Supplementary Materials for "The (anti-)conformity strategy space: how flexible is social learning when group membership and incentive alignment vary?"

Lukas von Flüe[1,*], Sonja Vogt[1,*], and Charles Efferson[1,*]

[1]Faculty of Business and Economics, University of Lausanne

[*]Address correspondence to *lukas.vonfluee@unil.ch*, *sonja.vogt@unil.ch*, and *charles.efferson@unil.ch*.

This document can be found on `https://doi.org/10.17605/OSF.IO/MZFQ5`.

S1 Statistical Models

Pure behavior: The analysis of pure behavior allows us to investigate whether social learners are able to adjust their learning strategies symmetrically between the four informationally equivalent treatments. The following OLS model shows a regression of y_{ij}, a binary variable representing the choice between the two urns, on 7 dummies representing treatment combinations, all relative to an omitted category, plus interactions between these dummies and the social information x_i, i.e. the proportion of demonstrators choosing the *right* urn in the fourth round centered at zero. Centering at zero means that if p_{ij} is the proportion of demonstrators choosing the right urn in the fourth round of a given block, for observation i for social learner j, we transform this as $x_{ij} = p_{ij} - 0.5$. We have that $x_{ij} \in [-0.5, 0.5]$, where $x_{ij} = -0.5$ means that all demonstrators have chosen the *left* urn, and $x_{ij} = 0.5$ means that all demonstrators have chosen the *right* urn. We define the urn choice as the dependent variable and we have the following notation: $y_{ij} \in \{0, 1\}$, with $y_{ij} = 1 \equiv$ "right urn". We use the following notation for the dummy variable to indicate whether a social learner observes ingroup or outgroup demonstrators: $d_{ij,in} \in \{0, 1\}$, with $d_{ij,in} = 1 \equiv$ "ingroup". We use μ to label similarity to ingroup and we use the following notation for the dummy variable: $d_{ij,09} \in \{0, 1\}$, with $d_{ij,09} = 1 \equiv \mu = 0.9$. Finally, we use the following notation to indicate whether a social learner is in the cognitive load treatment or not: $d_{j,c} \in \{0, 1\}$, with $d_{j,c} = 1 \equiv$ "cognitive load". Let [*] denote treatment combinations, where [conditions true] = 1, and [¬(conditions true)] = 0. For example, $[d_{ij,in} = 0, d_{ij,09} = 1, d_{j,c} = 1]$ would stand for the treatment in which social learners observe *outgroup* demonstrators, the similarity to the ingroup is $\mu = 0.9$, and the "cognitive load" treatment is implemented. We have the following model,

$$
\begin{aligned}
y_{ij} = \{controls\} &+ \beta_0 + \beta_1 x_{ij} + \beta_2 \left[d_{ij,in} = 0, d_{ij,09} = 0, d_{j,c} = 1\right] \\
&+ \beta_3 \left[d_{ij,in} = 0, d_{ij,09} = 1, d_{j,c} = 0\right] + \beta_4 \left[d_{ij,in} = 0, d_{ij,09} = 1, d_{j,c} = 1\right] \\
&+ \beta_5 \left[d_{ij,in} = 1, d_{ij,09} = 0, d_{j,c} = 0\right] + \beta_6 \left[d_{ij,in} = 1, d_{ij,09} = 0, d_{j,c} = 1\right] \\
&+ \beta_7 \left[d_{ij,in} = 1, d_{ij,09} = 1, d_{j,c} = 0\right] + \beta_8 \left[d_{ij,in} = 1, d_{ij,09} = 1, d_{j,c} = 1\right] \\
&+ \beta_9 \left[d_{ij,in} = 0, d_{ij,09} = 0, d_{j,c} = 1\right] x_{ij} + \beta_{10} \left[d_{ij,in} = 0, d_{ij,09} = 1, d_{j,c} = 0\right] x_{ij} \\
&+ \beta_{11} \left[d_{ij,in} = 0, d_{ij,09} = 1, d_{j,c} = 1\right] x_{ij} + \beta_{12} \left[d_{ij,in} = 1, d_{ij,09} = 0, d_{j,c} = 0\right] x_{ij} \\
&+ \beta_{13} \left[d_{ij,in} = 1, d_{ij,09} = 0, d_{j,c} = 1\right] x_{ij} + \beta_{14} \left[d_{ij,in} = 1, d_{ij,09} = 1, d_{j,c} = 0\right] x_{ij} \\
&+ \beta_{15} \left[d_{ij,in} = 1, d_{ij,09} = 1, d_{j,c} = 1\right] x_{ij} + \epsilon_i.
\end{aligned} \tag{S1}
$$

Optimal behavior: We use a logistic regression with robust clustered standard errors to model optimal behavior. The choice of the optimal urn is represented by z_{ij}. In this first version of the regression we don't control for w_{ij}, the centered proportion of demonstrators choosing optimally, and we don't have interaction terms with w_{ij}, because we focus on how the different treatments affect social learners' abilities to symmetrically adjust between informationally equivalent treatments.

$$
\begin{aligned}
z_{ij} = \{controls\} &+ \beta_0 + \beta_1 \left[d_{ij,in} = 0, d_{ij,09} = 0, d_{j,c} = 1\right] \\
&+ \beta_2 \left[d_{ij,in} = 0, d_{ij,09} = 1, d_{j,c} = 0\right] + \beta_3 \left[d_{ij,in} = 0, d_{ij,09} = 1, d_{j,c} = 1\right] \\
&+ \beta_4 \left[d_{ij,in} = 1, d_{ij,09} = 0, d_{j,c} = 0\right] + \beta_5 \left[d_{ij,in} = 1, d_{ij,09} = 0, d_{j,c} = 1\right] \\
&+ \beta_6 \left[d_{ij,in} = 1, d_{ij,09} = 1, d_{j,c} = 0\right] + \beta_7 \left[d_{ij,in} = 1, d_{ij,09} = 1, d_{j,c} = 1\right] + \epsilon_i
\end{aligned} \tag{S2}
$$

S2 Additional analysis

S2.1 Demonstrator behavior

Demonstrators identified their own optimal urns in the final periods of the blocks at a rate that is highly significantly above chance. Results differ slightly when comparing the data from Zurich and Nairobi. In Zurich, in the fourth round of a block, at least three of the five demonstrators chose the optimal urn in 92.8% of group-by-block observations. This is close to the 95.5% that Efferson et al. (2016) find. In Nairobi, it was 70.2%. The majority-optimal statistic we measure here pools 3–2, 4–1, and 5–0 splits and collapses across blocks in which the left versus right urn was optimal. Below we show the results of a regression of demonstrator's choice optimum on the last round of a given block, and on a dummy indicating whether the right urn was optimal. We include this dummy variable to test whether demonstrators exhibit a bias for the right or left urn, which does not seem to be the case.

Table S1: Logistic regression showing whether demonstrators chose optimally. "Last round of block" is a dummy variable indicating whether the choice of the demonstrator was made in the fourth round of a given block. "Right urn optimal" is a dummy variable indicating whether the right urn was optimal in a given round when the focal demonstrator chose between the two urns. The third column shows 95% confidence intervals. Those are based on robust standard errors clustered on the demonstrator shown in parentheses in the second column. Asterisks denote the level of significance of the p-values: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

	Zurich		**Nairobi**	
Parameter	Estimate	95 % CI	Estimate	95 % CI
Intercept	0.697 ***	[0.538, 0.857]	0.192 **	[0.064, 0.321]
	(0.081)		(0.066)	
Last round of block	0.679 ***	[0.505, 0.853]	0.153 **	[0.038, 0.268]
	(0.089)		(0.059)	
Right urn optimal	-0.173	[-0.440, 0.094]	0.148	[-0.055, 0.350]
	(0.136)		(0.103)	

S2.2 Social learner behavior

S2.2.1 Manipulation check

In Table S2 we show the results for a treatment manipulation check. With this manipulation check we wanted to test whether the "cognitive load" treatment was effective in reducing the social learners' ability to recall the multi-digit numbers correctly when compared to the social learners in the "no cognitive load" treatment. The dummy representing the cognitive load treatment is highly significant in reducing the likelihood of a social learner to recall a number correctly in both the Zurich data and the Nairobi data.

Table S2: Likelihood of recalling number correctly (logistic regression). Asterisks denote the level of significance of the p-values: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Robust standard errors in parentheses are clustered on the social learner.

	Zurich		**Nairobi**	
Parameter	Estimate	95 % CI	Estimate	95 % CI
Intercept	4.97 ***	[4.48, 5.47]	4.20 ***	[3.79, 4.62]
	(0.251)		(0.214)	
Cognitive load dummy	-3.39 ***	[-3.95, -2.83]	-2.90 ***	[-3.37, -2.44]
	(0.285)		(0.236)	

S2.2.2 Exploratory heterogeneity by engagement with the memorisation task

A remaining concern is that the null average effect of cognitive load on social learning behaviour could mask heterogeneity within the cognitive-load treatment if some social learners disengaged from the memorisation task, especially when the optimal response required the less direct mapping from outgroup demonstrators to one's own incentives. To assess this possibility, we conducted exploratory analyses restricted to social learners in the cognitive-load treatment. We used three participant-level measures of engagement with the memorisation task: (i) objective recall success across the 20 blocks, (ii) the self-reported number of times the number was memorized correctly (Question 5 in Section S4.1), and (iii) the

self-reported allocation of cognitive resources between the memorisation task and the urn choice task (Question 6 in Section S4.1, reverse-coded so that higher values indicate greater focus on memorisation). Because these engagement measures are themselves post-treatment outcomes, we treat the following analyses as exploratory.

The descriptive results provide little indication that social learners simply gave up on the memorisation task. Mean objective recall success in the cognitive-load treatment was 16.59 correct recalls out of 20 in Zurich and 15.72 out of 20 in Nairobi, and in neither site did any social learner have zero correct recalls across all 20 blocks. Table S3 reports site-specific logistic regressions predicting optimal choice within the cognitive-load treatment. In each model, we interacted the outgroup indicator with one of the three engagement measures while controlling for the centred proportion of demonstrators choosing their optimal urn, block number, similarity to ingroup, the outgroup indicator, and the similarity $\times$ outgroup interaction.

Across both sites, none of the interaction terms was statistically distinguishable from zero. In Zurich, the estimated interaction coefficients were $b = 0.082$ ($SE = 0.148$, $p = 0.580$) for objective recall success, $b = 0.061$ ($SE = 0.137$, $p = 0.655$) for self-reported recall success, and $b = 0.088$ ($SE = 0.126$, $p = 0.485$) for self-reported focus on memorisation. In Nairobi, the corresponding estimates were $b = 0.102$ ($SE = 0.107$, $p = 0.341$), $b = 0.012$ ($SE = 0.095$, $p = 0.898$), and $b = 0.023$ ($SE = 0.073$, $p = 0.757$). These exploratory results therefore do not support the interpretation that the null average effect of cognitive load is driven by a subgroup of social learners who disengaged from the memorisation task and consequently showed a disproportionately strong difficulty with outgroup demonstrators.

Table S3: Exploratory heterogeneity within the cognitive-load treatment

Site	Moderator	Estimate	SE	p-value
Zurich	Objective recall success	0.082	0.148	0.580
Zurich	Self-reported recall (Question 5)	0.061	0.137	0.655
Zurich	Self-reported memory focus (Question 6)	0.088	0.126	0.485
Nairobi	Objective recall success	0.102	0.107	0.341
Nairobi	Self-reported recall (Question 5)	0.012	0.095	0.898
Nairobi	Self-reported memory focus (Question 6)	0.023	0.073	0.757

Note. Each row reports the coefficient on the interaction between observing outgroup demonstrators and a participant-level measure of engagement with the memorisation task, estimated in a site-specific logistic regression restricted to the cognitive-load treatment. The cognitive-load samples comprise 91 social learners in Zurich and 81 social learners in Nairobi. All models control for the centered proportion of demonstrators choosing their optimal urn, block number, similarity to ingroup, observing the outgroup, and the similarity $\times$ outgroup interaction. Standard errors are clustered on social learners. The three engagement measures were standardised within site before entering the regressions. Objective recall success is the total number of correctly recalled numbers across the 20 blocks. Question 5 is the self-reported number of times the number was recalled correctly. Question 6 is reverse-coded so that higher values indicate greater focus on memorisation rather than on the urn choice task. Mean objective recall totals were 16.59 out of 20 in Zurich and 15.72 out of 20 in Nairobi; in neither site did any social learner in the cognitive-load treatment have zero correct recalls across all 20 blocks.

S2.2.3 Optimal behavior

Table S4 shows the results associated with the model shown in equation S2 in section S1. Table S5 extends this model by adding the centered proportion of demonstrators choosing their optimal urn (abbreviated as "centered prop. dem. opt." in the table) as a control. Table S6 extends this by adding interaction terms between the centered proportion of demonstrators choosing optimally and the treatment dummies.

Table S4: Likelihood of a social learner choosing optimally (log. regression). The omitted category is dissimilar to ingroup, observing the outgroup, and no cognitive load. The third and fifth columns show 95% confidence intervals. Those are based on robust standard errors clustered on the social learner (shown in parentheses in the second and fourth columns). Asterisks denote the level of significance of the p-values: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

	Zurich		**Nairobi**	
Parameter	Estimate	95 % CI	Estimate	95 % CI
Intercept	0.729 ***	[0.461, 0.996]	-0.042	[-0.241, 0.156]
	(0.136)		(0.101)	
Similar to ingroup, observe ingroup, no cognitive load	0.597 ***	[0.251, 0.944]	0.351 *	[0.080, 0.621]
	(0.177)		(0.138)	
Dissimilar to ingroup, observe ingroup, no cognitive load	0.407 **	[0.121, 0.693]	0.028	[-0.233, 0.289]
	(0.146)		(0.133)	
Similar to ingroup, observe outgroup, no cognitive load	0.386 *	[0.087, 0.684]	0.141	[-0.144, 0.426]
	(0.152)		(0.145)	
Similar to ingroup, observe ingroup, cognitive load	0.538 **	[0.174, 0.902]	0.166	[-0.127, 0.459]
	(0.186)		(0.150)	
Dissimilar to ingroup, observe ingroup, cognitive load	0.379	[-0.018, 0.776]	-0.012	[-0.299, 0.275]
	(0.202)		(0.146)	
Similar to ingroup, observe outgroup, cognitive load	-0.009	[-0.396, 0.378]	0.037	[-0.246, 0.321]
	(0.197)		(0.145)	
Dissimilar to ingroup, observe outgroup, cognitive load	-0.078	[-0.453, 0.297]	-0.191	[-0.473, 0.091]
	(0.191)		(0.144)	

Table S5: Likelihood of a social learner choosing optimally (log. regression). The omitted category is dissimilar to ingroup, observing the outgroup, and no cognitive load. The proportion of demonstrators choosing their optimal urn is centred by subtracting 0.5. The third and fifth columns show 95% confidence intervals. Those are based on robust standard errors clustered on the social learner (shown in parentheses in the second and fourth columns). Asterisks denote the level of significance of the p-values: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

	Zurich		**Nairobi**	
Parameter	Estimate	95 % CI	Estimate	95 % CI
Intercept	0.066	[-0.198, 0.330]	-0.058	[-0.262, 0.147]
	(0.135)		(0.104)	
Centered proportion of	2.522 ***	[2.012, 3.032]	0.138	[-0.232, 0.507]
dem. choosing optimally	(0.260)		(0.188)	
Similar to ingroup, observe	0.624 ***	[0.272, 0.977]	0.350 *	[0.080, 0.620]
ingroup, no cognitive load	(0.180)		(0.138)	
Dissimilar to ingroup, observe	0.377 *	[0.086, 0.669]	0.029	[-0.233, 0.291]
ingroup, no cognitive load	(0.149)		(0.134)	
Similar to ingroup, observe	0.359 *	[0.055, 0.664]	0.144	[-0.143, 0.431]
outgroup, no cognitive load	(0.155)		(0.146)	
Similar to ingroup, observe	0.566 **	[0.195, 0.937]	0.165	[-0.128, 0.458]
ingroup, cognitive load	(0.189)		(0.150)	
Dissimilar to ingroup, observe	0.351	[-0.058, 0.760]	-0.011	[-0.299, 0.276]
ingroup, cognitive load	(0.208)		(0.147)	
Similar to ingroup, observe	-0.053	[-0.456, 0.350]	0.041	[-0.244, 0.326]
outgroup, cognitive load	(0.205)		(0.145)	
Dissimilar to ingroup, observe	-0.083	[-0.474, 0.308]	-0.191	[-0.474, 0.091]
outgroup, cognitive load	(0.199)		(0.144)	

Table S6: Likelihood of a social learner choosing optimally (log. regression with interaction terms). The abbreviations "IG", "OG", "CL", and "NCL" refer to "ingroup", "outgroup", "cognitive load", and "no cognitive load" respectively. The omitted category is dissimilar IG, observe OG, NCL. The third and fifth columns show 95% confidence intervals, which are based on robust standard errors clustered on the social learner (shown in parentheses in the second and fourth columns). Asterisks denote the level of significance of the p-values: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

	Zurich		**Nairobi**	
Parameter	Estimate	95 % CI	Estimate	95 % CI
Intercept	0.119 (0.160)	[-0.207, 0.446]	-0.012 (0.109)	[-0.229 , 0.204]
Centered prop. dem. opt.	2.311 *** (0.504)	[1.213, 3.409]	-0.273 (0.449)	[-1.224, 0.678]
Similar to IG, observe IG, and NCL	0.263 (0.234)	[-0.201, 0.726]	0.074 (0.157)	[-0.221, 0.369]
Dissimilar to IG, observe IG, and NCL	0.214 (0.243)	[-0.239, 0.667]	-0.119 (0.155)	[-0.427, 0.190]
Similar to IG, observe OG, and NCL	0.338 (0.241)	[-0.108, 0.784]	0.194 (0.153)	[-0.115, 0.503]
Similar to IG, observe IG, and CL	0.060 (0.244)	[-0.403, 0.522]	-0.014 (0.158)	[-0.341, 0.313]
Dissimilar to IG, observe IG, and CL	0.456 (0.250)	[-0.044, 0.957]	-0.030 (0.157)	[-0.350, 0.289]
Similar to IG, observe OG, and CL	0.164 (0.244)	[-0.324, 0.652]	0.096 (0.154)	[-0.207, 0.398]
Dissimilar to IG, observe OG, and CL	0.214 (0.234)	[-0.241, 0.670]	-0.079 (0.157)	[-0.390, 0.232]
Centered prop. dem. opt. × similar IG, observe IG, NCL	1.610 * (0.783)	[-0.044, 3.264]	2.561 *** (0.668)	[1.105, 4.017]
Centered prop. dem. opt. × dissimilar IG, observe IG, NCL	0.643 (0.773)	[-0.844, 2.130]	1.409 * (0.663)	[0.068, 2.750]
Centered prop. dem. opt. × similar IG, observe OG, NCL	0.088 (0.758)	[-1.497, 1.673]	-0.651 (0.662)	[-2.269, 0.967]
Centered prop. dem. opt. × similar IG, observe IG, and CL	2.279 ** (0.833)	[0.382, 4.176]	1.576 * (0.653)	[0.107, 3.045]
Centered prop. dem. opt. × dissimilar IG, observe IG, CL	-0.397 (0.781)	[-2.124, 1.329]	0.160 (0.667)	[-1.335, 1.655]
Centered prop. dem. opt. × similar IG, observe OG, CL	-0.775 (0.752)	[-2.409, 0.859]	-0.746 (0.666)	[-2.303, 0.812]
Centered prop. dem. opt. × dissimilar IG, observe OG, CL	-1.138 (0.726)	[-2.662, 0.387]	-1.031 (0.649)	[-2.415, 0.353]

Table S7: Likelihood of a social learner choosing optimally (log. regression with control variables). The abbreviations "IG", "OG", "CL", and "NCL" refer to "ingroup", "outgroup", "cognitive load", and "no cognitive load" respectively. The omitted category is "dissimilar IG, observe OG, NCL". We grouped study subjects into three categories. The omitted category for the "study subject" variable is "Humanities". The third and fifth columns show 95% confidence intervals, which are based on robust standard errors clustered on the social learner (shown in parentheses in the second and fourth columns). Asterisks denote the level of significance of the p-values: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

	Zurich		**Nairobi**	
Parameter	Estimate	95 % CI	Estimate	95 % CI
Intercept	-0.521 (0.609)	[-1.715, 0.672]	0.325 (0.488)	[-0.632, 1.281]
Age	-0.012 (0.017)	[-0.039, 0.029]	-0.013 (0.022)	[-0.055, 0.031]
Gender 1 = "Man" ("Woman" = omitted)	0.104 (0.133)	[-0.157, 0.364]	-0.105 (0.079)	[-0.262, 0.051]
Gender 2 = "Other" ("Woman" = omitted)	0.062 (0.392)	[-0.706, 0.829]	NA (NA)	[NA, NA]
Natural Sciences ("Humanities" = omitted)	0.707 * (0.309)	[0.102, 1.312]	-0.055 (0.106)	[-0.262, 0.151]
Social Sciences ("Humanities" = omitted)	0.659 * (0.329)	[0.014, 1.304]	-0.084 (0.103)	[-0.286, 0.117]
Centered proportion of dem. choosing optimally	2.522 *** (0.265)	[2.002, 3.042]	0.142 (0.189)	[-0.229, 0.513]
Similar to IG, observe IG, and NCL	0.630 *** (0.180)	[0.277, 0.984]	0.350 * (0.138)	[0.080, 0.621]
Dissimilar to IG, observe IG, and NCL	0.380 * (0.150)	[0.086, 0.674]	0.029 (0.134)	[-0.233, 0.292]
Similar to IG, observe OG, and NCL	0.363 * (0.156)	[0.057, 0.669]	0.145 (0.147)	[-0.143, 0.432]
Similar to IG, observe IG, and CL	0.609 ** (0.192)	[0.233, 0.985]	0.158 (0.149)	[-0.134, 0.450]
Dissimilar to IG, observe IG, and CL	0.392 (0.214)	[-0.027, 0.810]	-0.019 (0.145)	[-0.303, 0.265]
Similar to IG, observe OG, and CL	-0.016 (0.209)	[-0.425, 0.393]	0.033 (0.144)	[-0.248, 0.315]
Dissimilar to IG, observe OG, and CL	-0.046 (0.204)	[-0.445, 0.354]	-0.199 (0.144)	[-0.481, 0.083]

Table S8: Pairwise comparisons of treatments in Zurich based on regression (2) and Table S4 of the supplementary materials. Standard errors are clustered on social learners which addresses repeated observations from the same individuals. The column Δpp reports the percentage point difference in the predicted probability of choosing the optimal urn for treatment 1 compared with treatment 2. Confidence intervals are Wald 95% intervals computed with a cluster robust covariance matrix that clusters on participants. The column p_{Bonf} reports Bonferroni adjusted p values across the 28 contrasts reported in this table. The column labelled Sim. to in. refers to similarity to ingroup. Sim refers to treatments where participants share the same winning colour with their ingroup. Dis refers to treatments where they do not. The column labelled Obs. gr. refers to observed group. In refers to treatments where participants observe ingroup demonstrators. Out refers to treatments where they observe outgroup demonstrators. The column labelled Cog. load refers to cognitive load. CL stands for cognitive load and NCL stands for no cognitive load.

Treatment 1			Treatment 2			Zurich		
Sim. to in.	Obs. gr.	Cog. load	Sim. to in.	Obs. gr.	Cog. load	Δpp	95% CI	p-value Bonferroni
Sim	In	NCL	Dis	In	NCL	3.3	[−2.48, 9.14]	1.000
Sim	In	NCL	Sim	Out	NCL	3.7	[−2.05, 9.50]	1.000
Sim	In	NCL	Sim	In	CL	1.0	[−4.83, 6.83]	1.000
Sim	In	NCL	Dis	In	CL	3.9	[−2.94, 10.65]	1.000
Sim	In	NCL	Sim	Out	CL	11.8	[4.40, 19.13]	0.034
Sim	In	NCL	Dis	Out	CL	13.3	[6.13, 20.48]	0.005
Dis	In	NCL	Sim	Out	NCL	0.4	[−5.69, 6.48]	1.000
Dis	In	NCL	Sim	In	CL	−2.3	[−8.78, 4.11]	1.000
Dis	In	NCL	Dis	In	CL	0.5	[−6.81, 7.86]	1.000
Dis	In	NCL	Sim	Out	CL	8.4	[0.57, 16.30]	0.930
Dis	In	NCL	Dis	Out	CL	10.0	[2.29, 17.66]	0.292
Sim	Out	NCL	Sim	In	CL	−2.7	[−9.46, 4.01]	1.000
Sim	Out	NCL	Dis	In	CL	0.1	[−7.46, 7.72]	1.000
Sim	Out	NCL	Sim	Out	CL	8.0	[−0.06, 16.14]	1.000
Sim	Out	NCL	Dis	Out	CL	9.6	[1.65, 17.51]	0.498
Sim	In	CL	Dis	In	CL	2.9	[−4.12, 9.83]	1.000
Sim	In	CL	Sim	Out	CL	10.8	[3.91, 17.63]	0.043
Sim	In	CL	Dis	Out	CL	12.3	[5.06, 19.55]	0.020
Dis	In	CL	Sim	Out	CL	7.9	[−0.04, 15.86]	1.000
Dis	In	CL	Dis	Out	CL	9.5	[2.40, 16.50]	0.250
Sim	Out	CL	Dis	Out	CL	1.5	[−5.56, 8.64]	1.000
Sim	In	NCL	Dis	Out	NCL	11.6	[4.66, 18.48]	0.021
Dis	In	NCL	Dis	Out	NCL	8.2	[2.34, 14.13]	0.152
Sim	Out	NCL	Dis	Out	NCL	7.8	[1.73, 13.96]	0.322
Sim	In	CL	Dis	Out	NCL	10.6	[3.28, 17.86]	0.106
Dis	In	CL	Dis	Out	NCL	7.7	[−0.37, 15.80]	1.000
Sim	Out	CL	Dis	Out	NCL	−0.2	[−8.76, 8.37]	1.000
Dis	Out	CL	Dis	Out	NCL	−1.7	[−10.14, 6.67]	1.000

Table S9: Pairwise comparisons of treatments in Nairobi based on regression (2) and Table S4 of the supplementary materials. Standard errors are clustered on social learners which addresses repeated observations from the same individuals. The column Δpp reports the percentage point difference in the predicted probability of choosing the optimal urn for treatment 1 compared with treatment 2. Confidence intervals are Wald 95% intervals computed with a cluster robust covariance matrix that clusters on participants. The column p_{Bonf} reports Bonferroni adjusted p values across the 28 contrasts reported in this table. The column labelled Sim. to in. refers to similarity to ingroup. Sim refers to treatments where participants share the same winning colour with their ingroup. Dis refers to treatments where they do not. The column labelled Obs. gr. refers to observed group. In refers to treatments where participants observe ingroup demonstrators. Out refers to treatments where they observe outgroup demonstrators. The column labelled Cog. load refers to cognitive load. CL stands for cognitive load and NCL stands for no cognitive load.

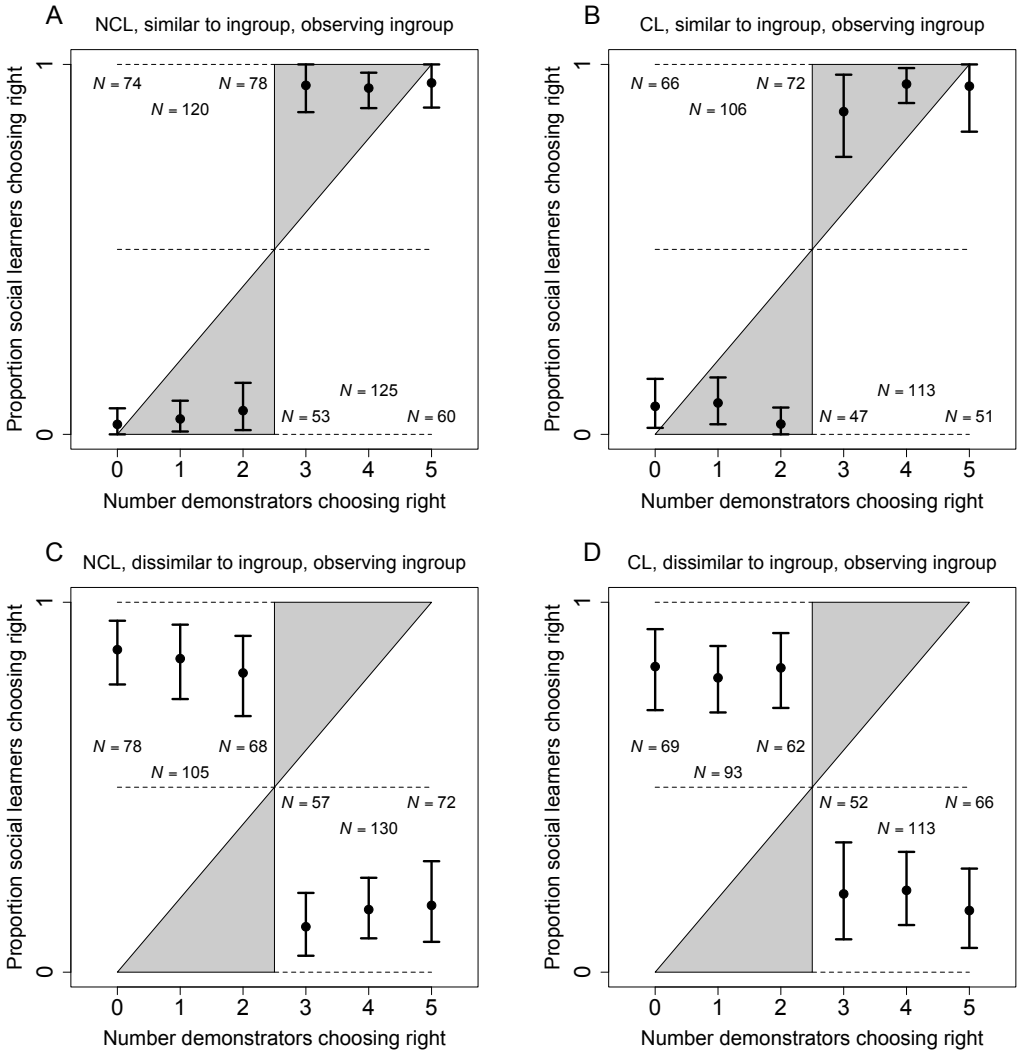
Treatment 1			Treatment 2			Nairobi		
Sim. to in.	Obs. gr.	Cog. load	Sim. to in.	Obs. gr.	Cog. load	Δpp	95% CI	p-value Bonferroni
Sim	In	NCL	Dis	In	NCL	8.0	[0.98, 15.02]	0.699
Sim	In	NCL	Sim	Out	NCL	5.2	[−2.14, 12.50]	1.000
Sim	In	NCL	Sim	In	CL	4.6	[−2.69, 11.81]	1.000
Sim	In	NCL	Dis	In	CL	9.0	[1.90, 16.11]	0.360
Sim	In	NCL	Sim	Out	CL	7.8	[0.76, 14.78]	0.821
Sim	In	NCL	Dis	Out	CL	13.4	[6.52, 20.38]	0.004
Dis	In	NCL	Sim	Out	NCL	−2.8	[−10.63, 4.99]	1.000
Dis	In	NCL	Sim	In	CL	−3.4	[−10.99, 4.11]	1.000
Dis	In	NCL	Dis	In	CL	1.0	[−6.41, 8.42]	1.000
Dis	In	NCL	Sim	Out	CL	−0.2	[−7.55, 7.09]	1.000
Dis	In	NCL	Dis	Out	CL	5.4	[−1.79, 12.69]	1.000
Sim	Out	NCL	Sim	In	CL	−0.6	[−8.29, 7.06]	1.000
Sim	Out	NCL	Dis	In	CL	3.8	[−3.71, 11.37]	1.000
Sim	Out	NCL	Sim	Out	CL	2.6	[−4.86, 10.05]	1.000
Sim	Out	NCL	Dis	Out	CL	8.3	[0.90, 15.65]	0.768
Sim	In	CL	Dis	In	CL	4.4	[−3.01, 11.90]	1.000
Sim	In	CL	Sim	Out	CL	3.2	[−3.59, 10.01]	1.000
Sim	In	CL	Dis	Out	CL	8.9	[2.32, 15.46]	0.220
Dis	In	CL	Sim	Out	CL	−1.2	[−8.95, 6.48]	1.000
Dis	In	CL	Dis	Out	CL	4.4	[−2.09, 10.98]	1.000
Sim	Out	CL	Dis	Out	CL	5.7	[−1.59, 12.95]	1.000
Sim	In	NCL	Dis	Out	NCL	8.7	[1.96, 15.45]	0.314
Dis	In	NCL	Dis	Out	NCL	0.7	[−5.87, 7.28]	1.000
Sim	Out	NCL	Dis	Out	NCL	3.5	[−3.64, 10.70]	1.000
Sim	In	CL	Dis	Out	NCL	4.1	[−3.23, 11.52]	1.000
Dis	In	CL	Dis	Out	NCL	−0.3	[−7.53, 6.93]	1.000
Sim	Out	CL	Dis	Out	NCL	0.9	[−6.20, 8.07]	1.000
Dis	Out	CL	Dis	Out	NCL	−4.7	[−11.80, 2.31]	1.000

S2.2.4 Pure behavior

Table S10: Likelihood of a social learner choosing the right urn (logit regression). The abbreviations "IG", "OG", "CL", and "NCL" refer to "ingroup", "outgroup", "cognitive load", and "no cognitive load" respectively. The omitted category is dissimilar IG, observe OG, NCL. "Dem. right" stands for the centered proportion of demonstrators choosing the right urn. The third and fifth columns show 95% confidence intervals, which are based on robust standard errors clustered on the social learner (shown in parentheses in the second and fourth columns). Asterisks denote the level of significance of the p-values: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

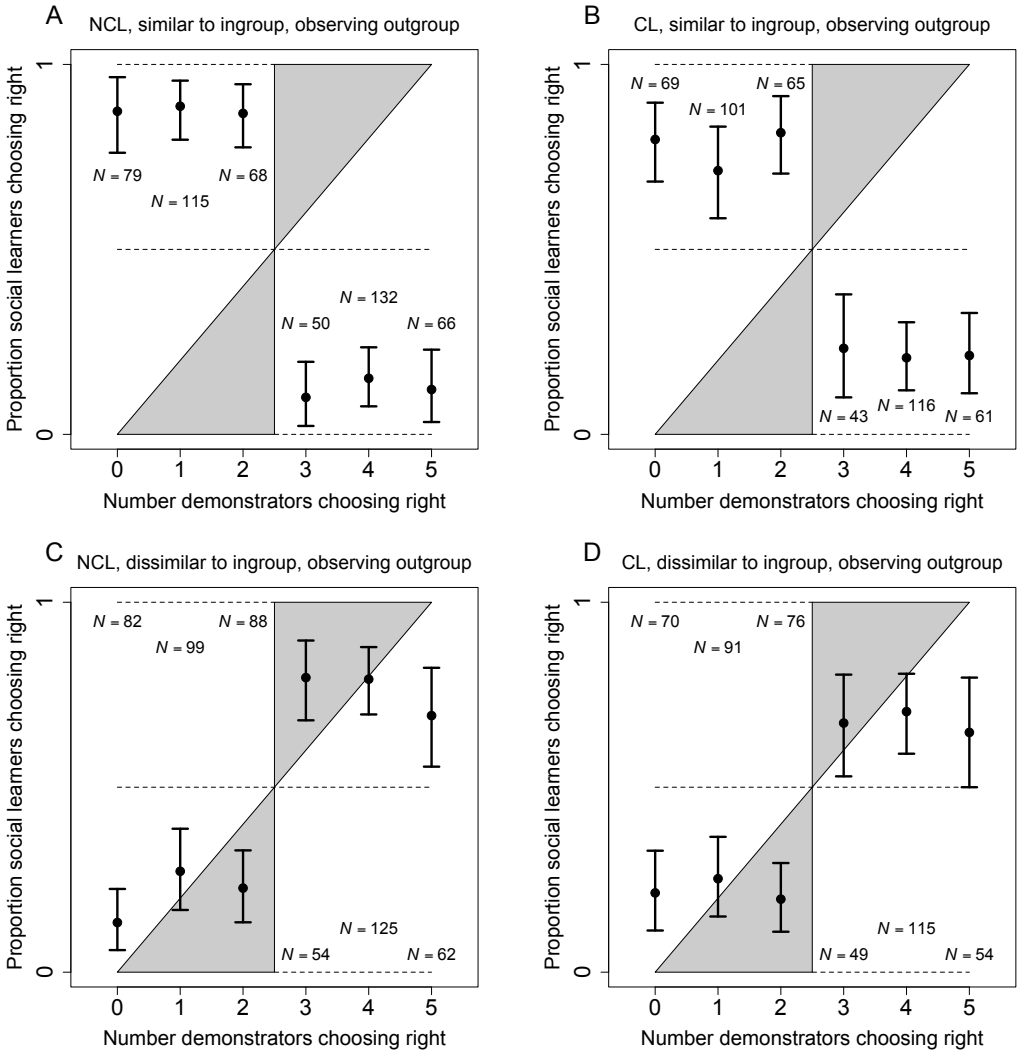
	Zurich		**Nairobi**	
Parameter	Estimate	95 % CI	Estimate	95 % CI
Intercept	-0.087 (0.087)	[-0.257, 0.084]	-0.035 (0.094)	[-0.219 , 0.149]
Dem. right	3.494 *** (0.639)	[2.241, 4.747]	-0.620 (0.475)	[-1.552, 0.312]
Similar to IG, observe IG, and NCL	-0.145 (0.183)	[-0.504, 0.214]	0.081 (0.128)	[-0.169, 0.331]
Dissimilar to IG, observe IG, and NCL	0.123 (0.150)	[-0.170, 0.417]	0.196 (0.129)	[-0.057, 0.449]
Similar to IG, observe OG, and NCL	0.169 (0.135)	[-0.095, 0.433]	0.074 (0.141)	[-0.202, 0.351]
Similar to IG, observe IG, and CL	-0.063 (0.150)	[-0.358, 0.232]	0.257 (0.152)	[-0.040, 0.554]
Dissimilar to IG, observe IG, and CL	0.147 (0.135)	[-0.117, 0.412]	0.268 (0.146)	[-0.017, 0.554]
Similar to IG, observe OG, and CL	0.068 (0.153)	[-0.232, 0.368]	0.188 (0.146)	[-0.098, 0.474]
Dissimilar to IG, observe OG, and CL	-0.164 (0.135)	[-0.428, 0.101]	0.045 (0.143)	[-0.236, 0.326]
Dem. right × similar IG, observe IG, NCL	6.693 *** (1.73)	[3.312, 10.075]	3.641 *** (0.787)	[2.098, 5.185]
Dem. right × dissimilar IG, observe IG, NCL	-8.241 *** (1.24)	[-10.675, -5.807]	-0.162 (0.718)	[-1.569, 1.246]
Dem. right × similar IG, observe OG, NCL	-9.059 *** (1.42)	[-11.849, -6.269]	1.393 (0.833)	[-0.238, 3.025]
Dem. right × similar IG, observe IG, and CL	4.923 ** (1.66)	[1.662, 8.184]	2.463 *** (0.729)	[1.034, 3.892]
Dem. right × dissimilar IG, observe IG, CL	-7.507 *** (0.944)	[-9.356, -5.657]	0.591 (0.721)	[-0.822, 2.005]
Dem. right × similar IG, observe OG, CL	-6.917 *** (0.906)	[-8.693, -5.141]	1.326 (0.773)	[-0.190, 2.841]
Dem. right × dissimilar IG, observe OG, CL	-0.767 (0.851)	[-2.436, 0.902]	-0.888 (0.704)	[-2.268, 0.491]

Figure S1: Urn choices by social learners observing ingroup (results from Zurich)



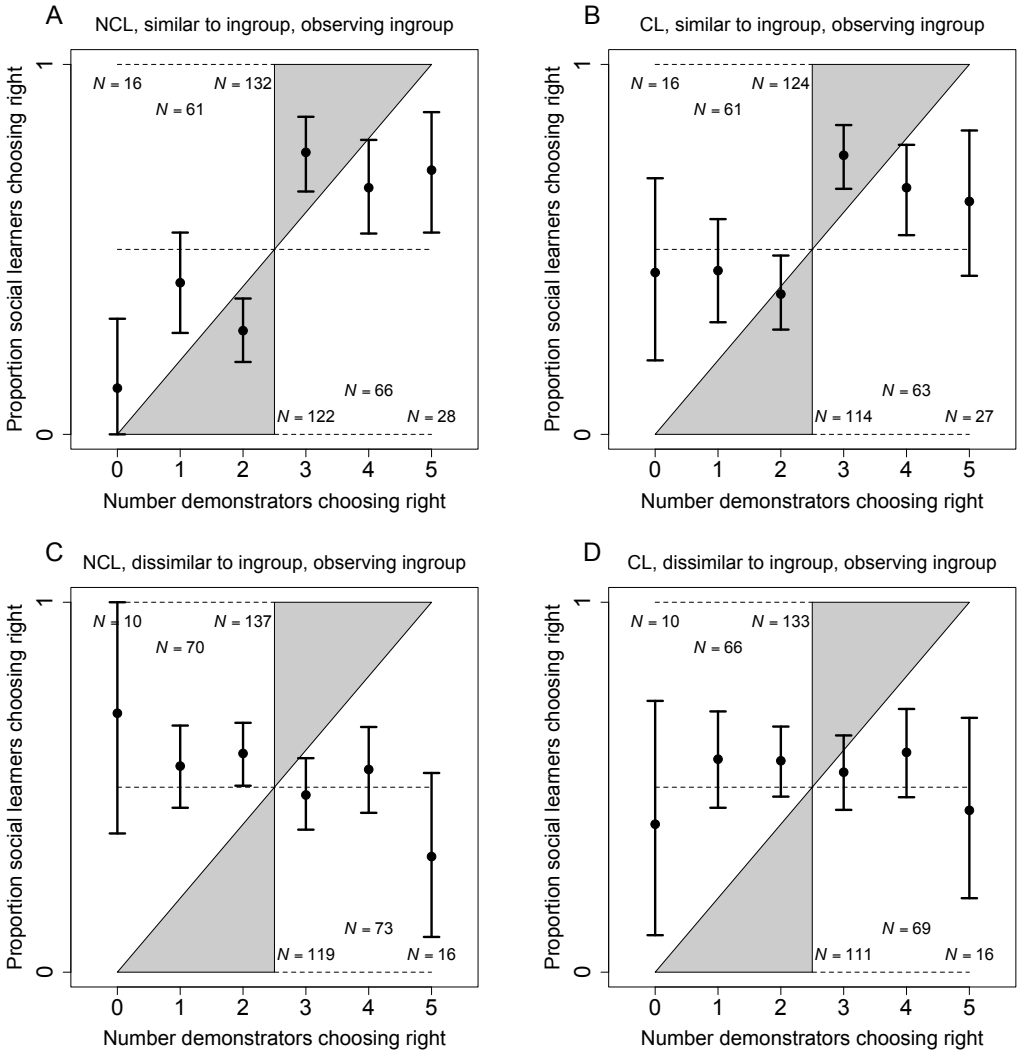
Note. The points show the rate at which social learners chose the right urn as a function of the number of demonstrators choosing the right urn in the final round of blocks. The N labels are vertically aligned with each point, and they count social-learner observations for a given number of observed demonstrators. The number of observations is vertically aligned with each point. **A** and **B** show treatments in which social learners are **similar** to ingroup demonstrators and they observe the ingroup. **C** and **D** show treatments in which social learners are **dissimilar** to their ingroup and they observe the ingroup. **A** and **C** represent "no cognitive load" treatments (NCL) and **B** and **D** show "cognitive load" treatments (CL). The error bars are 95% bootstrapped confidence intervals which we clustered on social learner. The gray area in each plot is consistent with conformist cultural transmission Boyd & Richerson (1985). The diagonal shown as a solid black line is consistent with unbiased social learning. Finally, the dashed lines provide additional points of reference at 0, 0.5, and 1.

Figure S2: Urn choices by social learners observing outgroup (results from Zurich)



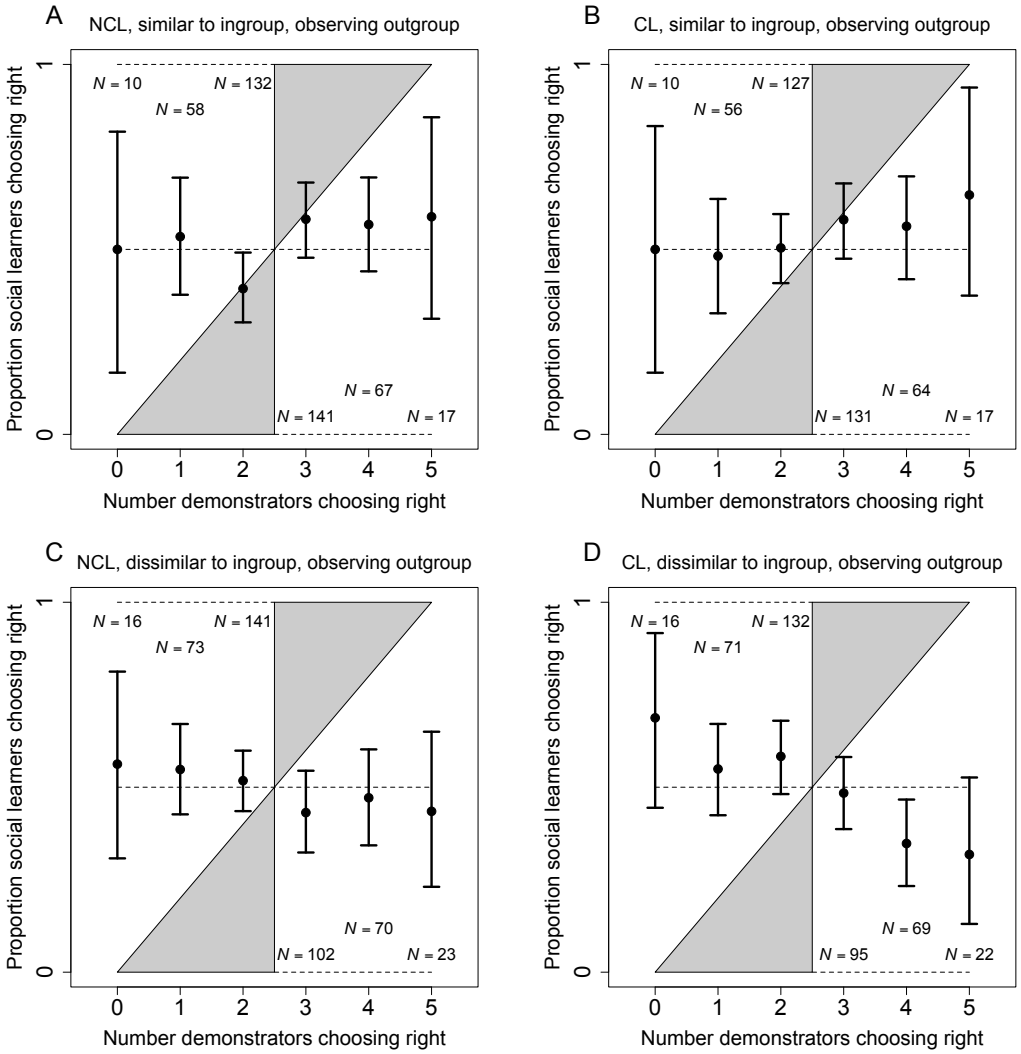
Note. The points show the rate at which social learners chose the right urn as a function of the number of demonstrators choosing the right urn in the final round of blocks. The N labels are vertically aligned with each point, and they count social-learner observations for a given number of observed demonstrators. **A** and **B** show treatments in which social learners are **similar** to the ingroup and they observe the outgroup. **C** and **D** show treatments in which social learners are **dissimilar** to the ingroup and they observe the outgroup. **A** and **C** represent "no cognitive load" treatments (NCL) and **B** and **D** show "cognitive load" treatments (CL). Error bars are 95% bootstrapped confidence intervals clustered on social learner. The gray area in each plot is consistent with conformist cultural transmission Boyd & Richerson (1985). The diagonal shown as a solid black line is consistent with unbiased social learning. Finally, the dashed lines provide additional points of reference at 0, 0.5, and 1.

Figure S3: Urn choices by social learners observing ingroup (results from Nairobi)



Note. The points show the rate at which social learners chose the right urn as a function of the number of demonstrators choosing the right urn in the final round of blocks. The N labels are vertically aligned with each point, and they count social-learner observations for a given number of observed demonstrators. **A** and **B** show treatments in which social learners are **similar** to ingroup demonstrators and they observe the ingroup. **C** and **D** show treatments in which social learners are **dissimilar** to their ingroup and they observe the ingroup. **A** and **C** represent "no cognitive load" treatments (NCL) and **B** and **D** show "cognitive load" treatments (CL). The error bars are 95% bootstrapped confidence intervals which we clustered on social learner. The gray area in each plot is consistent with conformist cultural transmission Boyd & Richerson (1985). The diagonal shown as a solid black line is consistent with unbiased social learning. Finally, the dashed lines provide additional points of reference at 0, 0.5, and 1.

Figure S4: Urn choices by social learners observing outgroup (results from Nairobi)



Note. The points show the rate at which social learners chose the right urn as a function of the number of demonstrators choosing the right urn in the final round of blocks. The N labels are vertically aligned with each point, and they count social-learner observations for a given number of observed demonstrators. **A** and **B** show treatments in which social learners are **similar** to the ingroup and they observe the outgroup. **C** and **D** show treatments in which social learners are **dissimilar** to the ingroup and they observe the outgroup. **A** and **C** represent "no cognitive load" treatments (NCL) and **B** and **D** show "cognitive load" treatments (CL). Error bars are 95% bootstrapped confidence intervals clustered on social learner. The gray area in each plot is consistent with conformist cultural transmission Boyd & Richerson (1985). The diagonal shown as a solid black line is consistent with unbiased social learning. Finally, the dashed lines provide additional points of reference at 0, 0.5, and 1.

S2.2.5 Signal quality across blocks and cross-site differences

One possible explanation for the larger share of unconditional and weakly frequency-dependent learners in Nairobi is that lower demonstrator performance generated a weaker experienced social signal across blocks (c.f. section S2.1 in this Supplementary Material). To examine this possibility, we exploited heterogeneity in the sequence of observed demonstrator choices over blocks. For each social learner and block $t > 1$, we computed (i) current signal strength, defined as the absolute value of the centred proportion of observed demonstrators choosing the right urn, (ii) the cumulative prior mean signal strength across blocks $1, \ldots, t-1$, and (iii) the cumulative prior share of weak 3–2 majorities across blocks $1, \ldots, t-1$. We then estimated site-specific logistic regressions predicting whether a learner followed the socially indicated response in block t—that is, the majority when the observed group was incentive-similar and the minority when it was incentive-dissimilar—and whether the learner chose the optimal urn in block t, controlling for treatment cell and block number.

Table S11 summarises the estimation sample after excluding first-block observations, for which prior signal history is undefined. The first two rows report the size of the estimation sample. The next two rows describe the experienced signal environment: prior mean signal strength was higher in Zurich than in Nairobi (0.292 vs. 0.202), whereas the prior share of weak 3–2 signals was substantially higher in Nairobi than in Zurich (0.553 vs. 0.276). The final two rows report mean values of the two dependent variables in the estimation sample. Relative to Nairobi, social learners in Zurich more often followed the socially indicated response (0.839 vs. 0.527) and more often chose the optimal urn (0.734 vs. 0.510). Thus, the descriptive summary indicates that the signal environment was stronger in Zurich than in Nairobi and that Zurich social learners were, on average, more responsive to social information.

Table S12 reports the corresponding regression results. Across both panels, the coefficients on the history variables are small and statistically indistinguishable from zero in both sites and for both outcomes. Neither prior mean signal strength nor prior exposure to weak 3–2 signals was associated with later reliance on social information or with later task performance

in either site. Current signal strength was not associated with following the socially indicated response in either site. For optimal choice, however, current signal strength was positively associated with accuracy in Zurich but not in Nairobi. This pattern does not alter the main conclusion of this subsection. First, the central question here is whether accumulated differences in the experienced quality of the social signal across earlier blocks help explain the cross-site difference in the distribution of social learning strategies, and the relevant history measures do not predict either later reliance on social information or later task performance in either site. Second, the statistically significant Zurich result concerns contemporaneous signal strength in the current block rather than cumulative signal history. It therefore indicates that clearer current demonstrations were more informative for realised accuracy in Zurich, but it does not imply that learners in Zurich became more reliant on social information because of stronger signals encountered earlier in the experiment. Taken together, the results suggest that variation in the experienced history of signal quality across blocks is unlikely to account for the larger share of unconditional and weakly frequency-dependent learners in Nairobi.

Table S11: Descriptive summary of signal history variables used in the sequence-based analysis. Observations from the first block are excluded because prior signal history is undefined for those observations.

Statistic	Zurich	Nairobi
Observations	3667	3154
Social learners	193	166
Mean prior signal strength	0.292	0.202
Mean prior share of weak signals	0.276	0.553
Proportion following socially indicated response	0.839	0.527
Proportion choosing optimal urn	0.734	0.510

Note. Observations from block 1 are excluded because prior signal history is undefined for those observations. After this exclusion, each social learner contributes 19 observations. The first two rows report the number of observations and distinct social learners in the estimation sample. Prior mean signal strength is the average across preceding blocks of the current-block signal-strength measure, defined as the absolute value of the centred proportion of observed demonstrators choosing the right urn (0.1 = 3–2, 0.3 = 4–1, 0.5 = 5–0). Prior share of weak signals is the proportion of preceding blocks with a weak 3–2 majority. The final two rows report the mean of the two dependent variables in the estimation sample. "Following the socially indicated response" equals majority-following when the observed group was incentive-similar and minority-following when it was incentive-dissimilar. "Choosing the optimal urn" records realised task performance in the current block.

Table S12: Sequence-based analyses of signal history. Logistic regressions were estimated separately by site with robust standard errors clustered by social learner. All models include the eight treatment-cell controls and block number. Only coefficients of interest are shown.

Outcome	Predictor	Zurich		Nairobi	
		Est.	95% CI	Est.	95% CI
Panel A. Prior mean signal strength					
Follow socially indicated response	Current signal strength	-0.052	[-0.686, 0.581]	-0.309	[-0.896, 0.277]
	Prior mean signal strength	-0.364	[-3.190, 2.460]	-0.330	[-1.820, 1.150]
Choose optimal urn	Current signal strength	1.512***	[0.974, 2.050]	-0.146	[-0.752, 0.460]
	Prior mean signal strength	0.744	[-0.731, 2.220]	-0.473	[-1.950, 1.010]
Panel B. Prior share of weak signals					
Follow socially indicated response	Current signal strength	-0.056	[-0.702, 0.590]	-0.316	[-0.904, 0.272]
	Prior share of weak signals	0.110	[-0.829, 1.050]	-0.016	[-0.370, 0.338]
Choose optimal urn	Current signal strength	1.526***	[0.988, 2.060]	-0.149	[-0.756, 0.459]
	Prior share of weak signals	-0.162	[-0.640, 0.317]	0.053	[-0.287, 0.394]

Note. Each panel reports a separate set of models. The coefficient on current signal strength therefore differs slightly across panels because it is estimated jointly with a different history measure in each panel. Panel A uses cumulative prior mean signal strength as the history measure. Panel B uses the cumulative prior share of weak 3–2 signals. Current signal strength is defined as the absolute value of the centred proportion of observed demonstrators choosing the right urn in the current block, such that 0.1 corresponds to a 3–2 split, 0.3 to a 4–1 split, and 0.5 to a 5–0 split. Prior mean signal strength is the average of this measure across preceding blocks. Prior share of weak signals is the proportion of preceding blocks with a weak 3–2 majority. "Follow socially indicated response" equals 1 when the learner followed the majority if the observed group was incentive-similar and followed the minority if the observed group was incentive-dissimilar. Asterisks denote significance levels (* $p < .05$, ** $p < .01$, *** $p < .001$). Exact p-values for the focal predictors are as follows. Panel A, Zurich: current signal strength $p = .8711022$ for following the socially indicated response and $p = .00000003706$ for choosing the optimal urn; prior mean signal strength $p = .8004843$ and $p = .323030$, respectively. Panel A, Nairobi: current signal strength $p = .30136$ and $p = .636456$; prior mean signal strength $p = .66273$ and $p = .531369$. Panel B, Zurich: current signal strength $p = .8648286$ and $p = .00000002768$; prior share of weak signals $p = .8181553$ and $p = .507840$. Panel B, Nairobi: current signal strength $p = .29237$ and $p = .631525$; prior share of weak signals $p = .92959$ and $p = .758967$.

S2.2.6 Regularised hierarchical estimation of social learning curves

We estimate continuous individual response curves that map observed demonstrator frequencies to the probability that a social learner chooses the right urn. Each participant contributes only a small number of trials per experimental cell. Consequently, estimating a separate logistic regression for each participant and condition can be noisy and can suffer from quasi-separation (leading to extremely large or undefined maximum-likelihood estimates). To obtain stable individual-level curves while still allowing heterogeneity across individuals, we fit a hierarchical (mixed-effects) logistic model. This approach is regularised in the sense that participant-level parameters are partially pooled (shrunk) toward population-level estimates, with the degree of shrinkage determined by the inferred between-participant variance and the amount of information available for each participant.

Model specification (trial-level). Let $y_{it} \in \{0, 1\}$ indicate whether social learner i chose the right urn in trial t. Let $k_{it} \in \{0, \ldots, 5\}$ denote the number of demonstrators (out of 5) choosing the right urn, and define the centered predictor $x_{it} = k_{it} - 2.5$. The value 2.5 is not an observed count; it is the midpoint between $k = 2$ and $k = 3$ (an even split, i.e., 50% choose right), so $x_{it} = 0$ corresponds to an evenly split demonstration. We estimate the following pooled logistic mixed-effects model with treatment-cell-specific fixed effects and participant-level random effects:

$$\operatorname{logit}\big(\Pr(y_{it} = 1)\big) = \alpha_{c_{it}} + u_i + \big(\beta_{c_{it}} + b_{i,\mathrm{ctx}_{it}}\big)\, x_{it}. \tag{S3}$$

Here c_{it} indexes the 8 treatment cells (Sim/Dis × In/Out × CL/NCL), and ctx_{it} indexes the 4 within-subject contexts (Sim_In, Dis_In, Sim_Out, Dis_Out). Because the random slopes are indexed by the four within-subject contexts rather than the full eight treatment cells, the same participant-specific slope deviation applies to the CL and NCL versions of a given context.

Meaning of α_c and β_c (population-level fixed effects). The parameters α_c and β_c are condition-specific fixed effects (population-level means) for treatment cell c. The intercept α_c is the baseline log-odds of choosing the right urn in condition c when $x_{it} = 0$, i.e., when demonstrators are evenly split. Thus, for an "average" participant (random effects equal to zero), the predicted probability of choosing right at an even split is $\text{logit}^{-1}(\alpha_c)$. The slope β_c captures how strongly, on average in condition c, the log-odds of choosing the right urn change as more demonstrators choose the right urn. Under $x_{it} = k_{it} - 2.5$, β_c is the change in log-odds associated with one additional demonstrator choosing right (a +1 change in k).

Random effects, diagonal covariance, and "regularisation". The random intercept u_i captures persistent individual differences in baseline choice propensity (e.g., a general tendency for participant i to choose the right urn irrespective of demonstrator composition). With the intercept included, participant i's baseline log-odds at an even split are $\alpha_c + u_i$. The context-specific random slopes $b_{i,\text{ctx}}$ capture heterogeneity in responsiveness to social information, allowing participants to differ in how strongly they react to x_{it}, potentially differently across the four within-subject contexts. Under our primary (diagonal) random-effects specification, we estimate a separate slope-variance parameter for each context and impose zero correlations among random effects. Concretely, we assume $u_i \sim \mathcal{N}(0, \sigma_u^2)$, $b_{i,\text{Sim_In}} \sim \mathcal{N}(0, \sigma^2_{\text{Sim_In}})$, $b_{i,\text{Dis_In}} \sim \mathcal{N}(0, \sigma^2_{\text{Dis_In}})$, $b_{i,\text{Sim_Out}} \sim \mathcal{N}(0, \sigma^2_{\text{Sim_Out}})$, and $b_{i,\text{Dis_Out}} \sim \mathcal{N}(0, \sigma^2_{\text{Dis_Out}})$.

All random effects are assumed mutually independent; that is, $\text{Cov}(u_i, b_{i,\text{ctx}}) = 0$ and $\text{Cov}(b_{i,\text{ctx}}, b_{i,\text{ctx}'}) = 0$ for $\text{ctx} \neq \text{ctx}'$. This diagonal structure reduces the number of random-effects covariance parameters to be estimated. The analysis code also supports a full covariance specification that allows correlations across context-specific random slopes. Unless stated otherwise, we report results from the diagonal specification.

This hierarchical structure induces partial pooling: participant-level random effects are shrunk toward zero, especially when a participant provides limited information. In curve

space, zero random effects correspond to the fixed-effects-only population curve. The variance parameters σ_u^2 and $\sigma^2_{\text{Sim_In}}, \sigma^2_{\text{Dis_In}}, \sigma^2_{\text{Sim_Out}}, \sigma^2_{\text{Dis_Out}}$ quantify between-participant heterogeneity in baseline propensity and in social responsiveness, respectively.

In both subject pools we begin by fitting the hierarchical model in Eq. S3 with a participant random intercept u_i and context-specific random slopes $b_{i,\text{ctx}}$. In the Zurich sample, the estimated random-intercept variance collapses to (near) zero (a boundary/singular fit), so we report the refitted model without the random intercept. In the Nairobi sample, the estimated random-intercept variance is non-zero, so we retain the random intercept in the reported specification.

From parameters to curves. For each condition, we obtain participant-level curves by evaluating the fitted model over a fine grid of k values from 0 to 5 (and converting to x accordingly), using the condition-specific fixed effects and each participant's estimated random effects. The blue curves in the figures are these participant-level (partially pooled) predicted response curves.

The fitted model also yields two natural population-level curves for each condition. First, one can set all random effects to zero and obtain a fixed-effects-only curve. This is the curve corresponding to an "average" participant and, in parameter space, it is the curve toward which individual random effects are shrunk. Second, one can average the participant-level predicted probabilities at each value of k, yielding a marginal mean curve. Because the logistic link is nonlinear, these two population-level curves need not coincide. In the figures we plot the marginal mean as the solid black curve, because it directly summarises the average predicted behaviour across participants and therefore provides the most intuitive descriptive comparison to the blue participant-level curves. We also estimated the fixed-effects-only population curve, but do not show it in the final figures in order to avoid clutter. Each panel reports $N =$ the number of participants contributing data to that treatment condition. The plots are shown in the main text.

Table S13: Hierarchical logistic mixed-effects model estimates for the Zurich social learners (trial-level binary outcome: choosing the right urn). Estimates correspond to the model in Eq. S3. Panel A reports condition-level fixed effects: $\hat{\alpha}$ is the log-odds of choosing right when social information is neutral ($x = 0$, corresponding to an evenly split demonstration; midpoint between $k = 2$ and $k = 3$), and $\hat{\beta}$ is the change in log-odds per additional demonstrator choosing right (per $+1$ in k). Panel B reports estimated between-participant standard deviations (and variances) of the random effects used for partial pooling. Under the diagonal specification, random-effects correlations are constrained to zero; Panel B therefore reports one SD/variance for each context-specific random slope (Sim_In, Dis_In, Sim_Out, Dis_Out). In the Zurich specification the participant random intercept is omitted because its estimated variance collapsed to (near) zero (singular fit). Predictor: $x = k - 2.5$.

Panel A: Population-level fixed effects

Condition	$\hat{\alpha}$ (SE)	$\hat{\beta}$ (SE)	95% CI for β
Sim_In_NCL	-0.06 (0.37)	11.86 (2.63)	[6.72, 17.01]
Dis_In_NCL	0.14 (0.20)	-3.95 (0.59)	[-5.10, -2.80]
Sim_Out_NCL	-0.14 (0.24)	-5.50 (0.90)	[-7.25, -3.75]
Dis_Out_NCL	0.15 (0.18)	2.98 (0.49)	[2.03, 3.94]
Sim_In_CL	0.06 (0.31)	11.37 (2.69)	[6.11, 16.64]
Dis_In_CL	-0.01 (0.19)	-3.27 (0.56)	[-4.37, -2.17]
Sim_Out_CL	0.23 (0.22)	-4.11 (0.75)	[-5.57, -2.65]
Dis_Out_CL	-0.30 (0.17)	2.19 (0.46)	[1.29, 3.08]

Panel B: Between-participant variability (random effects)

Random effect	SD	Variance
Sim_In	10.23	104.62
Dis_In	3.67	13.50
Sim_Out	5.14	26.47
Dis_Out	3.21	10.28

Table S14: Hierarchical logistic mixed-effects model estimates for the Nairobi (Busara) social learners (trial-level binary outcome: choosing the right urn). Estimates correspond to the model in Eq. S3. Panel A reports condition-level fixed effects: $\hat{\alpha}$ is the log-odds of choosing right when social information is neutral ($x = 0$, corresponding to an evenly split demonstration; midpoint between $k = 2$ and $k = 3$), and $\hat{\beta}$ is the change in log-odds per additional demonstrator choosing right (per +1 in k). Panel B reports estimated between-participant standard deviations (and variances) of the random effects used for partial pooling. Under the diagonal specification, random-effects correlations are constrained to zero; Panel B therefore reports one SD/variance for each context-specific random slope and the participant random intercept. Predictor: $x = k - 2.5$.

Panel A: Population-level fixed effects

Condition	$\hat{\alpha}$ (SE)	$\hat{\beta}$ (SE)	95% CI for β
Sim_In_NCL	0.15 (0.12)	1.10 (0.20)	[0.71, 1.48]
Dis_In_NCL	0.18 (0.11)	-0.22 (0.13)	[-0.47, 0.03]
Sim_Out_NCL	0.02 (0.13)	0.39 (0.24)	[-0.07, 0.85]
Dis_Out_NCL	-0.01 (0.11)	-0.16 (0.12)	[-0.40, 0.08]
Sim_In_CL	0.28 (0.12)	0.73 (0.19)	[0.36, 1.10]
Dis_In_CL	0.24 (0.12)	-0.01 (0.13)	[-0.26, 0.24]
Sim_Out_CL	0.13 (0.13)	0.33 (0.24)	[-0.14, 0.80]
Dis_Out_CL	0.04 (0.12)	-0.38 (0.13)	[-0.63, -0.13]

Panel B: Between-participant variability (random effects)

Random effect	SD	Variance
Intercept	0.30	0.09
Sim_In	1.19	1.42
Dis_In	0.66	0.44
Sim_Out	1.75	3.07
Dis_Out	0.67	0.45

S2.2.7 Social learning strategies disaggregated by social learner

Discrete strategy classification (Maj/Min/U) and per-learner slope estimates. In addition to the hierarchical response-curve analysis above, we report a complementary disaggregated summary of individual social-learning strategies. The goal is descriptive, namely to identify social learners whose behaviour is perfectly consistent with simple deterministic rules, and to summarise the remaining (non-deterministic) learner-condition cells with a single slope parameter.

Fix a treatment cell (condition) c (Sim/Dis × In/Out × CL/NCL). Within this cell, index social learners by k and let $y_{kj} \in \{0, 1\}$ indicate whether learner k chose the right urn in social-learning block j. Let $d_{kj} \in \{0, \ldots, 5\}$ denote the number of demonstrators (out of 5) choosing the right urn in that block, and define the centred demonstrator proportion

$$x_{kj} \;=\; \frac{d_{kj}}{5} - 0.5.$$

Because there are 5 demonstrators, $x_{kj} \neq 0$ in the data; $x_{kj} > 0$ means a strict majority of demonstrators chose the right urn and $x_{kj} < 0$ means a strict majority chose the left urn.

For each learner-cell pair (k, c) we classify behaviour using the following perfect (deterministic) criteria: (i) **Unconditional (U)** if the learner always chose the same urn in that cell (i.e., $y_{kj} = 0$ for all j or $y_{kj} = 1$ for all j); (ii) **Majority-following (Maj)** if $y_{kj} = 1$ for all blocks with $x_{kj} > 0$ and $y_{kj} = 0$ for all blocks with $x_{kj} < 0$; (iii) **Minority-following (Min)** if $y_{kj} = 1$ for all blocks with $x_{kj} < 0$ and $y_{kj} = 0$ for all blocks with $x_{kj} > 0$. All remaining learner-cell pairs are labelled **Other**.

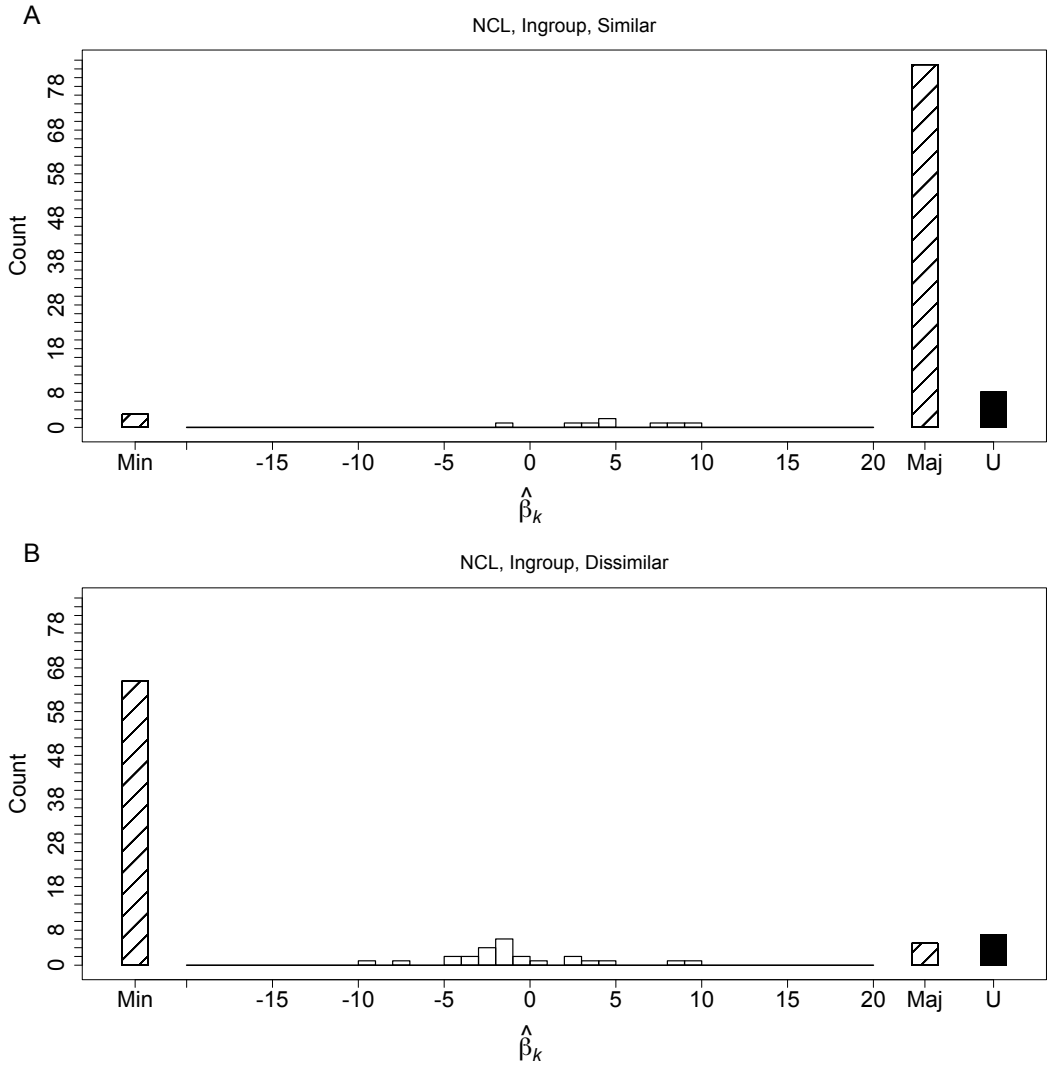
For the **Other** learner-cell pairs, we estimate a single per-learner slope parameter β_k by maximum likelihood from a logistic model without an intercept,

$$\operatorname{logit}\big(\Pr(y_{kj} = 1)\big) \;=\; \beta_k \, x_{kj}, \tag{S4}$$

which implies $\Pr(y_{kj} = 1) = \exp(\beta_k x_{kj})/(1 + \exp(\beta_k x_{kj}))$ and fixes $\Pr(y_{kj} = 1) = 0.5$ at $x_{kj} = 0$ (an evenly split demonstration). This specification omits an intercept and therefore does not separately model moderate left/right side biases within a learner-condition cell. We report $\hat{\beta}_k$ along with its Wald p-value from the fitted GLM. Positive $\hat{\beta}_k$ indicates that the learner is more likely to choose the right urn when demonstrator support for the right urn increases (alignment with the demonstrators' majority), while negative $\hat{\beta}_k$ indicates the opposite (tending to choose against the demonstrators' majority). The magnitude $|\hat{\beta}_k|$ reflects the steepness of the estimated response.

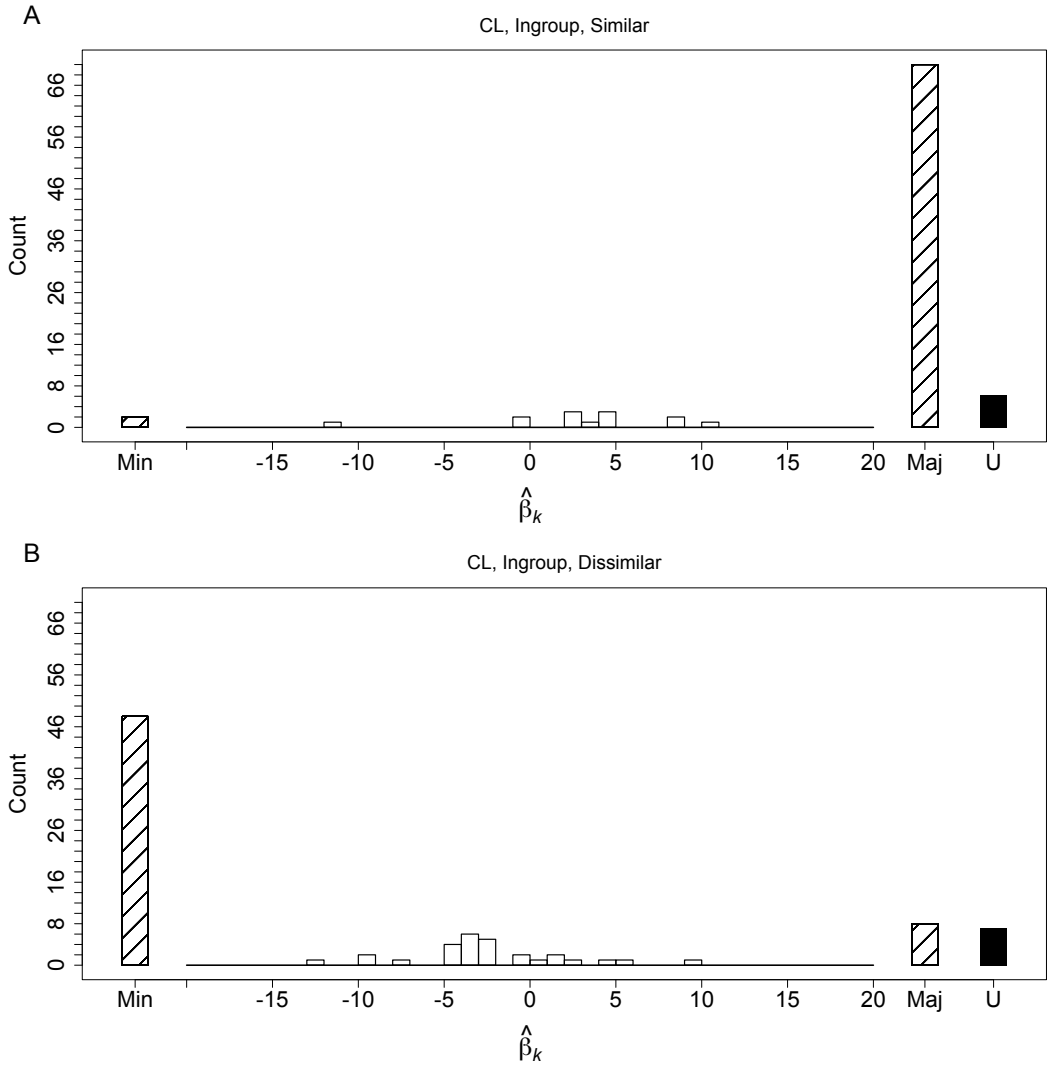
This disaggregated analysis is shown in Figures S5–S12. Because each learner contributes only a small number of blocks per treatment cell, the per-learner GLM estimates in (S4) are primarily descriptive and can be unstable (e.g., large estimates can arise under quasi-separation). For inference and for stable individual response curves, our main results rely on the hierarchical model reported above.

Figure S5: Disaggregated social learning strategies (Zurich; no cognitive load; observing ingroup).



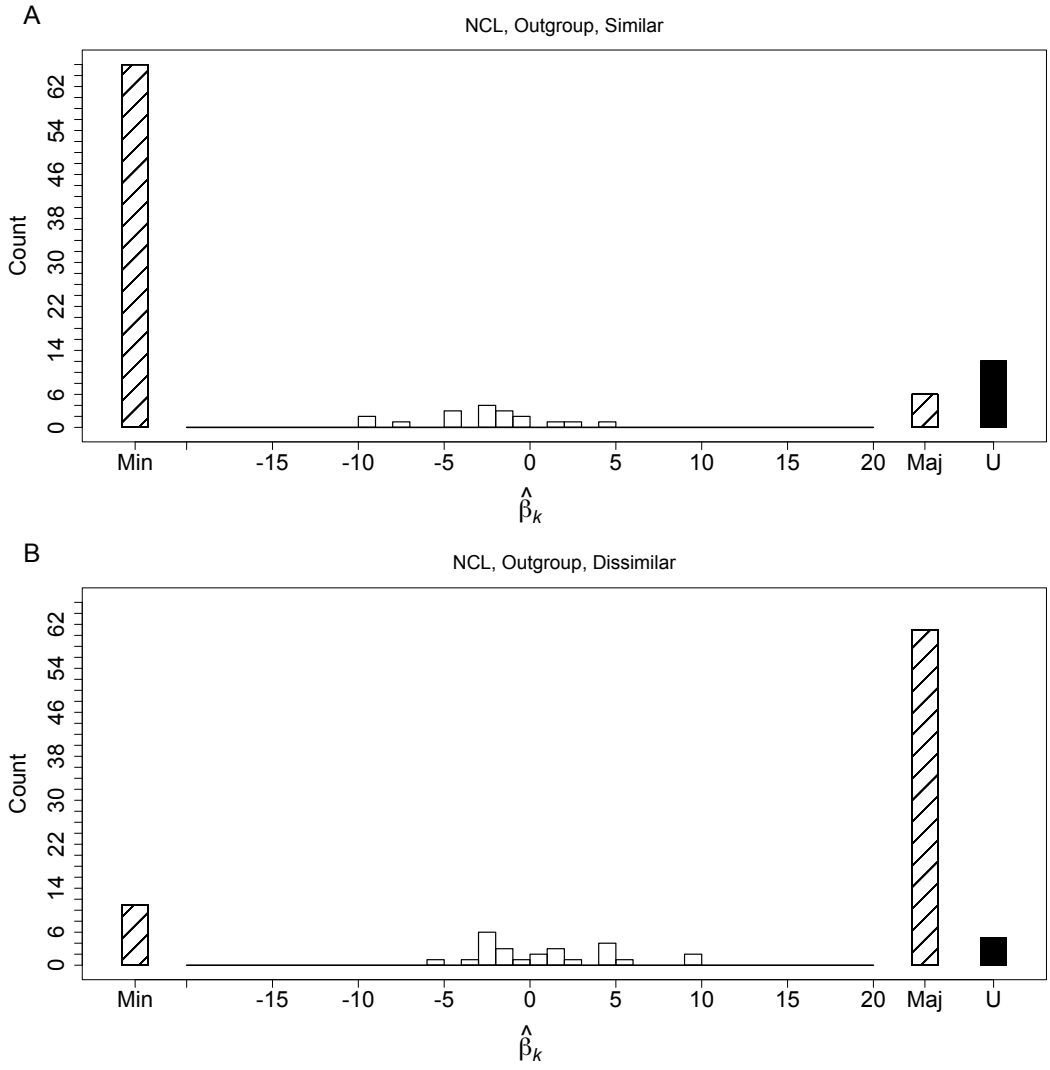
Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

Figure S6: Disaggregated social learning strategies (Zurich; cognitive load; observing ingroup).



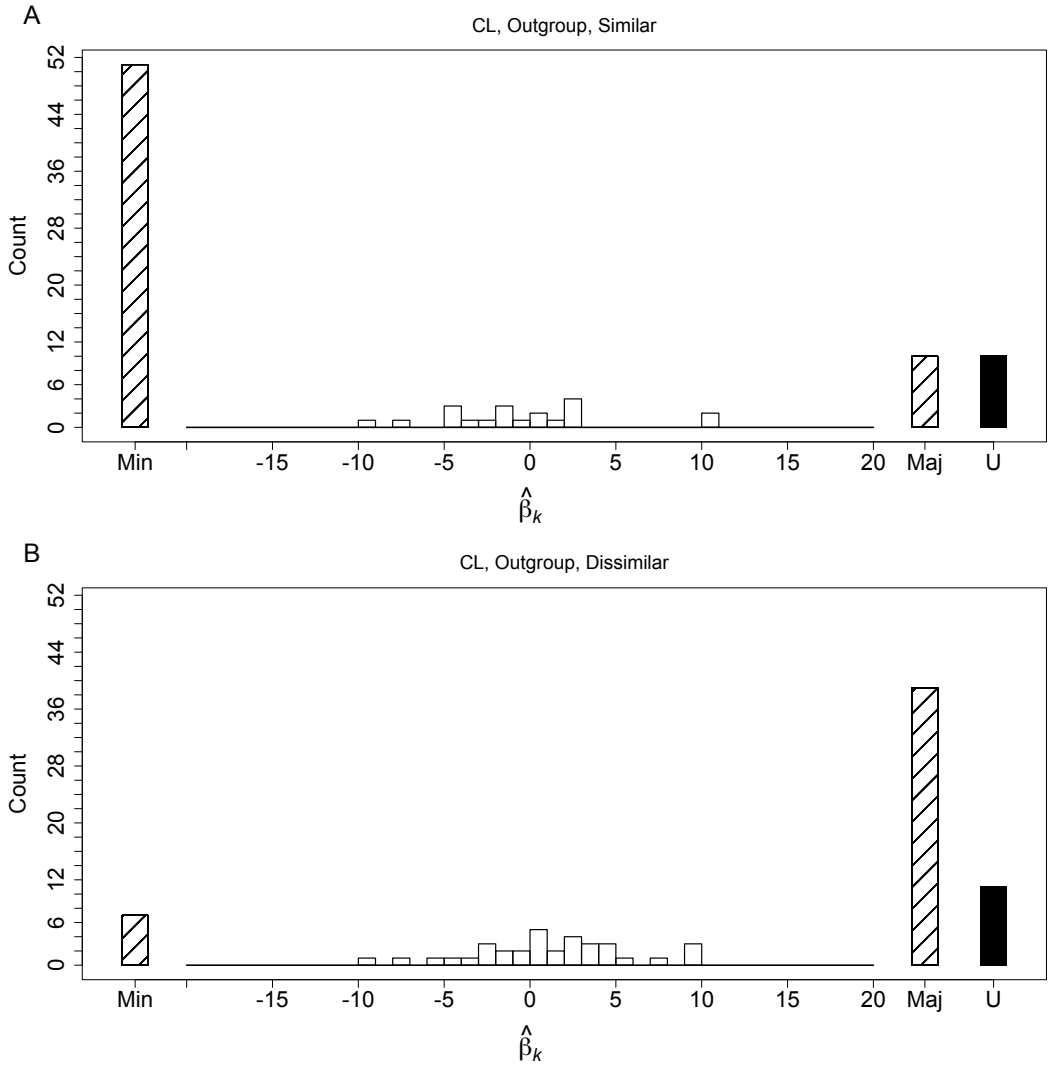
Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

Figure S7: Disaggregated social learning strategies (Zurich; no cognitive load; observing outgroup).



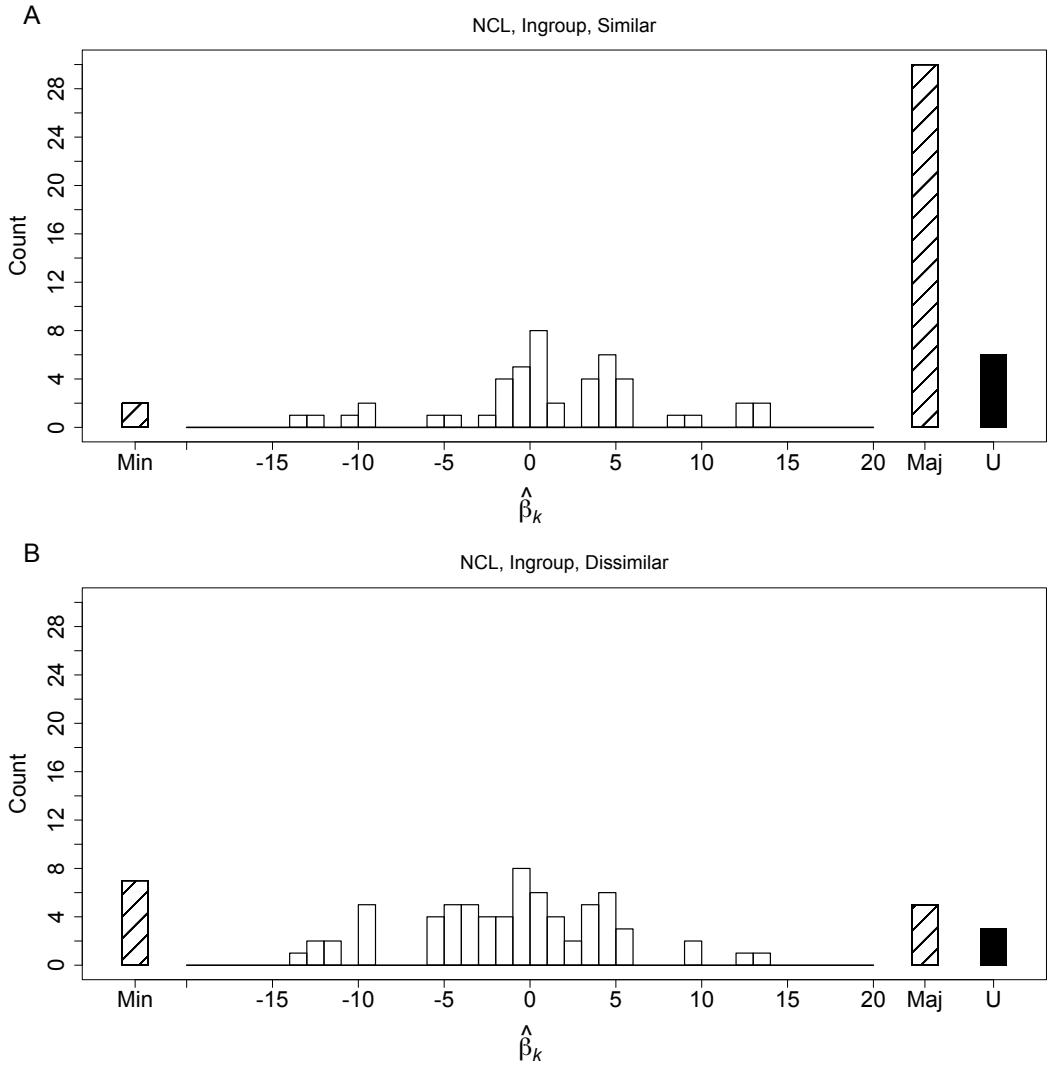
Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

Figure S8: Disaggregated social learning strategies (Zurich; cognitive load; observing outgroup).



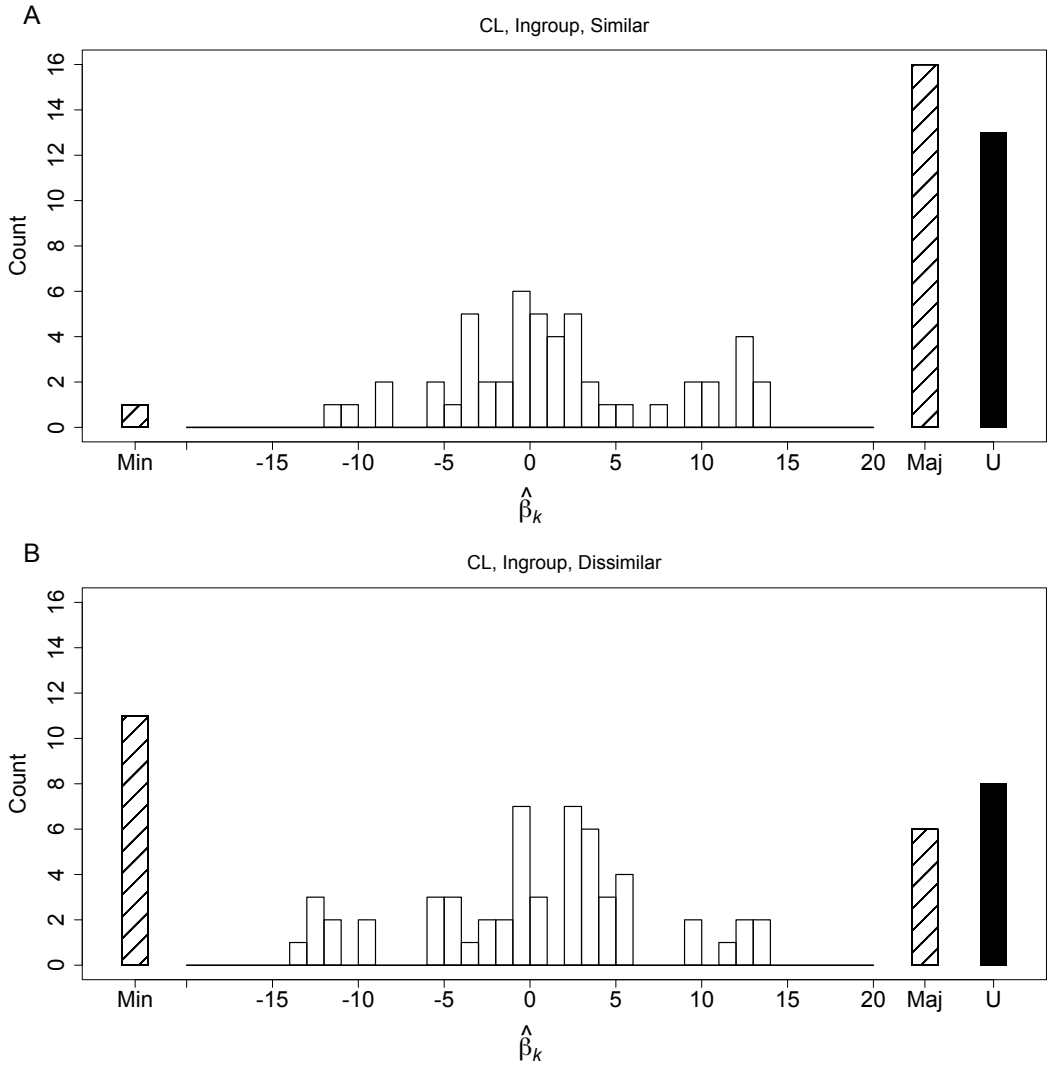
Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

Figure S9: Disaggregated social learning strategies (Nairobi; no cognitive load; observing ingroup).



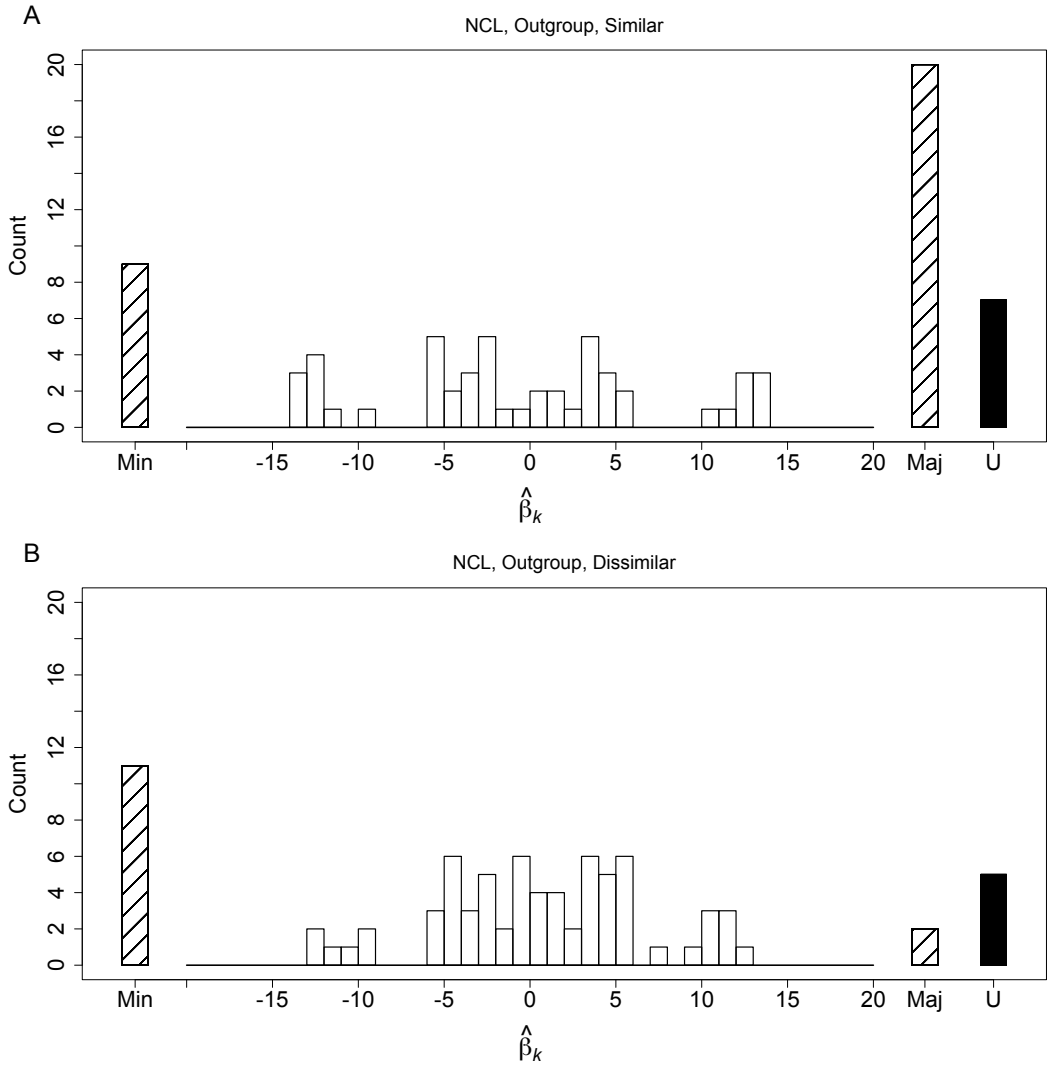
Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

Figure S10: Disaggregated social learning strategies (Nairobi; cognitive load; observing ingroup).



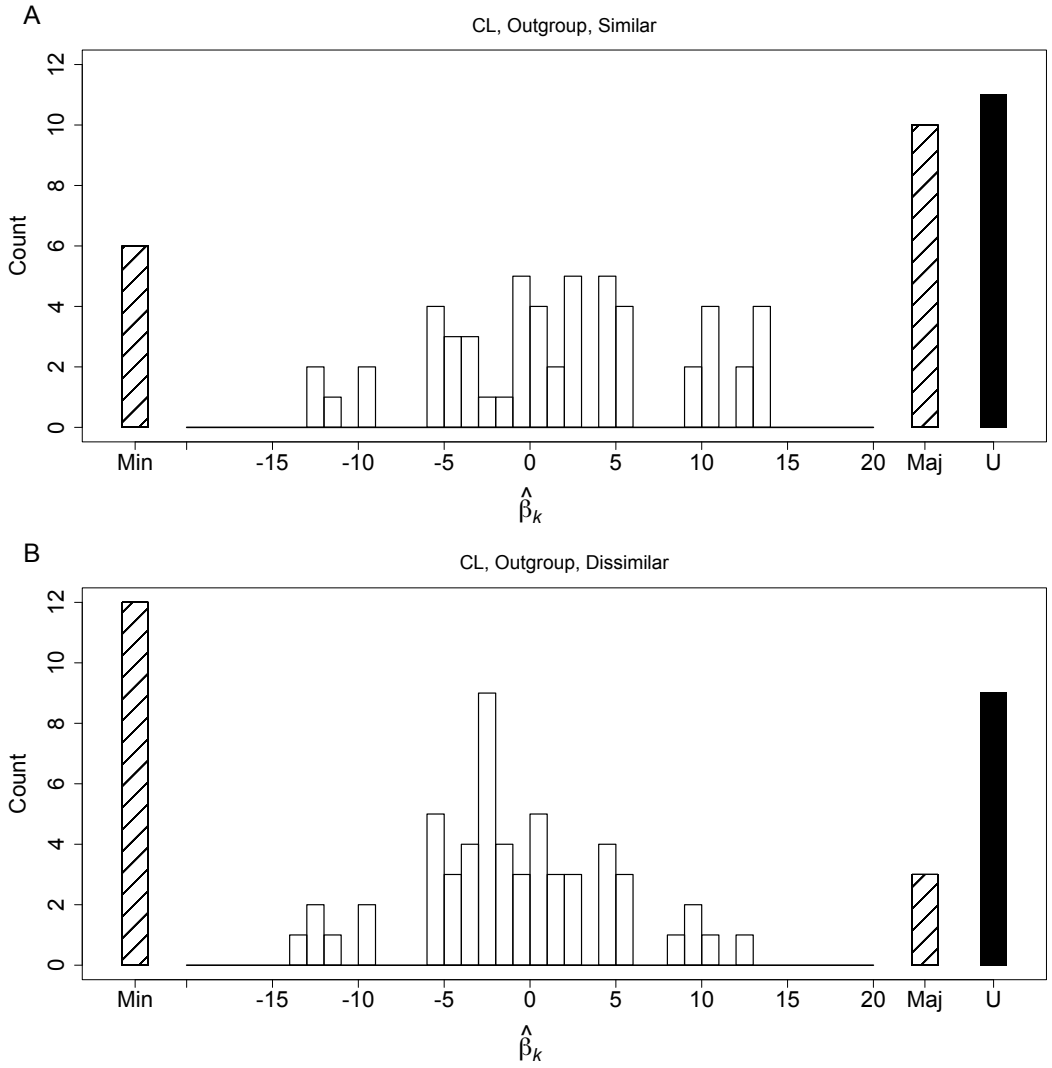
Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

Figure S11: Disaggregated social learning strategies (Nairobi; no cognitive load; observing outgroup).



Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

Figure S12: Disaggregated social learning strategies (Nairobi; cognitive load; observing outgroup).



Note. Each panel summarises learner-condition cells within the stated treatment. Counts at **Min** and **Maj** indicate the number of learner-cell pairs that perfectly follow the minority or majority rule, respectively (hatched bars). The count at **U** indicates unconditional learners who always chose the same urn in that cell (solid black bar). All remaining learner-cell pairs (**Other**) are summarised by an intercept-free logistic regression, $\text{logit}(\Pr(y_{kj} = 1)) = \beta_k x_{kj}$ with $x_{kj} = d_{kj}/5 - 0.5$ (Eq. (S4)); the histogram shows the distribution of the resulting $\hat{\beta}_k$ estimates. Grey shading within histogram bins indicates $p < 0.05$ for β_k (Wald test), while unshaded bins indicate $p \geq 0.05$. Panels **A** and **B** correspond to the Similar and Dissimilar conditions, respectively.

S3 Instructions

In the following, we present key parts of the instructions that the participants went through before the actual experiment started.

S3.1 Group and type assignment

At the beginning of the experiment, participants were assigned to a group and type. The following shows an example of how the assignment was presented to a participant.

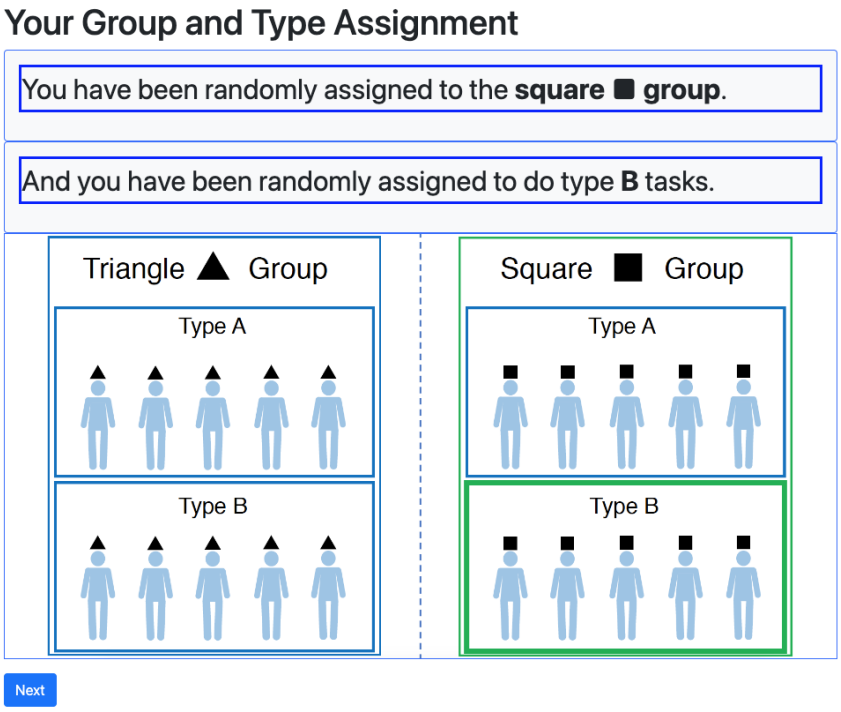


Figure S13: Example of group and type assignment.

S3.2 Experiment Task

Figure S14 shows a screenshot of the page where participants could watch a short video clip explaining the experiment task. The same animation can be watched by clicking on Figure S15 below.

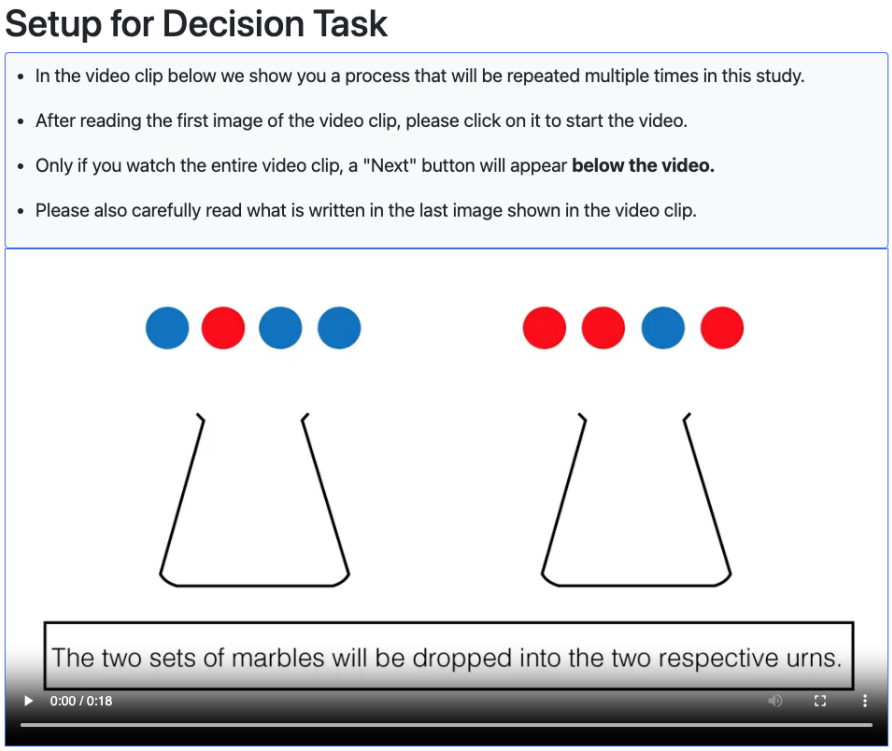


Figure S14: Screenshot showing video that explains the setup for the decision task.

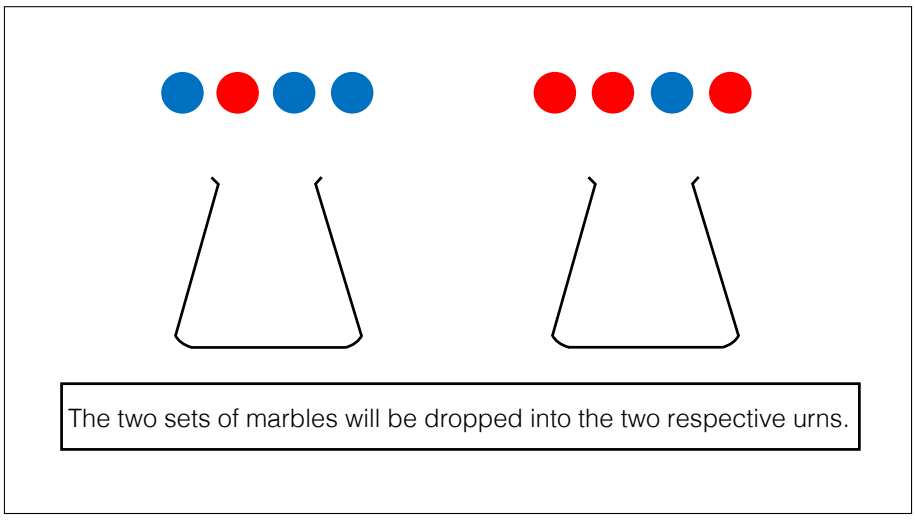


Figure S15: Animation shown to participants to explain choice task set up.

S3.3 Social information

Here we show the four different combinations of social information. The choice distributions shown in the figures are examples, and actual choice distributions reflected the choices of demonstrators during the experiment which could vary between $0 - 5$ for each urn.

S3.4 Triangle social learner

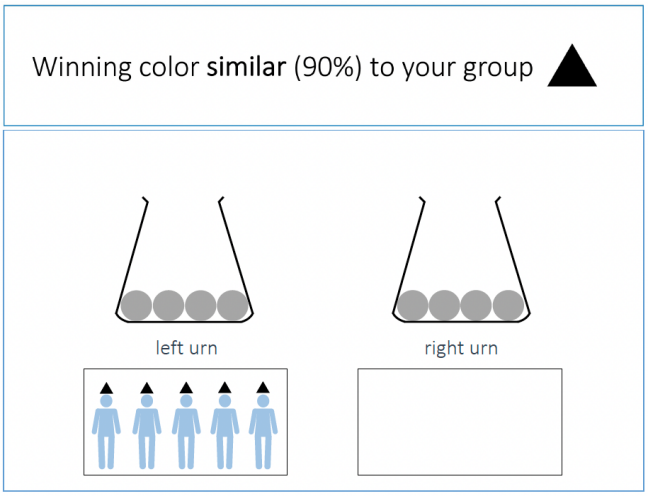


Figure S16: The social learner is similar to the ingroup and observes the ingroup.

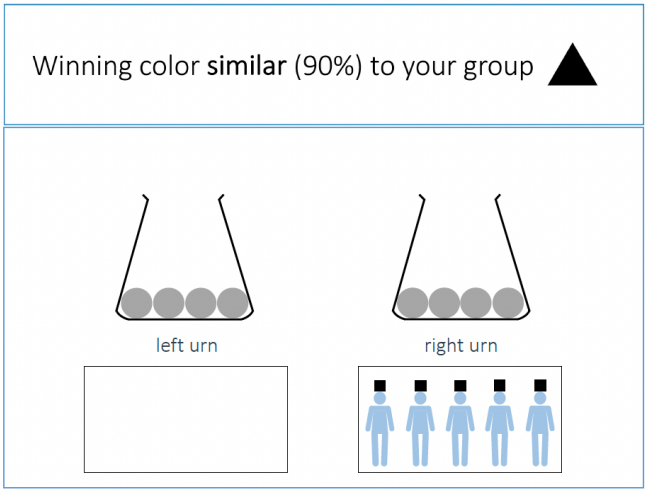


Figure S17: The social learner is similar to the ingroup but observes the outgroup.

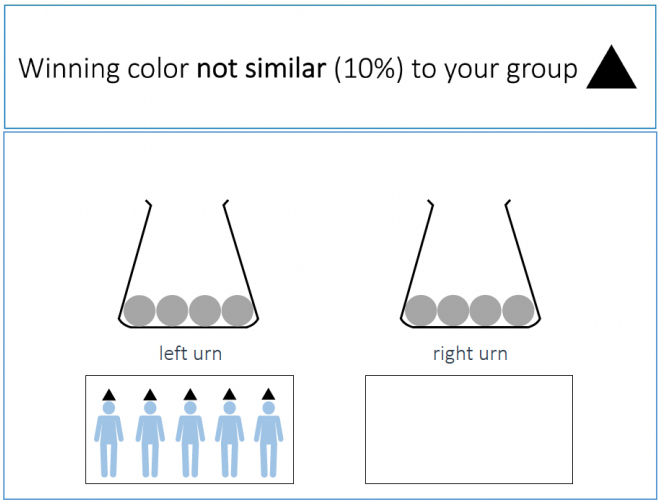


Figure S18: The social learner is dissimilar to the ingroup and observes the ingroup.

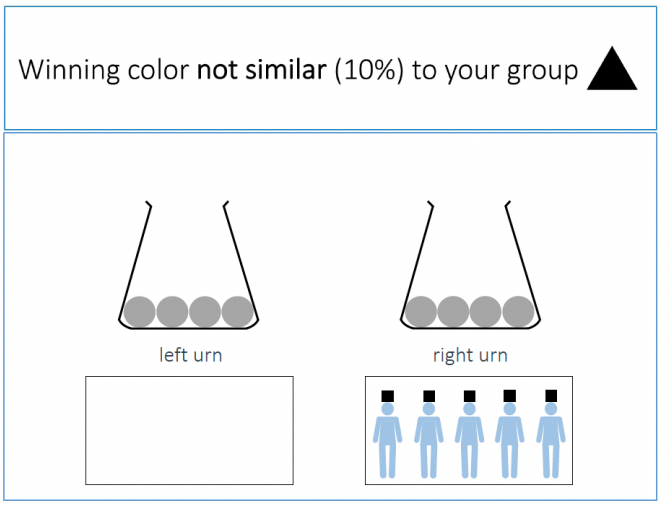


Figure S19: The social learner is dissimilar to the ingroup and observes the outgroup.

S3.5 Square social learner

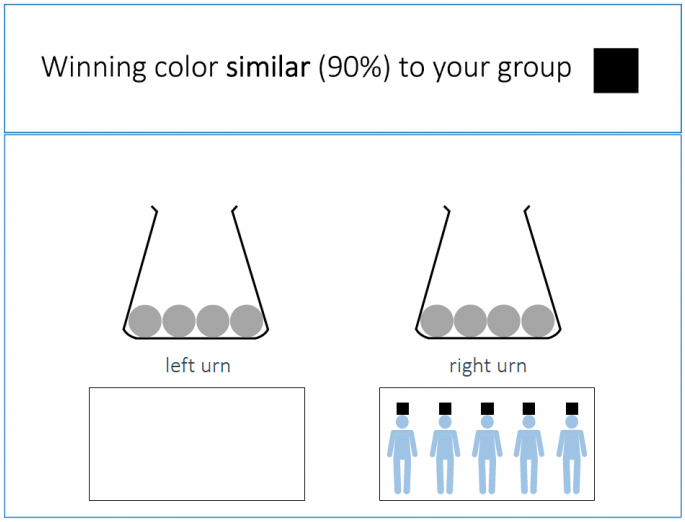


Figure S20: The social learner is similar to the ingroup and observes the ingroup.

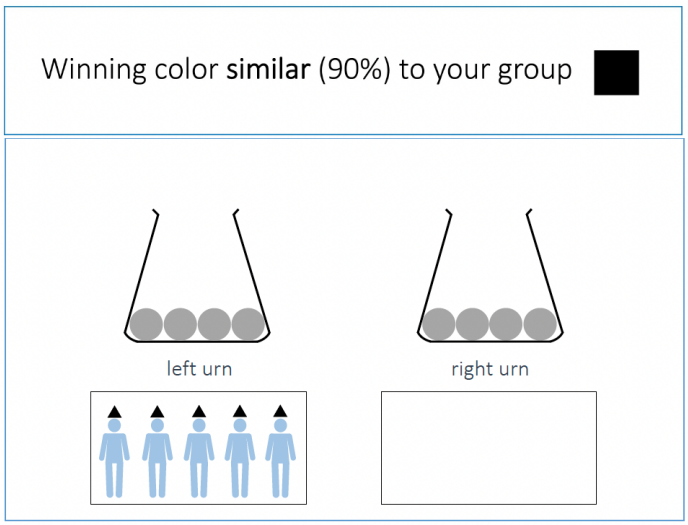


Figure S21: The social learner is similar to the ingroup but observes the outgroup.

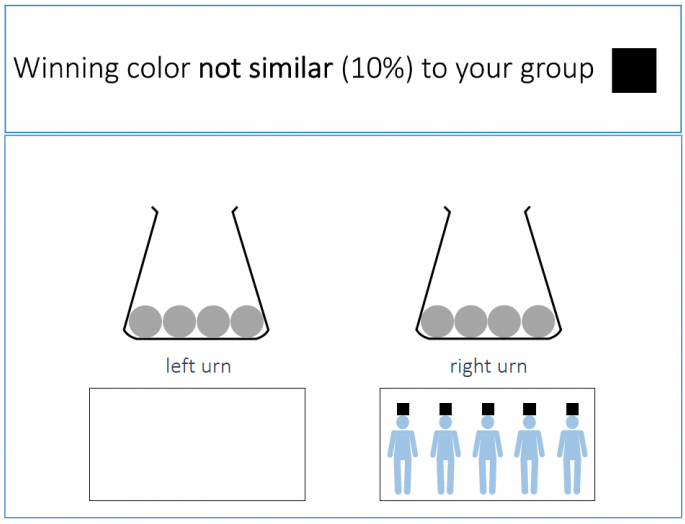


Figure S22: The social learner is dissimilar to the ingroup and observes the ingroup.

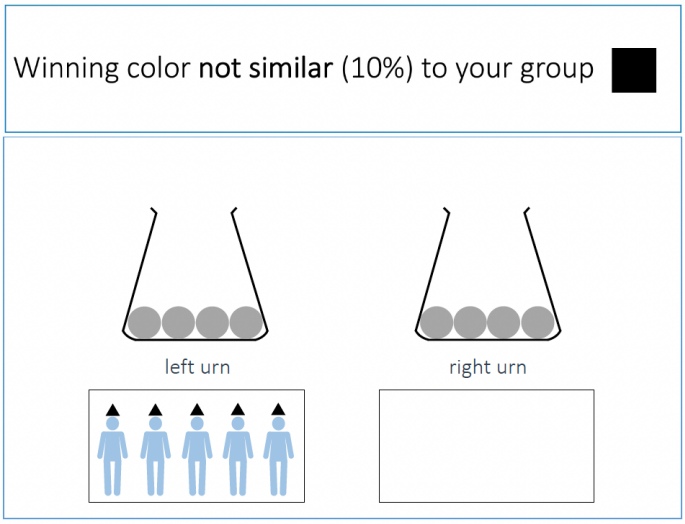


Figure S23: The social learner is dissimilar to the ingroup and observes the outgroup.

S3.6 Practice rounds

After going through the instructions that explain the choice tasks, and before actually starting the experiment, both individual learners and social learners went through practice rounds. In the following we show screenshots of the interface shown to participants during the practice rounds. The interface during the actual experiment was identical to the one shown during the practice rounds. The below screenshot shows the urn choice task. Individual learners went through four rounds of this task before social learners got to make their own urn choice. Participants could not click on the urns but had to click on the buttons "Right urn" and "Left urn", which were arranged vertically. In each round it was randomly determined which button was on top and which on the bottom. This was done to avoid any behavioral bias that could have potentially existed in terms of choosing left or right. The fact that the buttons were randomly arranged in each round was explained to the participants during the instructions.

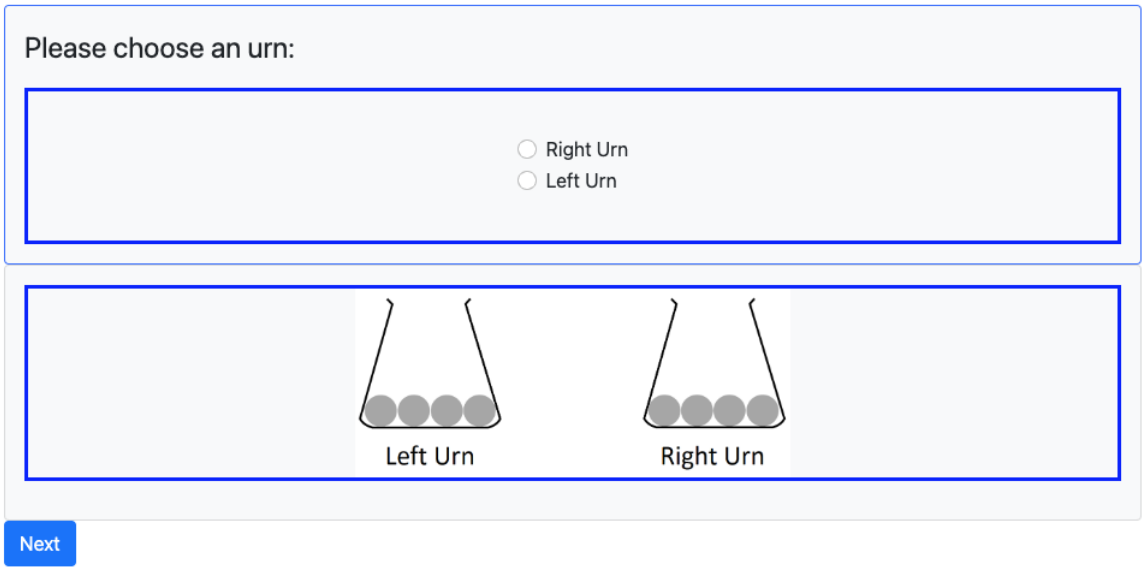


Figure S24: Urn choice

S3.6.1 Information for individual learners

Figure S25 shows an example of a private information that an individual learner saw during the first practice round after choosing an urn.

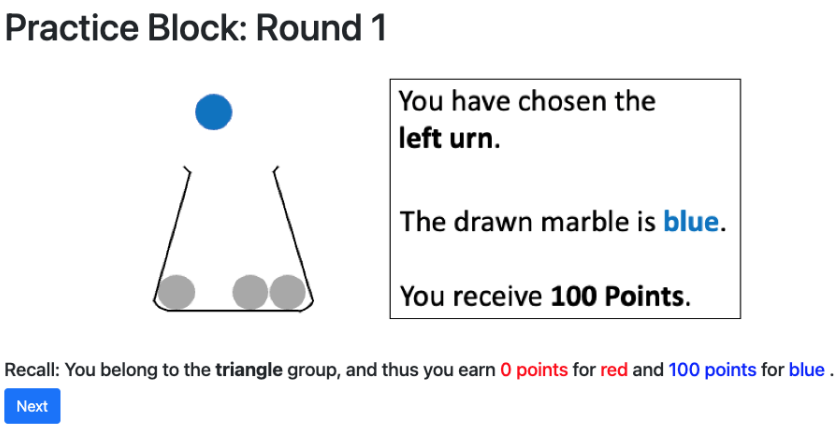


Figure S25: Urn choice

S3.6.2 Information for social learners

Figure S26 shows an example of a multi-digit number that a social learner saw before getting to see the social information. The multi-digit number was shown to social learners in both the "cognitive load" and the "no cognitive load" treatment. However, the social learners in the "no cognitive load" treatment knew that they did not have to memorize the number because they were shown the number when having to recall it.

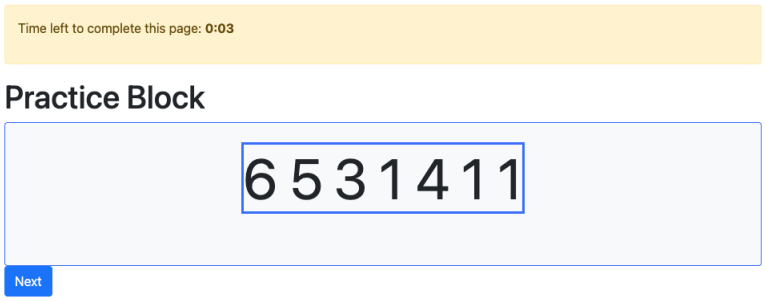


Figure S26: Example of a randomly generated multi-digit number shown to social learners.

Figure S27 shows an example of a social information that a social learner saw during the practice block before the decision task.

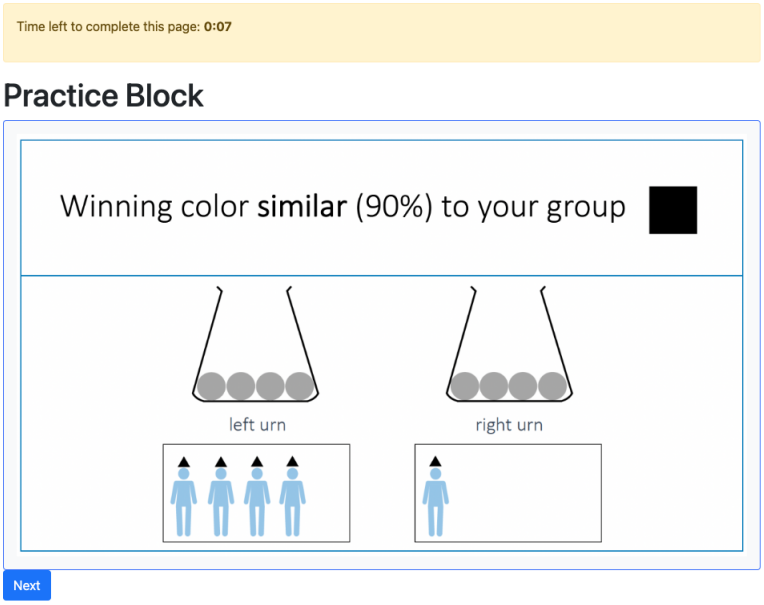


Figure S27: Example of social information shown during practice block.

A social learner in the cognitive load treatment had to recall the multi-digit number from memory and enter it. This is shown in Figure S28.

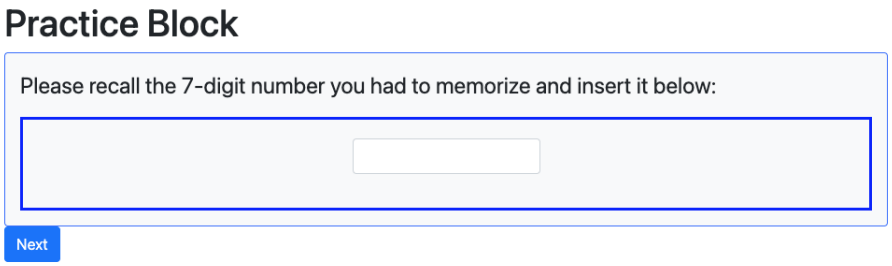


Figure S28: Screenshot showing how a social learner in the practice block had to recall and enter the memorized number.

In contrast to the social learners in the cognitive load treatment, social learner in the "no cognitive load" treatment did not have to recall the number from memory. Instead, they were shown the number. This is shown in Figure S29.

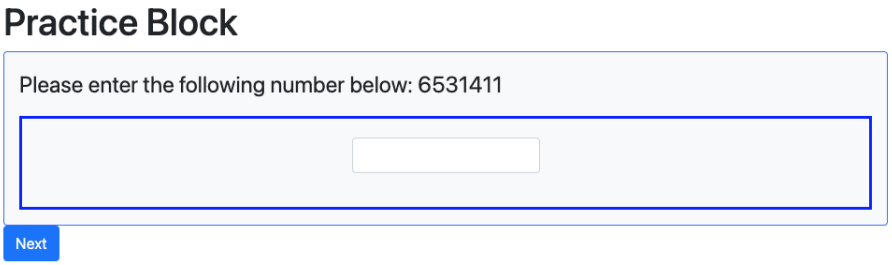


Figure S29: Screenshot showing the task of recalling the number for a social learner who was not in the cognitive load treatment.

Finally, social learners made their own urn choice. The interface is the same as shown in Figure S24 further above. When a social learner chose an urn, four marbles were randomly drawn with replacement from that chosen urn. We did this to ensure that social learners could earn the same amount of points on average as the individual learners.

S3.7 Testing the participants' understanding of the instructions

At the end of the instructions, we tested the participants' understanding of the instructions with a series of questions. Participants were only allowed to proceed if they answered every quiz question correctly. If they answered a question incorrectly, they were shown an explanation why their answer is incorrect and they had to try again (see Figure S30 below). We also informed the participants that they can raise their hand in case they did not understand something, so that we could explain it to them. The participants had to answer the following questions, some of which were multiple choice (see for example Figure S31) while for other questions they had to enter a number (see for example Figure S30):

- Question 1: What is the number of red marbles in the right urn, if there are 3 blue marbles in the left urn? (open)

- Question 2: What is the number of blue marbles in the right urn, if there is red marble in the right urn? (open)
- Question 3: Can your group membership and your type assignment change throughout the experiment? (multiple choice)
- Question 4 (individual learners): For how many rounds during one given block can you choose between the two urns and try to find out which urn has more marbles of your winning colours? (multiple choice)
- Question 4 (triangle social learners): Do you always get points for the same colour of marbles as type A participants of your own group, the triangle group? (multiple choice)
- Question 4 (square social learners): Do you always get points for the same colour of marbles as type A participants of your own group, the square group? (multiple choice)
- Question 5 (individual learners): What is the probability that a red marble will get drawn from the left urn, if there are 3 blue marbles in the right urn? (open)
- Question 5 (social learner): For which colour do type A participants of the two groups, triangle and square, earn points? (multiple choice)
- Question 6 (individual learner): What is the probability that a blue marble will get drawn from the right urn, if there are 3 red marbles in the right urn? (open)
- Question 6 (social learner): If you would only focus on either the urn choice task or memorizing the multi digit numbers, what is the worst possible outcome in terms of the amount of points you can earn? (open)
- Question 7 (social learner): If in a given block, the left urn has more marbles of the winning color for the type A participants of the triangle group, then which urn is optimal for type A participants of the square group? (multiple choice)

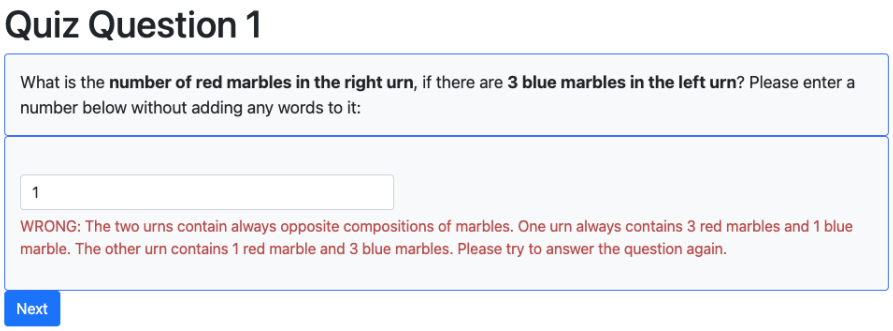


Figure S30: Example of answering a question wrongly.

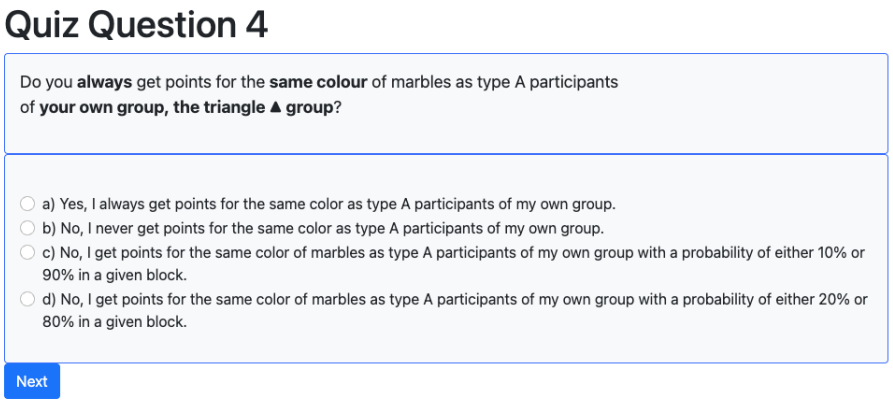


Figure S31: Example of multiple choice question.

The correct answers to the questions are the following:

- Answer 1: 3
- Answer 2: 3
- Answer 3: No
- Answer 4 (individual learners): 4
- Answer 4 (social learners): No, in some rounds the probability of getting points for the same colour of marbles as type A participants of a social learner's ingroup was 10%, and in other rounds it was 90%.
- Answer 5 (individual learner): 75%
- Answer 5 (social learner): They earn points for the opposite color.
- Answer 6 (individual learner): 25%
- Answer 6 (social learner): 0
- Answer 7 (social learner): The right urn

We measured whether each participant answered each quiz question correctly on the first attempt. The Zurich sample contains 193 social learners and the Nairobi sample contains 166 social learners. Table S15 reports the overall fraction of correct first quiz attempts for each question in Zurich and Nairobi.

Table S15: Fraction of correct first quiz attempts by social learners in Zurich and Nairobi

Quiz Question	**Zurich**	**Nairobi**
Question 1	0.927	0.777
Question 2	0.798	0.735
Question 3	0.870	0.398
Question 4	0.653	0.789
Question 5	0.845	0.627
Question 6	0.979	0.663
Question 7	0.808	0.422

To clarify whether these overall fractions reflect item difficulty or differences in field-of-study composition, we provide two complementary breakdowns by field of study. First, Tables S16 (Zurich) and S17 (Nairobi) report, within each field of study, the fraction of participants who answered each quiz question correctly on the first attempt. Thus, these entries should be interpreted as within-field percentages, while the "Total" column reports the overall site-level fraction correct.

Second, Tables S18 (Zurich) and S19 (Nairobi) decompose the overall site-level fraction correct by field of study. Each cell reports the share of the total sample that both belongs to the indicated field of study and answered the corresponding question correctly on the first attempt. Accordingly, the field-specific columns sum to the "Total" column, up to rounding.

Table S16: Fraction of correct first quiz attempts by question and field of study among Zurich social learners (within-field percentages)

Quiz Question	Humanities	Natural Sciences	Social Sciences	Total
Question 1	0.818	0.921	0.977	0.927
Question 2	0.455	0.827	0.791	0.798
Question 3	0.727	0.878	0.884	0.870
Question 4	0.727	0.669	0.581	0.653
Question 5	0.545	0.871	0.837	0.845
Question 6	0.909	0.986	0.977	0.979
Question 7	0.727	0.827	0.767	0.808

Note. Zurich sample sizes by field of study were as follows. Humanities $N = 11$, Natural Sciences $N = 139$, and Social Sciences $N = 43$.

Table S17: Fraction of correct first quiz attempts by question and field of study among Nairobi social learners (within-field percentages)

Quiz Question	Humanities	Natural Sciences	Social Sciences	Total
Question 1	0.733	0.714	0.835	0.777
Question 2	0.667	0.857	0.709	0.735
Question 3	0.378	0.452	0.380	0.398
Question 4	0.756	0.833	0.785	0.789
Question 5	0.689	0.524	0.646	0.627
Question 6	0.622	0.619	0.709	0.663
Question 7	0.356	0.381	0.481	0.422

Note. Nairobi sample sizes by field of study were as follows. Humanities $N = 45$, Natural Sciences $N = 42$, and Social Sciences $N = 79$.

In Tables S18 and S19 we additionally show, for each question, the share of the total sample with a correct first quiz attempt that belongs to a given field of study. The field-specific shares sum to the overall total.

Table S18: Decomposition of the overall fraction of correct first quiz attempts by question and field of study among Zurich social learners (shares of the total sample)

Quiz Question	Humanities	Natural Sciences	Social Sciences	Total
Question 1	0.047	0.663	0.218	0.929
Question 2	0.026	0.596	0.176	0.797
Question 3	0.042	0.632	0.197	0.870
Question 4	0.042	0.482	0.130	0.654
Question 5	0.031	0.627	0.187	0.845
Question 6	0.052	0.710	0.218	0.980
Question 7	0.042	0.596	0.171	0.809

Note. Zurich sample sizes by field of study were as follows. Humanities $N = 11$, Natural Sciences $N = 139$, and Social Sciences $N = 43$.

Table S19: Decomposition of the overall fraction of correct first quiz attempts by question and field of study among Nairobi social learners (shares of the total sample)

Quiz Question	Humanities	Natural Sciences	Social Sciences	Total
Question 1	0.199	0.181	0.398	0.777
Question 2	0.181	0.217	0.337	0.735
Question 3	0.102	0.114	0.181	0.398
Question 4	0.205	0.211	0.373	0.789
Question 5	0.187	0.133	0.307	0.627
Question 6	0.169	0.157	0.337	0.663
Question 7	0.096	0.096	0.229	0.422

Note. Nairobi sample sizes by field of study were as follows. Humanities $N = 45$, Natural Sciences $N = 42$, and Social Sciences $N = 79$.

S3.7.1 Quiz performance and task performance.

To assess whether participants who struggled with the comprehension quiz performed differently in the main task, we relate quiz performance during the instruction phase to subsequent task performance in the social-learning rounds. Quiz performance is measured as the proportion of correct first quiz attempts (`quiz_correct_prop` $\in [0, 1]$). Task performance is measured at the trial level as whether the social learner chose the objectively correct urn (`chosen_correctly` $\in \{0, 1\}$).

As a descriptive participant-level summary, we compute each participant's mean task performance and correlate it with quiz performance (c.f. our data analysis R-scripts on `https://doi.org/10.17605/OSF.IO/MZFQ5`). In Zurich, quiz performance is positively associated with task performance (OLS estimate 0.227, $p = 0.003$). In Nairobi, the corresponding association is near zero (OLS estimate -0.008, $p = 0.859$).

For inference at the trial level, we first estimate logistic regressions of `chosen_correctly` on `quiz_correct_prop` with standard errors clustered by participant. We include treatment-cell indicators (7 dummies; omitted category: Dis_Out_NCL) and also report a specification that adds our standard control for the quality of social information in the block (`propIndCorrCentered`, the centred proportion of demonstrators choosing the correct urn) and demographic controls (age, gender, and study field). We then estimate an interaction specification that allows the quiz-performance slope to differ across treatment cells. Because the coefficients in that interaction model are parametrised relative to the omitted cell, we interpret the model through the implied cell-specific quiz-performance slopes rather than through the raw interaction coefficients. Tables S20–S23 report the non-interaction specifications. Tables S24 and S25 report the cell-specific slopes from the interaction specifications.

In Zurich, quiz performance positively predicts task performance even after controlling for treatments (Table S20, $\hat{\beta} = 1.120$, $p = 0.004$) and remains positive when adding controls (Table S21, $\hat{\beta} = 0.901$, $p = 0.062$). The interaction specification provides limited evidence that this association is concentrated in a particular treatment cell: only two of the eight im-

plied cell-specific slopes are distinguishable from zero, both under cognitive load (Table S24). In Nairobi, quiz performance is not predictive of task performance in the non-interaction specifications (Tables S22 and S23; $p \geq 0.95$). The interaction specification shows that this overall null masks heterogeneous cell-specific associations: the implied quiz-performance slopes are significantly negative in Dis_Out_NCL and Sim_In_NCL, significantly positive in Sim_Out_NCL, and not distinguishable from zero in the remaining cells (Table S25). This mixed-sign pattern is difficult to reconcile with a simple account in which weaker initial comprehension uniformly depresses performance in the less direct outgroup conditions. Across both sites, the treatment coefficients remain substantively similar when quiz performance is included, so differential first-attempt quiz performance cannot on its own explain the main treatment effects.

Table S20: Likelihood of a social learner choosing optimally as a function of quiz performance (Zurich; logit model). Standard errors are clustered on the social learner (in parentheses). The omitted treatment cell is Dis_Out_NCL. Asterisks denote significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Term	Estimate	Std. Error	p-value
(Intercept)	-0.211	0.359	0.55668
quiz_correct_prop	1.120	0.394	0.0045**
Sim_In_NCL	0.601	0.178	0.0007***
Dis_In_NCL	0.409	0.147	0.0054**
Sim_Out_NCL	0.388	0.153	0.0113*
Sim_In_CL	0.550	0.186	0.0031**
Dis_In_CL	0.389	0.200	0.0510
Sim_Out_CL	-0.001	0.195	0.9947
Dis_Out_CL	-0.071	0.193	0.7139

Table S21: Likelihood of a social learner choosing optimally as a function of quiz performance (Zurich; logit model with controls). Controls include age, gender, study field, and `propIndCorrCentered` (centred proportion of demonstrators choosing the correct urn). Standard errors are clustered on the social learner. The omitted treatment cell is Dis_Out_NCL. For gender, the omitted category is *Woman*, so the reported gender coefficients compare *Man* and *Other* to *Woman*. For study field, the omitted category is *Humanities*, so the reported study-field coefficients compare *Natural Sciences* and *Social Sciences* to *Humanities*. Asterisks denote significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Term	Estimate	Std. Error	p-value
(Intercept)	-1.391	0.682	0.0414*
Age	0.004	0.018	0.8288
Man	0.104	0.130	0.4243
Other	0.140	0.415	0.7362
Natural Sciences	0.603	0.343	0.0791
Social Sciences	0.577	0.355	0.1035
quiz_correct_prop	0.901	0.484	0.0624
propIndCorrCentered	2.519	0.269	0.0000***
Sim_In_NCL	0.633	0.181	0.0005***
Dis_In_NCL	0.382	0.151	0.0111*
Sim_Out_NCL	0.364	0.157	0.0200*
Sim_In_CL	0.602	0.193	0.0018**
Dis_In_CL	0.384	0.213	0.0714
Sim_Out_CL	-0.025	0.208	0.9057
Dis_Out_CL	-0.057	0.205	0.7814

Table S22: Likelihood of a social learner choosing optimally as a function of quiz performance (Nairobi; logit model). Standard errors are clustered on the social learner (in parentheses). The omitted treatment cell is Dis_Out_NCL. Asterisks denote significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Term	Estimate	Std. Error	p-value
(Intercept)	-0.036	0.140	0.7957
quiz_correct_prop	-0.010	0.172	0.9532
Sim_In_NCL	0.351	0.138	0.0111*
Dis_In_NCL	0.028	0.133	0.8320
Sim_Out_NCL	0.141	0.145	0.3311
Sim_In_CL	0.166	0.150	0.2674
Dis_In_CL	-0.012	0.147	0.9367
Sim_Out_CL	0.038	0.145	0.7946
Dis_Out_CL	-0.190	0.144	0.1866

Table S23: Likelihood of a social learner choosing optimally as a function of quiz performance (Nairobi; logit model with controls). Controls include age, gender, study field, and `propIndCorrCentered` (centred proportion of demonstrators choosing the correct urn). Standard errors are clustered on the social learner. The omitted treatment cell is Dis_Out_NCL. For gender, the omitted category is *Woman*, so the reported gender coefficient compares *Man* to *Woman*; no separate estimate for *Other* is available in the Nairobi estimation sample. For study field, the omitted category is *Humanities*, so the reported study-field coefficients compare *Natural Sciences* and *Social Sciences* to *Humanities*. Asterisks denote significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Term	Estimate	Std. Error	p-value
(Intercept)	0.336	0.494	0.4966
age	-0.012	0.022	0.5689
gender (Man)	-0.102	0.080	0.1998
Social Sciences	-0.106	0.103	0.3056
Natural Sciences	-0.068	0.104	0.5137
quiz_correct_prop	0.004	0.173	0.9801
propIndCorrCentered	0.142	0.189	0.4517
Sim_In_NCL	0.351	0.138	0.0111*
Dis_In_NCL	0.029	0.134	0.8261
Sim_Out_NCL	0.145	0.147	0.3239
Sim_In_CL	0.156	0.150	0.2966
Dis_In_CL	-0.020	0.145	0.8874
Sim_Out_CL	0.032	0.144	0.8254
Dis_Out_CL	-0.200	0.144	0.1648

Table S24: Treatment-specific quiz-performance slopes from the interaction model (Zurich social learners; clustered SE by participant)

Treatment	Estimate	Std. Error	p-value
Dis_Out_NCL	0.910	0.967	0.347
Sim_In_NCL	0.002	0.997	0.999
Dis_In_NCL	0.865	0.844	0.305
Sim_Out_NCL	0.975	1.042	0.349
Sim_In_CL	0.695	0.727	0.339
Dis_In_CL	2.669	0.819	0.001**
Sim_Out_CL	2.363	0.819	0.004 **
Dis_Out_CL	0.034	0.818	0.967

Table S25: Treatment-specific quiz-performance slopes from the interaction model (Nairobi social learners; clustered SE by participant)

Treatment	Estimate	Std. Error	p-value
Dis_Out_NCL	-0.961	0.385	0.013*
Sim_In_NCL	-1.208	0.415	0.004**
Dis_In_NCL	0.555	0.476	0.244
Sim_Out_NCL	1.195	0.447	0.007**
Sim_In_CL	0.425	0.548	0.439
Dis_In_CL	-0.382	0.514	0.457
Sim_Out_CL	0.870	0.658	0.186
Dis_Out_CL	-0.476	0.595	0.424

S4 Questionnaire

Social learners had to fill out a questionnaire at the end of the experiment. Most of the questions concern the participants' subjective evaluation of their performance in the memorisation task. For example, we wanted to check whether participants in the "no cognitive load treatment" thought the display of the 7 digit number was less distracting than the actual memorisation task for the social learners in the "cognitive load treatment". We also wanted to know whether social learners in the "no cognitive load treatment" tried to memorize the 7 digit number, even though they did not have to do so because they were shown the number when having to enter it.

S4.1 Questions

- Question 1: When choosing between the two urns, did you spend very little time on your choice or a lot of time (1=very little time, 7=a lot of time)?
- Question 2: How difficult was it for you to decide between the two urns (1=not difficult, 7=very difficult)?
- Question 3: When progressing through the 20 blocks, did it get harder to choose between the two urns or easier or did the difficulty not change at all (1=it got easier, 4=no change, 7=it got harder)?
- Question 4 (no cognitive load treatment): How distracting was it that we showed you the multi digit number every time before showing you the choices of type A participants (1=not distracting, 7=very distracting)?
- Question 4 (cognitive load treatment): How distracting was the memorisation task (1=not distracting, 7=very distracting)?
- Question 5 (no cognitive load treatment): Did you try to memorize the number? (1=never, 7=every time)?

- Question 5 (cognitive load treatment): How many times do you think you have memorized the number correctly? (please write a number between 0-20 in the box below)
- Question 6 (cognitive load treatment): How did you allocate your cognitive resources between the memorisation task and the urn choice task (0=all focus on memorizing, 100=all focus on urn choice task)?
- Question 7 (cognitive load treatment): How difficult was it for you to recall the multi digit numbers (1=not difficult, 7=very difficult)?
- Question 8 (cognitive load treatment): When progressing through the 20 blocks, did it get harder to memorize the multi digit number or easier or did the difficulty not change at all (1=it got easier, 4=no change, 7=it got harder)?

S4.2 Results

In the following we summarize the results of the questionnaire. The answers to question 3 indicate that participants did not find it more difficult to make choices in later blocks than in earlier blocks. We perform a logit regression of optimal choice on the same independent variables as in the model presented in the main text, with the addition of an index for the 20 blocks. The effect of this index is very small and not significant. This supports the participants' perception that their performance has not decreased over time. We observe that participants in Nairobi think they spent about equal amounts of time on average on the urn choice task and the memorisation task. In contrast, participants in Zurich seem to think they spent a little bit more time on the memorisation task than on the urn choice task. Generally, this suggests that our design of setting incentives to avoid that participants only focused on one of both tasks worked. A surprising result is that the participants from both Nairobi and Zurich in the "no cognitive load" treatment answered that they tried to memorize the numbers sometimes. Specifically, the mean answer to a 7-point likert scale (1 =never, 7 =every time) in Nairobi was 4.2 and the mean answer in Zurich was 3.2. Even

though the mode answer was 1 in both subject pools, the finding that participants in the "no cognitive load" treatment tried to memorize the numbers sometimes suggests a potential explanation for our finding that the cognitive load treatment did not have an effect. This explanation does not seem plausible, however, when looking at the answers to Question 4 shown in section S4.1. Participants in the "no cognitive load" treatment answered that they felt much less distracted by the number task than the participants in the "cognitive load" treatment. The mean answer of participants to in the "no cognitive load" treatment in Zurich was 1.5, while participants under cognitive load answered 4 on average. The mean answer of participants in the "no cognitive load" treatment in Nairobi is 2.8 and the mean answer of participants under cognitive load is 4.4. Below we show the means and standard deviations of the answers to all questions shown in section S4.1.

Table S26: Answers to the questionnaire. We show the means and the standard deviations (SD) rounded to two decimal places. "NCL" stands for the "no cognitive load" treatment. "CL" stands for the cognitive load treatment.

	Zurich		**Nairobi**	
	Mean	SD	Mean	SD
Question 1	2.42	1.09	2.99	1.59
Question 2	2.60	1.52	4.05	1.84
Question 3	3.67	1.30	3.78	2.00
Question 4 (NCL)	1.56	1.13	2.82	2.01
Question 4 (CL)	4.10	2.06	4.42	1.91
Question 5 (NCL)	3.23	1.89	4.20	2.56
Question 5 (CL)	15.55	4.08	14.68	3.72
Question 6 (CL)	32.25	19.14	48.95	18.87
Question 7 (CL)	3.24	1.64	3.44	1.66
Question 8 (CL)	3.95	1.46	3.38	1.60

S5 Agent-based simulation model

We implemented an agent-based simulation model using R (R Core Team 2023) to examine population-level dynamics under different frequency-dependent social learning rules. The code for this simulation model can be found on `https://doi.org/10.17605/OSF.IO/MZFQ5` in the .zip file called "conformism.zip".

Importantly, this simulation is not intended as a model of our experiment. It provides an illustration of how different frequency response functions, and mixtures of them, can generate different population dynamics. This motivates our focus on estimating individual level heterogeneity in frequency based responses in the experiment.

The population consists of $N = 10{,}000$ individuals who repeatedly choose between two options over $T = 10$ periods. At each time step, all individuals observe the frequency $p_t \in [0, 1]$ of one option ("Alt") in the previous period and decide whether to adopt it in the next round according to their individual learning rule.

Each individual is assigned one of four learning types, each of which maps the observed frequency p_t to a probability $\Pr(\text{Alt}_{t+1}) \in [0, 1]$ of choosing Alt in the next period:

$$
\begin{aligned}
\text{Linear:} \quad & \Pr(\text{Alt}_{t+1}) = p_t \\
\text{Conformist:} \quad & \Pr(\text{Alt}_{t+1}) = \frac{S_k(p_t) - S_k(0)}{S_k(1) - S_k(0)} \\
\text{Anti-conformist:} \quad & \Pr(\text{Alt}_{t+1}) = 1 - \frac{S_k(p_t) - S_k(0)}{S_k(1) - S_k(0)} \\
\text{Non-conformist:} \quad & \Pr(\text{Alt}_{t+1}) = \begin{cases} 0.5 \cdot (2p_t)^{\gamma} & \text{if } p_t \leq 0.5 \\ 1 - 0.5 \cdot [2(1 - p_t)]^{\gamma} & \text{if } p_t > 0.5 \end{cases}
\end{aligned}
$$

Here, $S_k(p) = \frac{1}{1+e^{-k(p-0.5)}}$ is a sigmoid function centred at 0.5, with steepness parameter $k = 10$ used for conformist and anti-conformist rules, and $\gamma = 0.45$ controls the curvature of the non-conformist rule.

We considered four different population compositions (mixtures of learning types), including a homogeneous conformist population and three heterogeneous mixtures. For each composition and each of four initial conditions ($p_0 \in \{0.05, 0.45, 0.55, 0.95\}$), we ran 1,000 independent simulations. In each simulation, agents update their choices synchronously based on the population-level frequency of Alt observed in the previous period. The simulations were run with a fixed seed (`set.seed(1234)`) to ensure reproducibility.

For each mixture and initial condition, we computed the average fraction of agents choosing *Alt* at each time point. The main paper presents these average dynamics across simulations. The accompanying R script also includes code to compute 95% bootstrapped confidence intervals. These intervals were vanishingly small across all conditions, reflecting the high precision of the simulation results due to the large number of agents and repetitions. For this reason, we display only the mean trajectories in the main paper to improve visual clarity. However, all plots, including those with bootstrapped confidence bands, can be reproduced using the code provided at `https://doi.org/10.17605/OSF.IO/MZFQ5`.

References

Boyd, R. & Richerson, P. J. (1985), *Culture and the Evolutionary Process*, University of Chicago press.

Efferson, C., Lalive, R., Cacault, M. P. & Kistler, D. (2016), 'The evolution of facultative conformity based on similarity', *PLOS One* **11**(12), e0168551.

R Core Team (2023), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*