Cite as: A. Pocheville *et al.*, *Science* 10.1126/science.aaw9549 (2019).

## **Response to Comment on "Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions"**

Arnaud Pocheville<sup>1\*</sup>, Sabine Nöbel<sup>1,2</sup>, Guillaume Isabel<sup>3</sup>, Etienne Danchin<sup>1\*</sup>

<sup>1</sup>Laboratoire Évolution et Diversité Biologique (EDB UMR 5174), Université de Toulouse, CNRS, IRD, 31062 Toulouse Cedex 9, France. <sup>2</sup>Université Toulouse 1 Capitole, Institute for Advanced Study in Toulouse (IAST), Toulouse, France. <sup>3</sup>Centre de Recherches sur la Cognition Animale (CRCA), Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, UPS, 31062 Toulouse Cedex 9, France.

\*Corresponding author. Email: arnaud@pocheville.science (A.P.); etienne.danchin@univ-tlse3.fr (E.D.)

Thornquist and Crickmore claim that systematic experimental error may explain the results of Danchin and colleagues. Their claim rests on mistakes in their analyses, for which we provide corrections. We reassert that conformity in fruitflies predicts long-lasting mate-preference traditions.

Thornquist and Crickmore (1) argue that from a Bernoulli process of average sample size (63 flies) and average effect (copying probability = 0.68), Danchin et al.'s (2, 3) data reject (at threshold = 0.05) the null model more often than expected. We believe they counted the same experiment twice (their 89% is consistent with 17 experiments, although there were 16), miscalculated the average sample size (average = 61.5, not 63), grounded their argument on a nonsignificant P value (their 89% corresponds to P = 0.11), and miscalculated their P value (P = 0.124, not their 0.11, taking variation in sample size into account). Thornquist and Crickmore argue that significant P values "clustered more closely to the usual threshold for accepting a result (P =0.05) ... than expected ..." However, in addition to the mistakes above, which were transferred into their Kolmogorov-Smirnov (KS) test, they wrongly compared Danchin et al.'s *P*-value distribution to a distribution computed for an experiment of average sample size, while P-value distributions precisely depend on sample sizes (which vary in this dataset). Our own computation of 10,000 two-sided KS tests performed on P-value distributions obtained on simulated sets of 16 experiments of sample sizes equal to ours yielded ~9500 tests with a *P* value > 0.05 and an average *P* value = 0.43 (Figs. 1 to 3). Thornquist and Crickmore argue that "[nonsignificant] P values [in the GLMM] are skewed strongly toward 1, indicating far lower variance in preference than would be expected from an unbiased set of experiments." We had long noticed that variance in mate copying was low relative to an average Bernoulli process. This is, however, exactly what is expected under a mixture of Bernoulli processes (Fig. 4). Several factors may affect mate copying, that will lead to a reduced variance among treatments that are performed in parallel. This is a well-known

mathematical feature of good experimental design. The variance will be even lower if females show anticonformity when stressed and a very high capacity to learn in the best conditions.

Thornquist and Crickmore argue that transmission chain results more closely resemble the predicted mean of a Bernoulli process than expected by chance. Their argument is grounded on a nonsignificant *P* value (their *P* = 0.07, from their 10,000 simulations) which we were again unable to reproduce. Our own simulations yield *P* = 0.44; that is, 44% of simulated transmission chains were closer to the mean trajectory than the observed trajectory (1,000,000 simulations; Fig. 5). This is to be compared with *P* < 10<sup>-6</sup> for the null hypothesis: Without copying, all simulated chains were closer to the mean null trajectory than the observed trajectory. More generally, we argue that Thornquist and Crickmore discarded results where data matched the model's predictions and that this is a problematic scientific procedure (Fig. 6).

Thornquist and Crickmore speculate that male coloring may be one source of experimental bias, the quantity of powder possibly affecting demonstration and test males of the same color. When females freely chose the males, demonstration and test males did not come from the same coloring vial. When females were introduced as already copulating with a male of a given color, this male was randomly imposed by the experimenter (4). Their speculation about a "functional criterion for knowing whether enough color remains [being] mate copying [itself]" is incoherent, because coloring is finished before mate copying is assessed.

Thornquist and Crickmore note that Danchin *et al.* reported a significant effect of experimenter identity on mate copying. They also reported, importantly, that controlling

for it did not change the results.

Thornquist and Crickmore propose a model of transmission chain based on that of Danchin *et al.* Their assumption that flies remain active at low air-pressure is surprising. Their model (their figure 2B) fits neither the Dagaeff *et al.* data (4) nor the Danchin *et al.* data (2, 3), where the inflection point occurs around 1007 hPa for short demonstration protocols and air pressure does not correlate with mate copying in long demonstration protocols. We thank Thornquist and Crickmore for their comment, and we reiterate that we welcome replications of this study. We invite researchers interested in studying social learning in *Drosophila* to visit us, as others have already done to develop together new approaches for invertebrate models of animal culture.

## REFERENCES

- S. C. Thornquist, M. A. Crickmore, Comment on "Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions". *Science* 365, eaaw8012 (2019).
- E. Danchin, S. Nöbel, A. Pocheville, A.-C. Dagaeff, L. Demay, M. Alphand, S. Ranty-Roby, L. van Renssen, M. Monier, E. Gazagne, M. Allain, G. Isabel, Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions. *Science* **362**, 1025–1030 (2018). <u>doi:10.1126/science.aat1590</u> <u>Medline</u>
- É. Danchin et al., Data from "Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions". Dryad Digital Repository (2018); https://doi.org/10.5061/dryad.683870h.
- A.-C. Dagaeff, A. Pocheville, S. Nöbel, A. Loyau, G. Isabel, E. Danchin, Drosophila mate copying correlates with atmospheric pressure in a speed learning situation. *Anim. Behav.* 121, 163–174 (2016). doi:10.1016/j.anbehav.2016.08.022

## ACKNOWLEDGMENTS

We thank M. Framit, A. Jeliazkov, M. Montévil, and Au. Pocheville for discussions and help with figures, tests, and simulations.

27 February 2019; accepted 14 June 2019 Published online 11 October 2019 10.1126/science.aaw9549



Fig. 1. Percentage of nonsignificant KS tests between Danchin et al.'s 16 P values distributions simulated for an average Bernoulli process depending on mate-copying index (MCI). The red dot corresponds to the observed average MCI = 0.68 (10,000 simulations).

Fig. 2. Probability distributions for Pvalues under H<sub>0</sub> if H<sub>1</sub> is true. Dark solid line represents exact values for an average Bernoulli process similar to Thornquist and Crickmore's (N = 62, P = 0.68). Red dots show the empirical distribution, which looks offset as a result of the differential impact of some experiments on the average MCI and the distribution of P values (a heterogeneity obscured by Thornquist and Crickmore's averaging). Green and pink lines represent the probability distributions computed for two sets of 16 digital experiments originating from an average process (MCI = 0.68, green) or a mixture of Bernoulli processes (MCI between 0.5 and 1 with an average of 0.68, pink).

100

80

60

40

0

non-significant KS-tests (%)



Fig. 3. Cumulative distribution functions for the empirical set of 16 *P* values (red dots), for Thornquist and Crickmore's theoretical average process (green triangles), and for one set of 16 simulated experiments (blue squares). In the background are the exact cumulative distributions for each experiment (open circles). The variation comes from their different sample sizes and effect.



Fig. 4. Variance will be lower under a mixture of Bernoulli processes than under an average Bernoulli process. Performing treatments in parallel instantiates this mathematical fact.



**Fig. 5. Transmission-chain model in Danchin** *et al.* Color scale shows exact probabilities. Symbols represent trajectories illustrated in the original paper. Contrary to Thornquist and Crickmore (gray area in their figure 1C), the density is not uniform, but highly skewed toward the mean.



Fig. 6. Thornquist and Crickmore's method of discarding results when data fit prediction doubles the rate of wrong rejection of correct models. Asymptotically, discarding results because data match or mismatch any moment of a predicted distribution raises the rejection rate to 100% of correct models.



## Response to Comment on "Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions"

Arnaud Pocheville, Sabine Nöbel, Guillaume Isabel and Etienne Danchin

Science **366** (6462), eaaw9549. DOI: 10.1126/science.aaw9549

ARTICLE TOOLS	http://science.sciencemag.org/content/366/6462/eaaw9549
RELATED CONTENT	http://science.sciencemag.org/content/sci/366/6462/eaaw8012.full http://science.sciencemag.org/content/sci/362/6418/1025.full
REFERENCES	This article cites 2 articles, 1 of which you can access for free http://science.sciencemag.org/content/366/6462/eaaw9549#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2019, American Association for the Advancement of Science