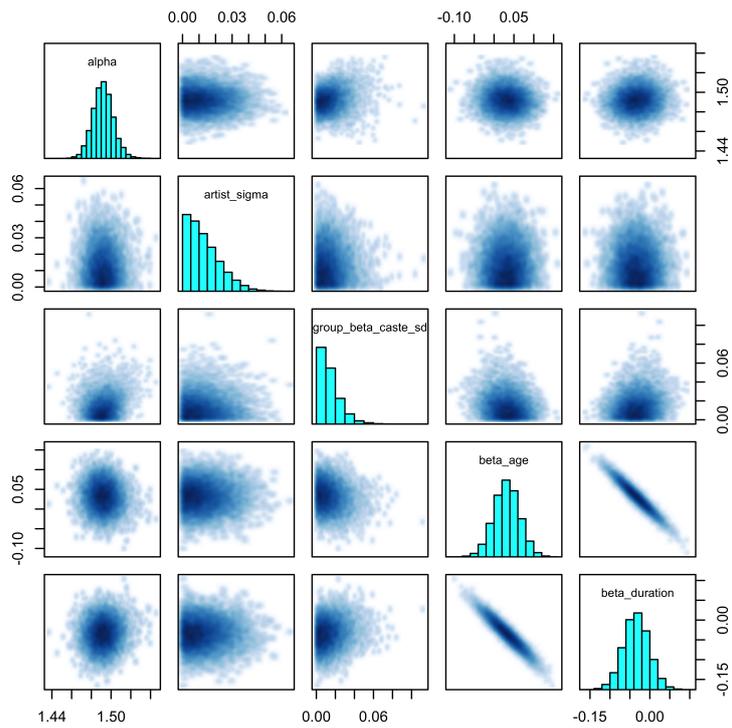


Figure S10: Pairs plot for the random intercept model on richness showing correlation among parameters and no sampling problems.

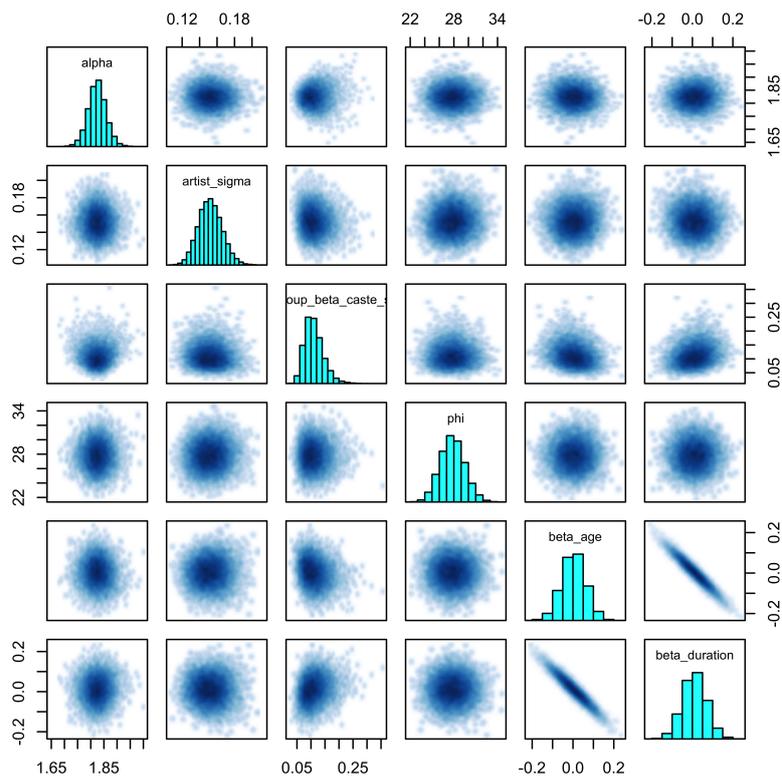


Figure S11: Pairs plot for the random intercept model on canvas size showing correlation among parameters and no sampling problems.

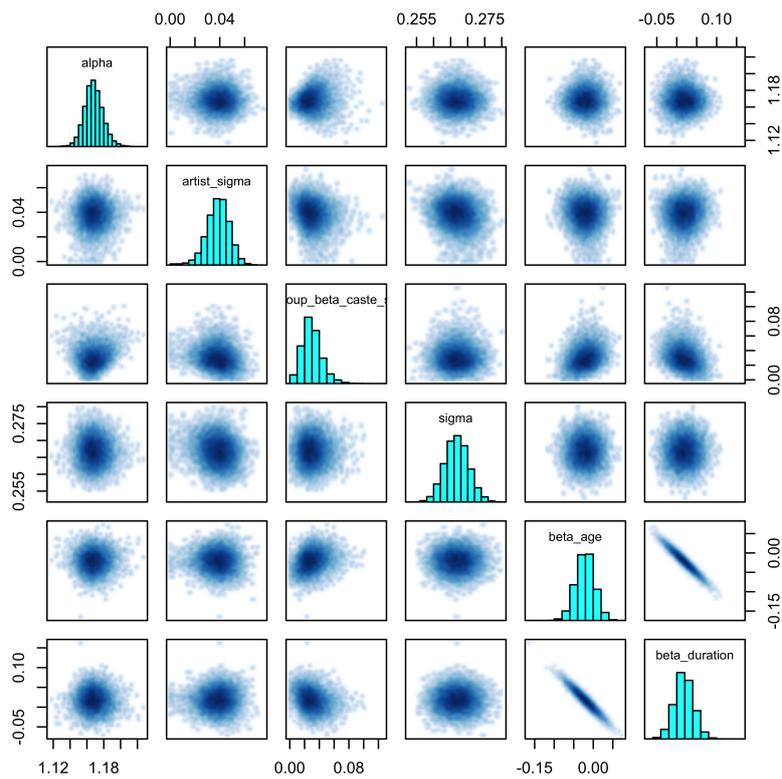


Figure S12: Pairs plot for the random intercept model on entropy showing correlation among parameters and no sampling problems.

120 **S3.1.3 Intraclass Correlation (ICC)**

121 The intraclass-correlation coefficient (ICC) can be calculated for Gaussian models
122 to determine the variance explained by random and fixed effects ([Gelman and Hill,](#)
123 [2006](#), p.258). Since our five models are non-Gaussian, we approximated the ICC
124 using variance decomposition based on the posterior predictive distribution. We first
125 drew from the posterior predictive distribution not conditioned on our fixed (i.e.,
126 age and duration of practice) and random effect (i.e., caste and individual variation)
127 terms and then drew from the posterior predictive distribution conditioned on all
128 fixed and random effects. Subsequently, we calculated the variances for both draws.
129 The pseudo-ICC is then the ratio between these two variances. Occasionally, the
130 variance ratio can be negative due to very large variance of the posterior predictive
131 distributions.

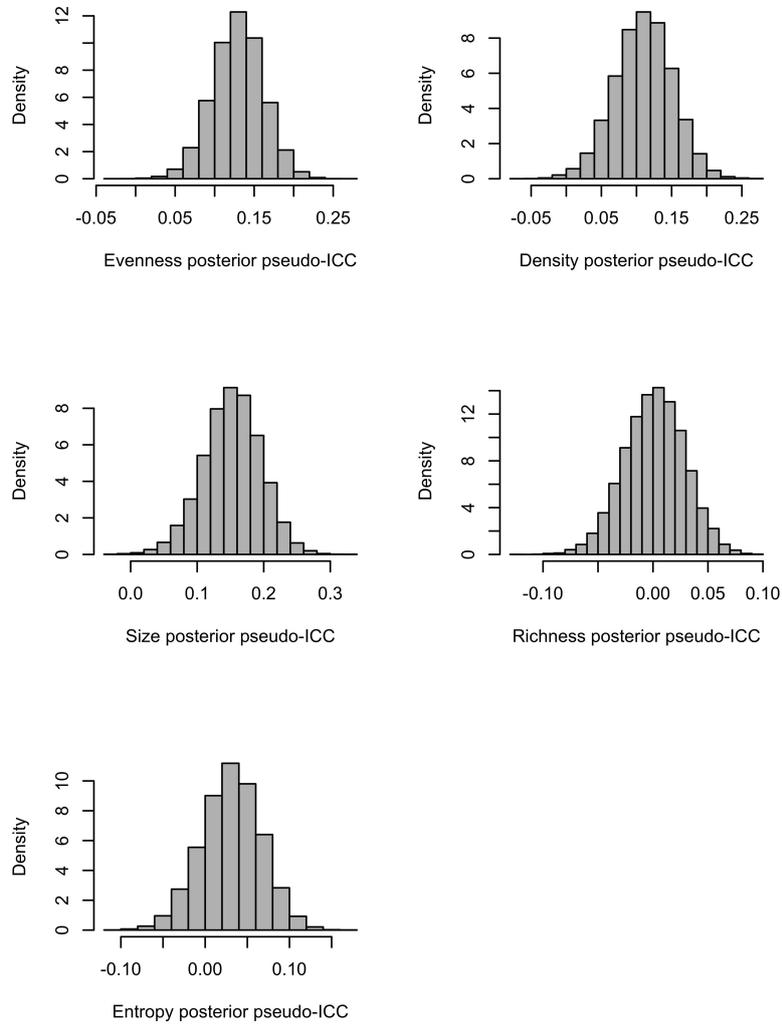


Figure S13: Intraclass Correlation Coefficients (ICC) for individual random-effect variances for the four outcome variables. The variance decomposition is based on the posterior predictive distribution, which is the correct way for Bayesian non-Gaussian models.

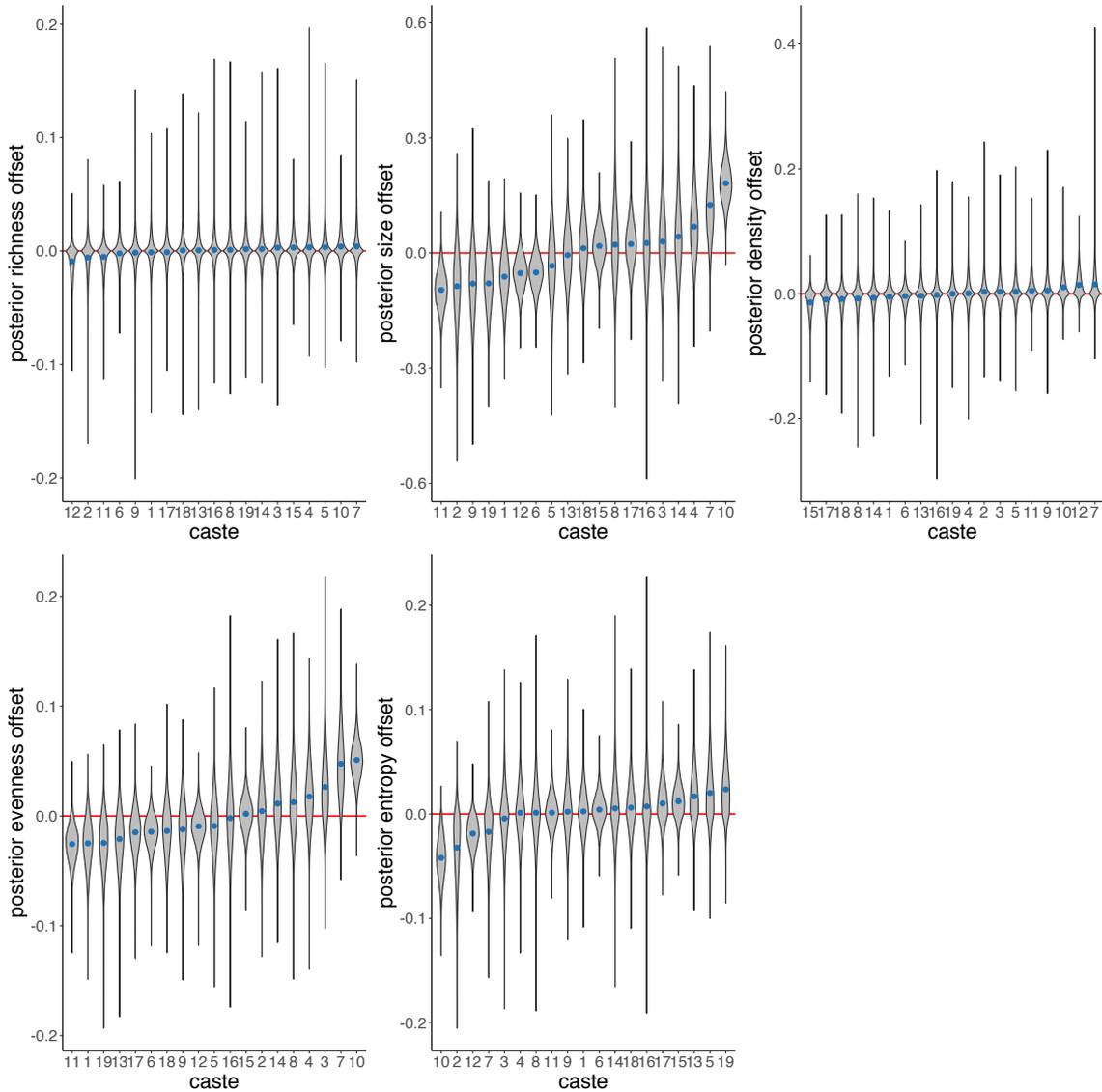


Figure S14: Caste random effect offsets for each outcome variable. The red line reflects zero offset. Each violin probability density plot displays the variation within each caste on the corresponding outcome variable (i.e., richness, canvas size, density, evenness and entropy). The posterior mean offset is illustrated in blue. The range of the violin plots reflect the 90% HPDI.

133 **S3.2 Gaussian Process Models**

134 To investigate whether the observed variation in structural and information-theoretic
135 properties of *kolam* drawings is structured by the residential or native place proximity
136 between individuals, we used a Gaussian process (GP) model to estimate a function
137 for the covariance between pairs of individuals at different spatial distances of their
138 residency as well as their native place. We fit the GP model on the five outcome
139 variables entropy, canvas size, gesture density, richness and evenness. To foreshadow
140 the conclusions drawn from the GP model, results from the GP model are in line with
141 our main results from the random-intercept models. The GP model further indicates
142 no major differentiation between artists originally from the community and those who
143 emigrated from other parts of India, and no spatial association between artist’s *kolam*
144 drawings.

145 **S3.2.1 Statistical Model**

146 We estimated unique intercepts for each individual with a varying effects approach
147 to continuous categories using a Gaussian process (GP) model. A GP can be seen as
148 a distribution of nonlinear functions. Placing a GP prior over the covariance, allows
149 us to estimate a function for the covariance between pairs of individuals at different
150 variable distances (McElreath, 2016). The individual-level covariance matrix α_i was
151 modeled using an exponentiated quadratic kernel. This covariance function implies
152 that the covariance between any two individuals j and k declines exponentially with
153 the squared distance between them. No variation between individuals would corre-
154 spond to a covariance of zero or close to zero. The parameter ρ^2 , also referred to
155 as length scale, determines the rate of decline. If ρ^2 is large, then the covariance
156 decreases slowly with squared distance, while if ρ^2 is small it decreases rapidly with
157 squared distance. This length-scale ρ^2 prior is constrained to be between zero and
158 one because the distances were normalized to be between zero and one. η^2 is the max-
159 imum covariance between any two individuals. Automatic relevance determination
160 (Neal, 1996) was performed where multiple predictors with corresponding length-scale
161 parameters for each dimension were merged into the GP.

162 Statistical models on the five structural and information-theoretic properties of
163 *kolam* drawings were implemented. The predictor variables to investigate the parti-
164 tion of variation across individuals and the population was the same in all models.
165 Age, duration of practice, caste, and residential and native place proximity between
166 individuals were used as predictor variables to explain individual variation on the
167 structural and information-theoretic properties. The residential and native place dis-
168 tances as well as age and practice duration were normalized, such that the minimum
169 value was zero and the maximum value was one. The pairwise Euclidean distances
170 between each pair of individuals along each predictor variable dimension was then
171 computed for age, duration of practice and residential and native place proximity
172 between individuals. GPS coordinates were used to compute residential and na-
173 tive place distance matrices between individuals. Distances correspond to Euclidean
174 distance between individuals for the specific variable dimension. For example, the
175 residence or native place distance matrix are spatial distances. Caste was modeled as
176 a hierarchical non-centred categorical variable with 19 categories.

$$\begin{aligned}
K_{jk} &= \eta^2 \exp \left[- \left(\frac{\text{Residence}^2}{2 \times \rho_R^2} + \frac{\text{Native Place}^2}{2 \times \rho_N^2} + \frac{\text{Age}^2}{2 \times \rho_A^2} + \frac{\text{Duration}^2}{2 \times \rho_D^2} \right) \right] + \delta_{jk} \times 0.001 \\
\text{Density} &\sim \text{Log-Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha_{density} + \alpha_j + \beta_{caste} \\
\alpha_j &\sim \text{MVNormal}(0, K(x)) \\
\beta_{caste} &= \sigma_{caste} \times z \\
\eta^2 &\sim \text{Normal}(5, 2) \\
\rho_R^2 &\sim \text{Beta}(1, 2) \\
\rho_N^2 &\sim \text{Beta}(1, 2) \\
\rho_A^2 &\sim \text{Beta}(1, 2) \\
\rho_D^2 &\sim \text{Beta}(1, 2) \\
\sigma &\sim \text{Normal}(0.5, 0.5) \\
\alpha_{density} &\sim \text{Normal}(0.5, 1) \\
\sigma_{caste} &\sim \text{Normal}(0.5, 0.5)
\end{aligned}$$

Evenness \sim Truncated Normal(μ_i, σ)[0, 1]

$$\mu_i = \alpha_{evenness} + \alpha_j + \beta_{caste}$$

$$\alpha_j \sim \text{MVNormal}(0, K(x))$$

$$K_{jk} = \eta^2 \exp \left[- \left(\frac{\text{Residence}^2}{2 \times \rho_R^2} + \frac{\text{Native Place}^2}{2 \times \rho_N^2} + \frac{\text{Age}^2}{2 \times \rho_A^2} + \frac{\text{Duration}^2}{2 \times \rho_D^2} \right) \right] + \delta_{jk} \times 0.001$$

$$\beta_{caste} = \sigma_{caste} \times z$$

$$\eta^2 \sim \text{Normal}(5, 2)$$

$$\rho_R^2 \sim \text{Beta}(1, 2)$$

$$\rho_N^2 \sim \text{Beta}(1, 2)$$

$$\rho_A^2 \sim \text{Beta}(1, 2)$$

$$\rho_D^2 \sim \text{Beta}(1, 2)$$

$$\sigma \sim \text{Normal}(0.5, 0.5)$$

$$\alpha_{evenness} \sim \text{Normal}(0.5, 1)$$

$$\sigma_{caste} \sim \text{Normal}(0.5, 0.5)$$

$$\begin{aligned}
& \text{Richness} \sim \text{Poisson}(\lambda_i) \\
& \log(\lambda_i) = \alpha_{richness} + \alpha_j + \beta_{caste} \\
& \alpha_j \sim \text{MVNormal}(0, K(x)) \\
K_{jk} = \eta^2 \exp & \left[- \left(\frac{\text{Residence}^2}{2 \times \rho_R^2} + \frac{\text{Native Place}^2}{2 \times \rho_N^2} + \frac{\text{Age}^2}{2 \times \rho_A^2} + \frac{\text{Duration}^2}{2 \times \rho_D^2} \right) \right] + \delta_{jk} \times 0.001 \\
& \beta_{caste} = \sigma_{caste} \times z \\
& \eta^2 \sim \text{Normal}(5, 2) \\
& \rho_R^2 \sim \text{Beta}(1, 2) \\
& \rho_N^2 \sim \text{Beta}(1, 2) \\
& \rho_A^2 \sim \text{Beta}(1, 2) \\
& \rho_D^2 \sim \text{Beta}(1, 2) \\
& \sigma \sim \text{Normal}(0.5, 0.5) \\
& \alpha_{richness} \sim \text{Normal}(0.5, 1) \\
& \sigma_{caste} \sim \text{Normal}(0.5, 0.5)
\end{aligned}$$

$$\begin{aligned}
& \text{Canvas Size}_i \sim \text{NegBinom}(\mu_i, \phi) \\
& \log(\mu_i) = \alpha_{size} + \alpha_j + \beta_{caste} \\
& \alpha_j \sim \text{MVNormal}(0, K(x)) \\
K_{jk} = \eta^2 \exp & \left[- \left(\frac{\text{Residence}^2}{2 \times \rho_R^2} + \frac{\text{Native Place}^2}{2 \times \rho_N^2} + \frac{\text{Age}^2}{2 \times \rho_A^2} + \frac{\text{Duration}^2}{2 \times \rho_D^2} \right) \right] + \delta_{jk} \times 0.001 \\
& \beta_{caste} = \sigma_{caste} \times z \\
& \eta^2 \sim \text{Normal}(5, 2) \\
& \rho_R^2 \sim \text{Beta}(1, 2) \\
& \rho_N^2 \sim \text{Beta}(1, 2) \\
& \rho_A^2 \sim \text{Beta}(1, 2) \\
& \rho_D^2 \sim \text{Beta}(1, 2) \\
& \sigma \sim \text{Normal}(0.5, 0.5) \\
& \alpha_{size} \sim \text{Normal}(0.5, 1) \\
& \sigma_{caste} \sim \text{Normal}(0.5, 0.5)
\end{aligned}$$

$$\begin{aligned}
& \text{Entropy} \sim \text{Truncated Normal}(\mu_i, \sigma)[0,] \\
& \mu_i = \alpha_{entropy} + \alpha_j + \beta_{caste} \\
& \alpha_j \sim \text{MVNormal}(0, K(x)) \\
K_{jk} = \eta^2 \exp & \left[- \left(\frac{\text{Residence}^2}{2 \times \rho_R^2} + \frac{\text{Native Place}^2}{2 \times \rho_N^2} + \frac{\text{Age}^2}{2 \times \rho_A^2} + \frac{\text{Duration}^2}{2 \times \rho_D^2} \right) \right] + \delta_{jk} \times 0.001 \\
& \beta_{caste} = \sigma_{caste} \times z \\
& \eta^2 \sim \text{Normal}(5, 2) \\
& \rho_R^2 \sim \text{Beta}(1, 2) \\
& \rho_N^2 \sim \text{Beta}(1, 2) \\
& \rho_A^2 \sim \text{Beta}(1, 2) \\
& \rho_D^2 \sim \text{Beta}(1, 2) \\
& \sigma \sim \text{Normal}(0.5, 0.5) \\
& \alpha_{entropy} \sim \text{Normal}(0.5, 1) \\
& \sigma_{caste} \sim \text{Normal}(0.5, 0.5)
\end{aligned}$$

177 S3.2.2 Estimation of Variation

178 The five statistical models were implemented in the probabilistic programming lan-
179 guage Stan 2.18 (Carpenter et al., 2017), using 6000 samples from four chains. Anal-
180 yses were performed in R (R Core Team, 2019). Data and analyses can be found here:
181 http://github.com/nhtran93/kolam_signaling. All R-hat values were less than
182 1.01, and visual inspection of trace plots and rank histograms indicated convergence
183 of all models.

184 Between-individual variation in entropy ($\eta^2 = 0.00$, 90% CI [0.00, 0.00]), density
185 ($\eta^2 = 0.01$, 90% CI [0.01, 0.01]), the evenness ($\eta^2 = 0$, 90% CI [0, 0]), and in the
186 richness ($\eta^2 = 0.00$, 90% CI [0, 0.01]) were estimated with high certainty to be very
187 small and close to zero (see left panel in Figure S15). The between-artist variability
188 is most pronounced in canvas size ($\eta^2 = 0.03$, 90% CI [0.02, 0.03]), while *kolam*
189 drawings show only small distinct variation between artists in the other structural and

190 information-theoretic properties. Our results further show no evidence that variation
191 in structural and information-theoretic properties of *kolam* drawings covary with the
192 spatial structure, age or practice duration (see Figure S15).

193 Prior-posterior plots of all five models show that the priors updated for all param-
194 eters except the length scale parameters ρ because there is barely any information to
195 explain individual variation with no individual variation present. We detected very
196 small effects of caste membership on the entropy, density, evenness, and richness, with
197 varying-effect deviations estimated near zero with high certainty (entropy: $\sigma_{caste} =$
198 0.03 , 90% CI $[0.01, 0.05]$; density: $\sigma_{caste} = 0.04$, 90% CI $[0.01, 0.07]$; evenness: σ_{caste}
199 $= 0.03$, 90% CI $[0.01, 0.04]$; and richness: $\sigma_{caste} = 0.01$, 90% CI $[0.00, 0.03]$ respec-
200 tively). We detected more pronounced effects of caste membership on canvas size
201 (canvas size: $\sigma_{caste} = 0.09$, 90% CI $[0.05, 0.14]$) as illustrated in Figure S15.

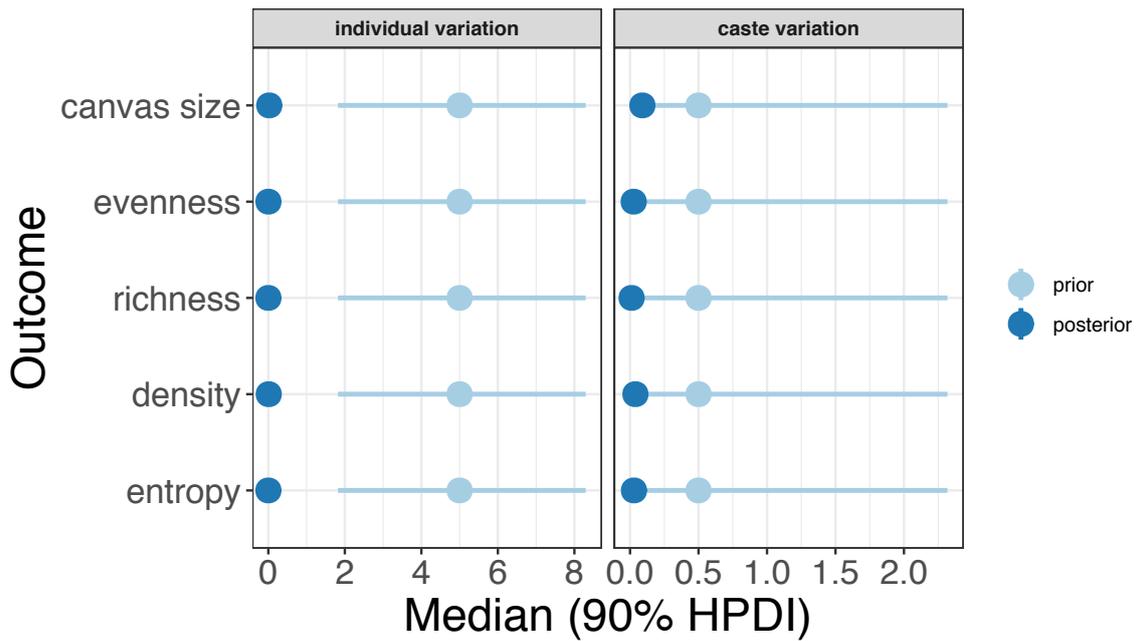


Figure S15: Prior-Posterior Coefficient Plots of Individual variation and Caste Variation. All panels have the same y-axis indicating the five models. The left panel (eta squared) illustrates the estimated individual variation (dark blue) in comparison to the prior (light blue) for each model. The right panel illustrates the estimated population-level standard deviation for the effect of caste (dark blue) in comparison to the prior (light blue) for each model. The 90% Highest Posterior Density Interval (HPDI) was computed for each posterior; however, the interval is very narrow.

202 References

- 203 Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M.,
 204 Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic pro-
 205 gramming language. *Journal of Statistical Software, Articles*, 76(1):1–32.
- 206 Frank, S. A. (2009). The common patterns of nature. *Journal of evolutionary biology*,
 207 22(8):1563–1585.
- 208 Gelman, A. and Hill, J. (2006). *Data Analysis Using Regression and Multi-*

- 209 *level/Hierarchical Models*. Analytical Methods for Social Research. Cambridge
210 University Press.
- 211 McElreath (2016). *Statistical Rethinking: A Bayesian Course with Examples in R*
212 *and Stan*. CRC Press.
- 213 Neal, R. M. (1996). Bayesian Learning for Neural Networks. In *Lecture Notes in*
214 *Statistics 118*. Springer, New York.
- 215 R Core Team (2019). R: A language and environment for statistical computing.
- 216 Ravallion, M. (2014). Income inequality in the developing world. *Science*,
217 344(6186):851–855.
- 218 Waring, T. M. (2012). Sequential Encoding of Tamil Kolam Patterns. *Forma*, 27:83–
219 92.
- 220 Zoli, C. (1999). Intersecting generalized Lorenz curves and the Gini index. *Social*
221 *Choice and Welfare*, 16(2):183–196.