Title
Group size and cultural complexity

Permalink
https://escholarship.org/uc/item/55b8w391

Journal
Nature, 511(7507)

ISSN
0028-0836

Authors
Andersson, C
Read, D

Publication Date
2014

DOI
10.1038/nature13411

Peer reviewed
A decade ago, Henrich’s proposed group size as a driver of cultural complexity. Derex et al. now present experimental results they say support this ‘group size hypothesis’ by seemingly showing that larger groups perform better than smaller groups under imitation-based cultural evolution. Our reanalysis of their experimental data, however, shows that larger groups actually perform worse than smaller groups. Thus, contrary to their claim, their data are consistent with empirical evidence discounting the group size hypothesis for non-food producing societies. There is a Reply to this Brief Communication Arising by Derex, M. et al. Nature 511, http://dx.doi.org/10.1038/nature13412 (2014).

In their argument, Derex et al. assume that showing that larger groups produce complex artefacts more often than do smaller groups (see their Fig. 1) suffices for supporting the group size hypothesis. But such an increase can arise just through a sample size effect (for example, we expect to roll more sixes the more dice we roll). Furthermore, the increases in probabilities associated with group size, as shown in their Fig. 1, are even less than what would be expected from the sampling size effect. We believe, therefore, that their data fail to support the group size hypothesis.

We show this statistically by comparing the null hypothesis, $H_0$: the probability $P$ of an individual drawing a simple artefact on the last step of the experiment is independent of the size of the group to which the individual belongs, with both the one-sided hypothesis that expresses the group size hypothesis, $H_{1a}$; the probability $p$ of an individual drawing a simple artefact on the last step of the experiment decreases with the size of the group to which the individual belongs, and the alternative one-sided hypothesis that expresses the opposite of the group size hypothesis, namely $H_{1b}$: the probability $p$ of an individual drawing a simple artefact on the last step of the experiment increases with the size of the group to which the individual belongs.

We first test whether we can reject $H_0$. To do this, we determine the expected performance for each group of size $N = 4, 8$ or $16$, using the data for groups of size $N = 2$ to compute a maximum likelihood estimate $P = 0.73$ for $P$, the probability that an individual in a group of size $N = 2$ draws a simple artefact on the last step of the experiment. We then compute for each group of size $N > 2$ the exact probability, using $P = 0.73$, that the observed performance for a group of size $N$ deviates from the expected performance under $H_0$ as much as, or more than, what Derex et al. observe in their experiments.

The probability $P$ of getting outcomes as extreme as, or more extreme than, the experimental data are given by $P = Pr(X_2 \geq 5)Pr(X_2 \geq 2) Pr(X_5 \geq 1) = (0.25)(0.252)(0.705) = 0.0045$ (where $X_2$ has a Bin($n$, $0.73^2$) distribution, $N = 4, 8$ or $16$), as (1) there were five groups of size $N = 4$, two groups of size $N = 8$, and one group of size $N = 16$ that only drew the simple artefact on the last step; and (2) the experiments for each group size are mutually independent. As $P < 0.005$, we reject $H_0$.

Next, we note that the direction of deviation from the expected performance under $H_0$ is that of worse performance (see Fig. 1, in which fewer of the groups with $N > 2$ have at least one individual drawing a complex artefact on the last step of the experiment than would be expected based on the performance of individuals in groups of size $N = 2$). Thus, we also reject $H_{1a}$ and so we conclude that the data support $H_{1b}$, namely that a larger group size decreases the likelihood of drawing the more complex artefact on the experiment’s last step.

In sum, the experimental results of Derex et al. appear to be contrary to the group size hypothesis. This conclusion also affects negatively their other claims and, in general, underscores the importance of identifying the null hypothesis with care.

Methods
We compute, for each group size $N > 2$, $Pr(X_N \geq x_N)$, in which $X_N$ has a Bin($n$, $0.73^N$) distribution, $N = 4, 8$ or $16$, $n$ is the number of groups of size $N$ ($n = 12$ for each of the groups of size $N = 4, 8$ or $16$), and $x_N$ is the number of groups of size $N$ that only draw arrowheads on the last step of the experiment.

Figure 1 | For each group size, observed and expected number of groups drawing only the simple artefact on the last step of the experiment. Expected values are from a maximum likelihood estimate $P = 0.73$ for the probability $P$ that an individual draws a simple artefact on the last step. The expected number (EN) of groups that only draw simple artefacts is given by $EN = P^N x$, in which $N$ is the group size and $n$ is the number of groups of size $N$ ($n = 15$ for groups of size $N = 2$, otherwise $n = 12$).
In the accompanying Comment1, Andersson & Read challenge our results2 that group size influences cultural complexity. Using a dual-task computer game, our experiment demonstrated that an increasing group size prevents the loss of cultural traits (simple and complex), promotes their improvements and prevents cultural richness to disappear3. Among these various effects, Andersson & Read4 are questioning the finding that larger groups favour the persistence of the complex trait.

Andersson & Read5 claim that the expected number of groups that only exploit the simple task is more than expected if individuals in larger groups were behaving as individuals in smaller groups. Thus, they conclude that group size negatively affects the individual ability to exploit the complex task. Instead our data show the opposite pattern: individuals in larger groups tend to be more able to exploit the complex task (Fig. 1), thus supporting the group size hypothesis.

The first issue with the analysis of Andersson & Read1 is their estimation of the individual probability to exploit the simple task. Indeed, their method is expected to provide a correct estimation only if individuals behave independently of their other group members. As they claim that individuals’ ability is affected by group size, this is quite problematic.

The expected number of groups exploiting only arrowheads should be computed from the probability that a single, isolated individual exploits the simple task. This probability is not available in our data, as individuals were always part of a group. To get around this, we used a generalized linear model on individual data to estimate this probability. We obtained a probability (P = 0.80) that a single individual exploits the simple task on the last step of the experiment. Using this probability, we computed the probability of getting outcomes as extreme as, or more extreme than, our observed data for each group size: \( Pr(X_5 \geq 8) = 0.873; Pr(X_4 \geq 5) = 0.595; Pr(X_2 \geq 2) = 0.629 \) and \( Pr(X_{16} = 1) = 0.295 \). Thus, for each group size, the expected number of groups that do not exploit the complex task is not significantly different from expected. We then combine P values using Fisher’s method6, and obtain an overall P value of 0.79.

Our initial analysis7 showed that the probability of maintaining the complex trait within a group is positively affected by group size. Even if explained by sample size effect, this supports the group size hypothesis: sample size effect is expected to be the main mechanism by which group size affects cultural evolution8–10. Sample size effect should promote cultural evolution, unless the individual probabilities to exploit the task ‘died’ during the game2 were excluded from the analysis (2-player groups: \( n = 26 \); 4-player groups: \( n = 44 \); 8-player groups: \( n = 86 \); 16-player groups: \( n = 174 \)).

![Figure 1](https://example.com/figure1.png)

**Figure 1** | Frequency of individuals exploiting the complex task according to group size. Group size had a positive and non-significant effect on the individual probability to exploit the complex task (linear: \( \chi^2 = 2.53, d.f. = 1, P = 0.11 \); quadratic: \( \chi^2 = 2.15, d.f. = 1, P = 0.14 \)). Older players tend to be less likely to exploit the complex task (\( \chi^2 = 3.47, d.f. = 1, P = 0.06 \)). Players who ‘died’ during the game2 were excluded from the analysis (2-player groups: \( n = 26 \); 4-player groups: \( n = 44 \); 8-player groups: \( n = 86 \); 16-player groups: \( n = 174 \)).


doi:10.1038/nature13412

---

**Maxime Derex**, Marie-Pauline Beugin, Bernard Godelle & Michel Raymond

1. University of Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France. email: maxime.derex@gmail.com
2. CNRS, Institute of Evolutionary Sciences, CC 065, Place Eugène Bataillon, Montpellier 34095, France.