LETTERS
https://doi.org/10.1038/s41562-019-0567-9

nature human behaviour

# Causal understanding is not necessary for the improvement of culturally evolving technology

Maxime Derex[1,2]*, Jean-François Bonnefon[3], Robert Boyd[4,5] and Alex Mesoudi[1]

**Bows and arrows, houses and kayaks are just a few examples of the highly optimized tools that humans have produced and used to colonize new environments[1,2]. Because there is much evidence that humans' cognitive abilities are unparalleled[3,4], many believe that such technologies resulted from our superior causal reasoning abilities[5–7]. However, others have stressed that the high dimensionality of human technologies makes them very difficult to understand causally[8]. Instead, they argue that optimized technologies emerge through the retention of small improvements across generations without requiring understanding of how these technologies work[1,9]. Here we show that a physical artefact becomes progressively optimized across generations of social learners in the absence of explicit causal understanding. Moreover, we find that the transmission of causal models across generations has no noticeable effect on the pace of cultural evolution. The reason is that participants do not spontaneously create multidimensional causal theories but, instead, mainly produce simplistic models related to a salient dimension. Finally, we show that the transmission of these inaccurate theories constrains learners' exploration and has downstream effects on their understanding. These results indicate that complex technologies need not result from enhanced causal reasoning but, instead, can emerge from the accumulation of improvements made across generations.**

According to the cognitive niche hypothesis, natural selection enhanced our ancestors' ability to think creatively, plan and engage in causal reasoning about their environment[5,6], and these enhancements enabled the production of more efficient technologies that powered human expansion[10,11]. Our remarkable reasoning abilities certainly contribute to the development of sophisticated technologies[12]. Yet, others have stressed that even in traditional societies human technology is often too complex to be the product of human ingenuity alone[8,9]. Constructing a well-designed bow, for example, requires solving a difficult multidimensional optimization problem[13]. The cultural niche hypothesis suggests that complex technologies such as bows result from the accumulation of many, mostly small, often poorly understood improvements made across generations linked by cultural transmission[1,9,14]. Over time, the selective retention of improvements gives rise to highly optimized solutions in the absence of explicit understanding about how these solutions work.

To test the hypothesis that the selective retention of beneficial changes over generations can produce cultural adaptations without individual understanding, we asked successive 'cultural generations' of participants (French university students) to optimize a physical system, and measured participants' understanding of how the device worked at each generation. The physical system was a wheel that travelled down a 1-m-long inclined track. The wheel had four radial spokes, and one weight could be moved along each spoke (Fig. 1a–d). Participants were organized into chains of five individuals. Each participant had five trials to minimize the time it took for the wheel to reach the end of the track. All participants (except those in the first generation) were provided with the last two configurations and associated scores of the previous participant in their chain to simulate overlapping generations. Participants were informed that their last two trials would be transmitted to the next participant in the chain, and that their reward depended both on their own performance and on the performance of the next participant in the chain. We collected data from 14 chains of 5 participants in this 'configurations' treatment.

The wheel system we used in this experiment suits our purpose for several reasons. First, it is unfamiliar (cognitive studies show that western students have poor understanding of wheel dynamics[15]), so participants cannot rely on acquired knowledge to solve the task. Second, the performance of the wheel depends solely on the laws of physics, and not on arbitrary principles that could compromise the ecological validity of our results. Finally, although the physics of the system is by no means trivial, the optimization problem is low-dimensional, which provides a conservative test of our hypotheses, compared with the many-dimensional problem of optimizing, for example, the performance of a bow[13].

The time required for the wheel to cover the track depends on just two variables: its moment of inertia and the position of its centre of mass (see Methods). This allowed us to rigorously measure participants' causal understanding of the system after they completed their five trials (Fig. 1e). Participants' understanding was evaluated by presenting them with pairs of wheels that differed in their configurations, and asking them to predict which wheel would reach the bottom of the rails first. A participant who understands the effects of varying the moment of inertia should predict that a wheel with four weights close to the axis would cover the track quicker than a wheel with four weights farther from the axis (Fig. 1a,b). Similarly, a participant who understands the effect of varying the position of the wheel's centre of mass should make correct predictions about the configurations displayed in Fig. 1c,d. The test comprised ten pairs of wheels: five in which wheels varied in their moment of inertia and five in which wheels varied in the position of their centre of mass.

The cultural niche hypothesis predicts that the speed of the wheel will increase with generations, while participants' understanding of the system will not improve over generations (pre-registered hypothesis 1).

[1]Human Behaviour and Cultural Evolution Group, Department of Biosciences, University of Exeter, Penryn, UK. [2]Laboratory for Experimental Anthropology, ETHICS (EA 7446), Catholic University of Lille, Lille, France. [3]Toulouse School of Economics (TSM Research), Centre National de la Recherche Scientifique, University of Toulouse Capitole, Toulouse, France. [4]Institute of Human Origins, Arizona State University, Tempe, AZ, USA. [5]School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA. *e-mail: maxime.derex@gmail.com

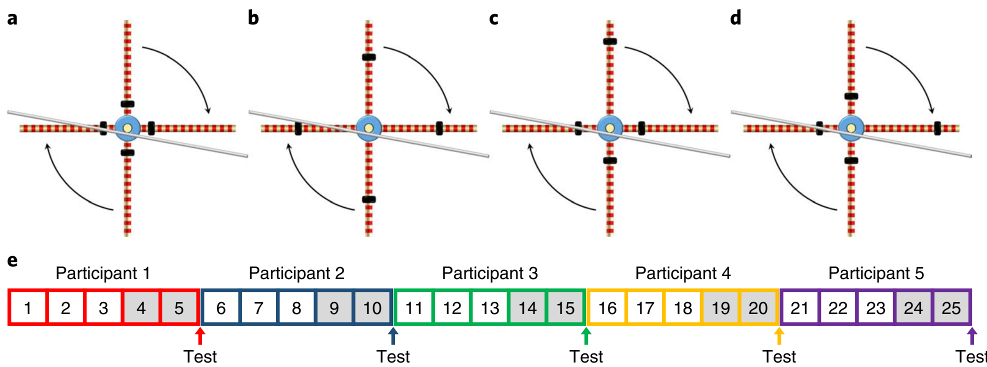


**Fig. 1 | Experimental task and design. a**, Illustration of the physical system used in the experiment. The wheel had four radial spokes, and one weight could be moved along each spoke. The time it took for the wheel to cover the track was determined by its moment of inertia and the position of its centre of mass. **a**,**b**, The moment of inertia depends on how mass is distributed around the axis. The wheel in **a** has a smaller moment of inertia and spins faster than the wheel in **b**. **c**,**d**, Asymmetrical wheels do not have their centre of mass on the axis of rotation, which can give wheels better initial acceleration. The wheel in **c** covers the distance faster than the wheel in **d** due to the higher initial position of its centre of mass. **e**, Participants were organized into chains of five individuals and had five trials each to improve their wheel. All participants (except those in the first generation) were provided with the last two configurations (shaded grey) and associated scores of the previous participant in the chain ('configurations' treatment). Participants' understanding was evaluated after they completed their five trials by asking them to predict which of two wheels would cover the distance faster (for example, **a** versus **b**, or **c** versus **d**).

The results confirmed these predictions. The average wheel speed (calculated as 1 m/descent time) increased across generations (generation 95% confidence interval (CI): 1.58 to 9.02; mean = 5.37 m h$^{-1}$; Fig. 2a) while participants' understanding did not (generation 95% CI: −0.34 to 0.25; mean = −0.04; Fig. 2b). The average wheel speed produced by first-generation participants on their last trial was 123.6 m h$^{-1}$ (95% highest posterior density interval (HPDI): 117.3 to 130.6) and their understanding score was 4.60 (95% HPDI: 3.83 to 5.53). After 5 generations, the average wheel speed increased to 145.7 m h$^{-1}$ (95% HPDI: 138.5 to 152.4) while participants' understanding remained the same (95% HPDI: 3.65 to 5.39; mean = 4.47). Given that the maximum possible speed was about 154 m h$^{-1}$, these results indicate an optimization of 71% after only 4 cultural generations. This confirms that the retention of improvements over generations produces highly optimized solutions and need not depend on the emergence of more accurate causal models.

To further investigate the relationship between cultural evolution and individual understanding, we ran a second 'configurations + theory' treatment with another 14 chains of 5 participants, in which participants could also formulate an explicit written theory about the physical system and transmit it to the next participant in the chain. The cultural transmission of explicit causal theories might affect both the optimization and the understanding of the physical system (pre-registered hypothesis 2). One possibility is that theory transmission increases both individual understanding and wheel performance. For example, participants who have a correct representation of the wheel dynamics might enhance others' performance by helping them notice the effects of varying specific parameters. However, the effects of theory transmission depend on the probability that participants generate useful theories. If participants produce incorrect theories, theory transmission would prevent individuals from noticing relevant parameters and detrimentally affect their performance. Inheriting a theory can also constrain participants' exploration behaviour (pre-registered hypothesis 3). For example, cognitive scientists have shown that children who are told the function of a toy engage in more limited exploration and are less likely to discover alternative functions than children ignorant of the toy's function[16] (see also ref. [17]). In our experiment, theory transmission might shape the exploration of parameter space and have negative downstream effects on participants' performance.

The results show that the average wheel speed increased at a similar rate in the 'configurations + theory' treatment as it did in the 'configurations' treatment (treatment: 95% CI: −10.76 to 18.13; mean = 3.52 m h$^{-1}$; generation × treatment: 95% CI: −7.07 to 2.52; mean = −2.23 m h$^{-1}$; Fig. 2a) and that participants' understanding again barely changed across generations, although participants in the very last generation had a slightly better understanding when they had inherited a theory (treatment: 95% CI: −2.54 to 0.31; mean = −1.14; generation × treatment: 95% CI: 0.03 to 0.81; mean = 0.44; Fig. 2b). Thus, these analyses do not provide substantial support for the idea that the transmission of explicit causal theories affects wheel optimization and individual understanding.

Exploratory analyses, however, reveal striking differences between treatments in participants' exploration behaviour (Supplementary Fig. 1). To investigate the effect of theory transmission, participants' theories were coded according to whether they contained information related to the moment of inertia, the position of the wheel's centre of mass, both or neither. Of the 56 participants who inherited a theory (all participants in the 'configurations + theory' treatment except first-generation participants), 15 inherited an inertia-related theory, 17 inherited a centre-of-mass-related theory, 6 inherited a full theory and 18 inherited diverse, irrelevant theories. Participants who inherited an inertia theory mainly produced compact and balanced wheels (that is, with a low moment of inertia; Fig. 3b,f). In contrast, participants who inherited a centre-of-mass theory produced unbalanced wheels with their top and right weights at extreme positions (that is, with better initial acceleration; Fig. 3c,g). The few participants who inherited a full theory produced compact and asymmetrical wheels (Fig. 3d,h). For comparison purposes, participants in the 'configurations' treatment (who did not inherit any theory) generated a greater range of wheels, although their centre of mass tended to be concentrated in the upper-right quadrant (Fig. 3a,e).

Furthermore, inherited theories strongly affected participants' understanding of the wheel system. Participants who did not inherit any theory ('configurations' treatment) scored similarly (and better than chance) on questions about inertia and questions about centre of mass (Fig. 3i). In comparison, participants who inherited an inertia- or centre-of-mass-related theory showed skewed understanding patterns. Inheriting an inertia-related theory increased their understanding of inertia, but decreased their understanding of centre of

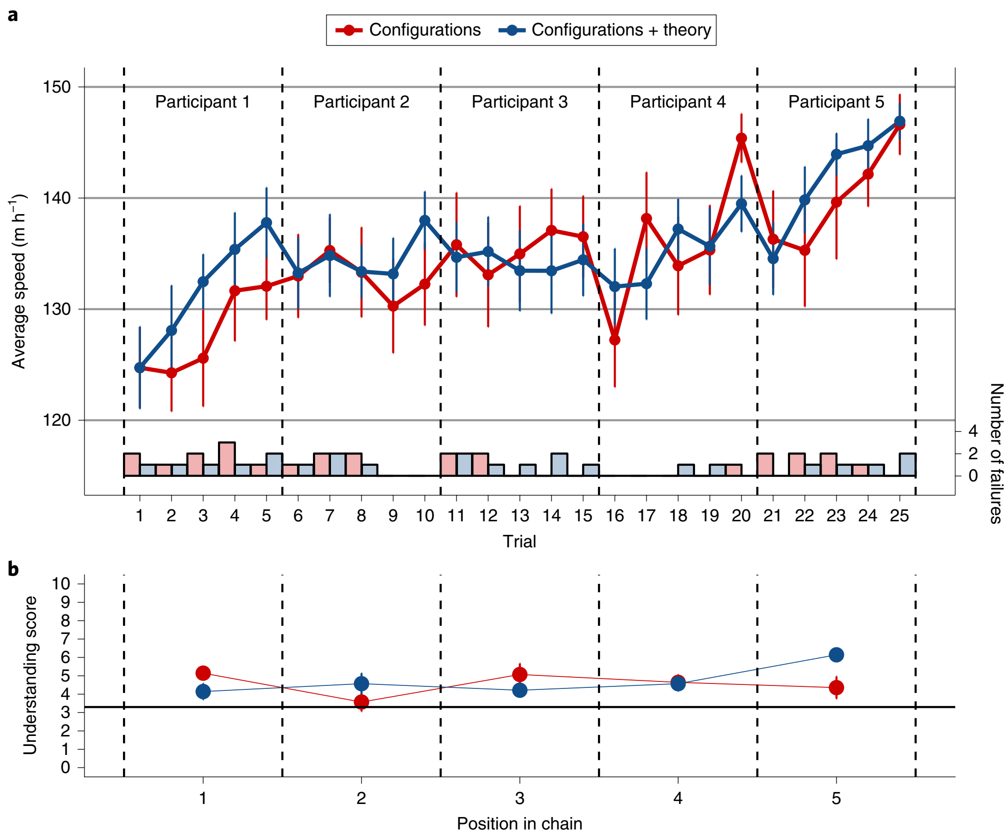


**Fig. 2 | Participants produce faster wheels across generations, but their understanding of the system does not increase. a**, Wheel performance across trials in the 'configurations' treatment (red bars and line) and 'configurations + theory' treatment (blue bars and line). Vertical bars (bottom) show the number of wheels that did not descend (that is, failures) at each trial in each treatment. Coloured lines (top) show the average speed for non-failure wheels at each trial in each treatment. **b**, Participants' understanding score across generations in each treatment. The horizontal line shows the expected score for random guessers. The error bars show s.e.m. Each treatment involved 14 chains, each containing 5 individuals.

mass. Symmetrically, inheriting a centre-of-mass-related theory increased their understanding of centre of mass, but decreased their understanding about inertia. One explanation for this pattern is that inheriting a unidimensional theory makes individuals focus on the effect of one parameter while blinding them to the effects of others. However, participants' understanding may also result from different exploration patterns. For instance, participants who received an inertia-related theory mainly produced balanced wheels (Fig. 3f), which could have prevented them from observing the effect of varying the position of the wheel's centre of mass. To test this mechanism, we grouped participants who did not inherit any theory (that is, from the 'configurations' treatment) into three categories: those who produced various types of wheels; those who only produced balanced wheels; and those who only produced unbalanced wheels. Participants who produced various types of wheels scored similarly on questions about inertia and centre of mass. However, participants who only produced balanced wheels showed better understanding of inertia than centre of mass, and participants who only produced unbalanced wheels showed better understanding of centre of mass than inertia (Supplementary Fig. 2). These results suggest that the understanding patterns observed in participants who received unidimensional theories are probably the result of the canalizing effect of theory transmission on exploration. Note that in the present case, this canalizing effect is performance-neutral: with our two-dimensional problem, better understanding of one dimension and worse understanding of another dimension simply compensate each other. However, for a many-dimensional problem, better understanding of one dimension is unlikely to compensate for worse understanding of all of the others.

As predicted by the cultural niche hypothesis[9], our experiment shows that highly optimized technologies can emerge from the accumulation of many improvements made across generations linked by cultural transmission, without the need for an accurate causal understanding of the system. Most participants actually produced incorrect or incomplete theories despite the relative simplicity of the physical system. These results are consistent with the view that individuals do not spontaneously create multidimensional representations of object motion[15]. Instead they mainly produce unidimensional models related to a specifically salient dimension[18]. Although examples of evidence of individuals' erroneous theories of motion are sometimes considered as experimental artefacts resulting from impoverished stimuli (such as using pictures to describe dynamical events[19]), our results show that incomplete representations commonly emerge even when individuals directly observe and modify an actual physical object. As a consequence, the transmission of explicit theories across generations did not help participants produce more efficient wheels: inheriting a theory mostly constrained participants' exploration, and prevented them from noticing the effects of relevant variables outside the theory they received.

It is worth noting that despite exhibiting poor understanding of the experimental physical system, participants did not randomly explore the parameter space. For example, in both treatments, wheels were much more likely to have their centre of mass at the centre of the wheel, or in the upper-right quadrant. This indicates that participants had appropriate intuitions about how to maximize acceleration, and sampled the parameter space fairly efficiently in that regard. Our ability to restrict exploration to potentially useful portions of the design space

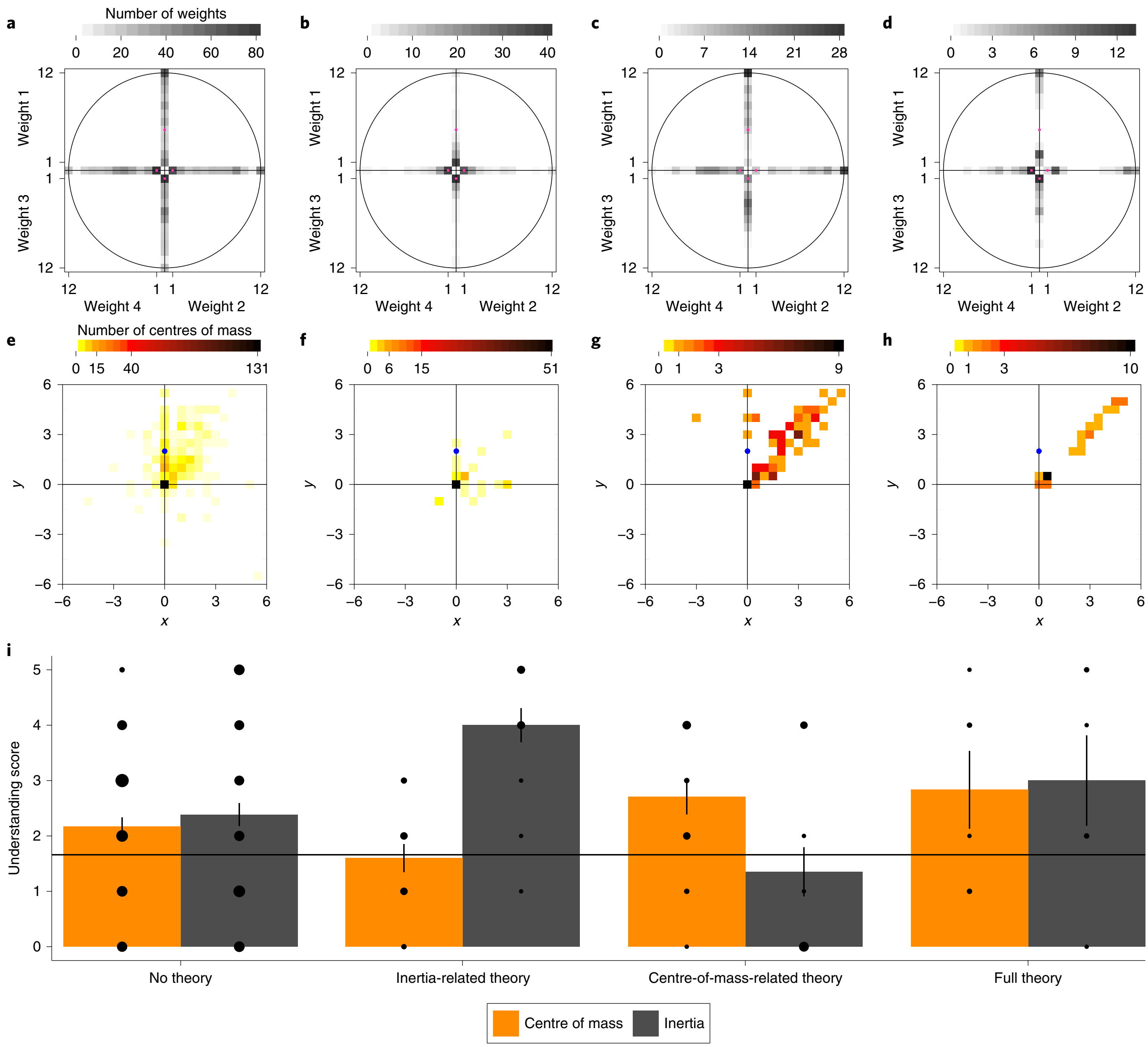


**Fig. 3 | Inheriting a theory affects both participants' exploration and understanding. a–h**, Heat maps illustrating the most frequent weight positions along each spoke (**a**–**d**) and the most frequent positions of the wheels' centres of mass (**e**–**h**) for no theory (**a** and **e**), inertia-related theory (**b** and **f**), centre-of-mass-related theory (**c** and **g**) and full theory (**d** and **h**). In **e**–**h**, the blue dot shows the optimum centre-of-mass position. Participants who did not inherit any theory ($n = 70$) sampled various positions along each spoke (**a**) and their wheels' centres of mass were concentrated in the upper-right quadrant (**e**). Participants who inherited an inertia theory ($n = 15$) mainly produced compact and balanced wheels (**b** and **f**). Participants who inherited a centre-of-mass theory ($n = 17$) produced unbalanced wheels with their top and right weights at extreme positions (**c** and **g**). The few participants who inherited a full theory ($n = 6$) produced compact and asymmetrical wheels (**d** and **h**). Values 1–12 in **a**–**d** describe the postions of weights 1–4.Values −6 to +6 in **e**–**h** describe the *x*- and *y*-coordinates of the centre of mass. **i**, Inheriting an inertia theory reduces understanding about centre of mass and increases understanding about inertia, while inheriting a centre-of-mass theory increases understanding about centre of mass and reduces understanding about inertia. The solid horizontal line shows the expected score for random guessers. Error bars show s.e.m. Black dots represent raw data, with dot size representing the number of observations.

certainly accelerated cultural evolution in our experiment. A greater focus on the determinants of biased exploration would be a fruitful area for further work. Here, we cannot tell whether participants' intuitions resulted from an implicit physics engine, past experience with analogous objects or Western formal education (although physics or engineering background had no effect on participants' understanding scores; Supplementary Fig. 3). Future cross-cultural work involving non-WEIRD participants (that is, not coming from Western, educated, industrialized, rich, and democratic societies) should tell us whether this selective exploration is culturally constructed or shared across populations[20]. In any case, our experiment indicates that one should be cautious when interpreting complex archaeological materials as evidence for sophisticated cognitive abilities (such as reasoning, problem solving or planning), since these abilities are not the sole driver of technological sophistication[1,9]. Understanding technological change demands a focus on individual cognition[5,6], but also requires us to give attention to factors affecting the pace of cultural accumulation, such as cultural transmission dynamics and demography[21–29].

## Methods

**Participants.** In total, 140 participants took part in the study (70 women and 70 men). Participants were randomly selected from a database managed by the Catholic University of Lille and recruited by email from various universities in Lille, France. The subjects ranged in age from 18–38 years (mean = 20.5 years; s.d. = 3.4 years). Participants received €3 for participating and an additional amount ranging from €0–26 depending on their own performance and the performance of the next participant in their chain (see below).

**Ethics statement.** The study was carried out in accordance with the ethical standards of the 1964 Declaration of Helsinki and the guidelines of the British Psychological Society's Code of Human Research Ethics. All methods were approved by the University of Exeter Biosciences Research Ethics Committee (2018/2310) and the Catholic University of Lille Research Ethics Committee (2018-01-31-E). All participants provided written, informed consent before taking part in the experiment.

**Experimental apparatus.** *Dynamics of the wheel.* The performance of the wheel depends on two variables: its moment of inertia and the position of its centre of mass. The wheel's moment of inertia depends on how mass is distributed around its axis of rotation. Wheels with a smaller moment of inertia (that is, wheels that have their weights closer to the axis) require less torque to increase angular momentum and spin faster (movies are available at https://osf.io/afwmr/). Asymmetrical wheels do not have their centre of mass on the axis of rotation, which can give wheels better initial acceleration. When the centre of mass of the wheel is in the wheel's upper-right quadrant, more energy is converted into angular kinetic energy so that the wheel will benefit from higher increases in angular momentum. Note that the same would occur with a centre of mass in the upper left quadrant. There, the wheel would rotate in the wrong direction and would go up on the rails (the kinetic energy would be converted into potential energy).

In our experiment, both the wheel's moment of inertia and the position of its centre of mass had to be taken into account to reach the best performance. A higher centre of mass can produce better acceleration but it will increase the wheel's moment of inertia and so there was a tradeoff between maximizing acceleration and minimizing inertia (Supplementary Fig. 4). Acceleration could be optimized in two different ways. One is keeping all weights close to the axis except the top one. The other is moving both the top and right weights away from the axis. This latter strategy can give the wheel better initial acceleration because the right weight has more leverage than the top weight to set the wheel in motion at its initial position (the top weight initially applies a vertical force on the axis, which does not affect the wheel's angular momentum). However, the right weight will only fall from half the height of the top weight (assuming both weights are equally far from the axis), so less energy will eventually be converted into kinetic energy.

*Building of the wheel.* The wheel was built around a tube clamp designed to form a 90° angle between a 28 mm tube (which passed through the clamp) and 4 other 28 mm tubes (with 90° angles between contiguous tubes; Fig. 1 and Supplementary Fig. 5a). The axis of the wheel was composed of a 10.5-cm-long bored-through wooden pole and an 8 mm threaded steel rod in its centre. The threaded steel rod protruded approximately 4 cm past the end of the wooden pole at each side and was covered with pieces of 3 cm rubber tube to prevent the wheel from sliding on the rails. Flat washers were positioned on either side of the pieces of rubber tube to guide the wheel along the rails and limit potential friction. Two nuts held the materials in position. Two 500 g weight plates were positioned along the axis of the wheel (one on each side of the clamp) to reduce the wheel's moment of inertia and limit the occurrence of motionless or back-spinning configurations. Two barbell clamp collar clips were used to lock the weight plates in position (Supplementary Fig. 5b). Four 28 mm wooden poles formed the spokes of the wheel and were 41 cm long from the centre of the wheel. Pieces of red tape were positioned every 28 mm along the spokes to signal 12 discrete weights' potential positions (the closest position to the axis was 6.5 cm from the centre of the wheel). Four barbell clamp collar clips were used as weights. Each was weighted with flat washers, screws and nuts (Supplementary Fig. 5c). The weight of a collar clip was about 100 g.

*Building of the rails.* Rails were built from 2-long plated steel slotted angles (20 mm wide). A steel and aluminium structure held the rails at an incline of 14°. Two push-button switches (made from computer mice) were located 92 cm apart on the rails and connected to a computer program (Supplementary Fig. 5b). Two arrows indicated the positions of the switches (starting/ending points; Supplementary Fig. 5a). A mechanical lever maintained the wheel motionless, with two of its spokes parallel to the ground at its starting position.

**Procedure.** The experiment took place in an experimental room at the Laboratory for Experimental Anthropology at the Catholic University of Lille. For each session (around 20 min long), a single individual was recruited and sat at a computer that was placed parallel to and at 2 from the experimental apparatus. Participants were randomly assigned to one condition of the experiment and one sex-segregated chain. Before starting the experiment, participants were asked to complete a consent form and were asked their age. At the end of the experiment, participants indicated whether they had an academic background in physics or engineering. Participants entered and left the room by two different doors to prevent any form of direct interaction between them. Participants came back to the laboratory a few days after the experiment to receive payment (once their final payoff was known; see below).

**Experimental design.** *Building phase.* Each participant had 5 trials to minimize the time it took the wheel to cover about 1 on an inclined track. Weights could be placed on one of 12 discrete positions along 4 spokes, which created a space of 20,736 unique configurations. Participants chose their configurations through a computer program using four sliders (Supplementary Fig. 6). Once the configuration was confirmed by the participant, the experimenter positioned the weights on the physical wheel accordingly (the computer screen was projected onto a wall to the right of the participant to allow the experimenter to see the chosen configuration without interacting with the participant; Supplementary Fig. 5a). The wheel was then positioned on the rails and held motionless by a mechanical system before being released. Once released, the time it took the wheel to descend the track was automatically recorded by the computer program. The wheel's average speed and associated payoff was then automatically displayed on the participant's screen. Participants could consult their two last configurations between any trials. They had as much time as they needed to consult these configurations and choose their next one. After three trials, participants were reminded that their last two configurations will be transmitted to the next participant in the chain. After five trials, the program automatically switched to the test phase.

*Testing phase.* After completing the task, participants were told that they would be presented with pairs of wheels and that they must guess which one of two wheels would cover the rails faster. They were also told that one of their answers will be randomly selected at the end of the test and that €5 will be added to their gain if that answer is correct. For each pair, participants could submit 3 possible answers: 'wheel 1', 'wheel 2' or 'no difference'. Participants could take as much time as needed before submitting their answer. Once an answer was submitted, another pair of wheels was displayed until participants compared ten pairs of wheels. In five pairs, wheels varied in their moment of inertia. In the other five, wheels varied in the position of their centre of mass (Supplementary Fig. 7). Participants were not told whether their guesses were correct. All participants were exposed to the same ten pairs of wheels in the same order.

*Experimental treatments.* Two treatments were run. In each treatment, participants were part of 14 chains, each containing 5 individuals (exclusively males or exclusively females). No statistical methods were used to pre-determine sample sizes, but our sample sizes (that is, number of independent chains) were larger than those reported in previous publications[30,31]. All participants except those in the first generation were provided with social information. In the 'configurations' treatment ($n = 70$), the last two configurations and associated scores of the previous participant in the chain were provided to the next participant in the chain. In the 'configurations + theory' treatment ($n = 70$), participants additionally received the previous participant's theory about the physical system. Participants were asked to write their theory after the test phase was completed. Participants could not transmit information about the performance of a specific configuration, to prevent individuals from extending the number of transmitted configurations compared with the 'configurations' treatment. Theories had to be less than 340 characters long and always started with 'The wheel covers the distance faster when…'. Social information was available all along the building phase and could be consulted between any trials in both treatments. The organization of experimental conditions was randomized. Data collection was not performed blind to the conditions of the experiments.

*Pre-experiment information.* The instructions could be read on a computer screen. They stated that the participants' task was to position four weights on a wheel to minimize the time it takes the wheel to cover an inclined track (Supplementary Methods). Participants were informed that they had five trials to do this and that their payoff would be determined by the performance of each of their wheels. Participants were told that they were part of a chain, so that the task was a collective one (despite being alone in the experimental room). They were informed that their last two configurations would be transmitted to the next participant in the chain, and all participants except those in the first generation were also told that they were going to be provided with the last two configurations of the previous participant in the chain. In the 'configurations + theory' treatment, participants were also informed that they could write/receive a theory. Finally, participants were told that their final gain would be determined by their own performance and the performance of the next participant in the chain. Participants did not know the length of the chain, nor the speed of the best possible wheel.

*Participants' payoff.* The following equation determined the payoff of each wheel:

$$[1-((\text{MaxSpeed} - \text{RecordedSpeed})/(\text{MaxSpeed} - \text{MinSpeed}))] \times 3 + \text{Bonus}$$

where MaxSpeed = 160 and MinSpeed = 96. RecordedSpeed was the recorded average speed of the wheel. Bonus took the value 0.2 for wheels that descended, and 0 otherwise.

Participants' final payoff corresponded to the sum of the payoff of each of their wheels plus the payoff of the next participant's first two wheels plus €5 if they correctly answered the randomly selected test. Final participants in chains had their last two payoffs doubled (although they were not aware of this as they did not know that the chain was about to end).

*Theory coding.* Five individuals blind to the research question were explained the dynamics of the wheel (that is, the respective role of inertia and centre of mass in the performance of the wheel) and were asked to code participants' theories according to whether they contain accurate information related to moment of inertia and/or centre of mass. A theory contained information related to the moment of inertia when it said that the wheel goes faster when its weights are close to the axis (for example, 'The wheel covers the track faster when its weights are balanced and close to the axis.'). A theory contained information related to centre of mass when it said that the wheel goes faster when its centre of mass is in the upper-right quadrant (for example, 'The wheel covers the track faster when its top and right weights are farther from the axis than its bottom and left weights.'). A few theories contained information about both principles (for example, 'The wheel covers the track faster when its weights are balanced and close to the axis. Furthermore, the wheel has a better initial acceleration when the top and right weights are slightly farther away from the axis.'). Cohen's kappa coefficients reveal almost perfect agreement between raters (0.81 for inertia and 0.85 for centre of mass).

**Statistical analyses and model outputs.** We ran a series of Bayesian multilevel models in R[32]. Models were fitted using map2stan in the rethinking package[33], and 95% credible intervals were used to make inferences.

*Analysis 1.* Pre-registered analysis 1 investigated the average speed of wheels across generations in the configurations treatment. Wheels that did not go down were attributed a speed of 0. Data were restricted to participants' last two trials to limit the occurrence of wheels that did not descend in the dataset. We fitted a linear model with 'speed' as the outcome variable, 'trial' and 'generation' as predictor variables and 'player's identity' and 'chain's identity' as random effects (see Supplementary Table 1 for the model output).

*Analysis 2.* Pre-registered analysis 2 investigated understanding across generations in the configurations treatment. We fitted a linear model with 'score' as the outcome variable, 'generation' as a predictor variable and 'chain's identity' as a random effect (see Supplementary Table 2 for model output).

*Analysis 3.* Pre-registered analysis 3 compared the average speed of wheels across generations between treatments. Wheels that did not go down were attributed a speed of 0. Data were restricted to participants' last two trials to limit the occurrence of wheels that did not descend in the dataset. We fitted a linear model with 'speed' as the outcome variable, 'trial', 'generation', 'treatment', 'trial:treatment' and 'generation:treatment' as predictor variables, and 'player's identity' and 'chain's identity' as random effects (see Supplementary Table 3 for model output). For this model, the chains were inefficient and the effective number of samples for one parameter was low (see Supplementary Table 3). The robustness of the model estimates was checked by running additional models (see Supplementary Results). Additional models with more efficient sampling confirmed the reported results (Supplementary Tables 4 and 5 and Supplementary Fig. 8).

*Analysis 4.* Pre-registered analysis 4 compared understanding across generations between treatments. We fitted a linear model with 'score' as the outcome variable, 'generation', 'treatment' and 'generation:treatment' as predictor variables and 'chain's identity' as a random effect (see Supplementary Table 6 for the model output).

*Deviation from pre-registered analyses.* In pre-registered analysis 4, the outcome variable was 'score' and each participant was associated with two values in the dataset: one score for inertia and the other for centre of mass. Compared with the analysis we ran, the pre-registered model included 'physical principle' and 'physical principle:treatment' as predictor variables, and 'player's identity' as a random effect. However, analyses revealed that understanding scores about inertia and centre of mass were negatively correlated (Supplementary Fig. 9 and Supplementary Table 7) and some individuals better understood inertia than centre of mass while others better understood centre of mass than inertia (Fig. 3i and Supplementary Fig. 2). As a result, the pre-registered model did not converge so we ran our analysis on aggregated score and removed the terms associated with the variable 'physical principle' in the reported model.

Data analyses were not performed blind to the conditions of the experiments. No data points were excluded from the analyses.

**Pre-registration.** The study was pre-registered (https://osf.io/ge7cs). Pre-registration of the study design, hypotheses and analysis plan took place before any data were collected.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Code availability

Codes used in this paper are available at https://osf.io/afwmr/.

## Data availability

The data that support the findings of this study are available at https://osf.io/afwmr/.

Received: 10 September 2018; Accepted: 25 February 2019; Published online: 1 April 2019

## References


1. Henrich, J. *The Secret of our Success: How Culture is Driving Human Evolution, Domesticating our Species, and Making us Smarter* (Princeton Univ. Press, 2015).
2. Richerson, P. J. & Boyd, R. *Not by Genes Alone* (Univ. of Chicago Press, 2005).
3. Povinelli, D. J. *World Without Weight: Perspectives on an Alien Mind* (Oxford Univ. Press, 2011).
4. Reader, S. M. & Laland, K. N. Social intelligence, innovation, and enhanced brain size in primates. *Proc. Natl Acad. Sci. USA* **99**, 4436–4441 (2002).
5. Pinker, S. The cognitive niche: coevolution of intelligence, sociality, and language. *Proc. Natl Acad. Sci. USA* **107**, 8993–8999 (2010).
6. Barrett, H. C., Cosmides, L. & Tooby, J. in *The Evolution of Mind: Fundamental Questions and Controversies* 241–248 (Guilford Press, 2007).
7. Bingham, P. M. Human uniqueness: a general theory. *Q. Rev. Biol.* **74**, 133–169 (1999).
8. Boyd, R., Richerson, P. J. & Henrich, J. in *Cultural Evolution: Society, Technology, Language, and Religion* 119–142 (The MIT Press, 2013).
9. Boyd, R., Richerson, P. J. & Henrich, J. The cultural niche: why social learning is essential for human adaptation. *Proc. Natl Acad. Sci. USA* **108**, 10918–10925 (2011).
10. Kyriacou, A. & Bruner, E. Innovation and the evolution of human behavior brain evolution, innovation, and endocranial variations in fossil hominids. *PaleoAnthropology* **2011**, 130–143 (2011).
11. Fuentes, A. *The Creative Spark: How Imagination Made Humans Exceptional* (Penguin, 2017).
12. Derex, M. & Boyd, R. The foundations of the human cultural niche. *Nat. Commun.* **6**, 8398 (2016).
13. Baker, T. in *The Traditional Bower's Bible* 43–116 (The Lyons Press, 1992).
14. Muthukrishna, M. & Henrich, J. Innovation in the collective brain. *Phil. Trans. R. Soc. B* **371**, 20150192 (2016).
15. Proffitt, D. R., Kaiser, M. K. & Whelan, S. M. Understanding wheel dynamics. *Cogn. Psychol.* **22**, 342–373 (1990).
16. Bonawitz, E. et al. The double-edged sword of pedagogy: instruction limits spontaneous exploration and discovery. *Cognition* **120**, 322–330 (2011).
17. Wood, L. A., Kendal, R. L. & Flynn, E. G. Does a peer model's task proficiency influence children's solution choice and innovation? *J. Exp. Child Psychol.* **139**, 190–202 (2015).
18. Proffitt, D. R. & Gilden, D. L. Understanding natural dynamics. *J. Exp. Psychol. Hum. Percept. Perform.* **15**, 384–393 (1989).
19. Kubricht, J. R., Holyoak, K. J. & Lu, H. Intuitive physics: current research and controversies. *Trends Cogn. Sci.* **21**, 749–759 (2017).
20. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83 (2010).
21. Henrich, J. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses: the Tasmanian case. *Am. Antiq.* **69**, 197–214 (2004).
22. Powell, A., Shennan, S. & Thomas, M. G. Late Pleistocene demography and the appearance of modern human behavior. *Science* **324**, 1298–1301 (2009).
23. Kline, M. A. & Boyd, R. Population size predicts technological complexity in Oceania. *Proc. R. Soc. B* **277**, 2559–2564 (2010).
24. Derex, M., Beugin, M.-P., Godelle, B. & Raymond, M. Experimental evidence for the influence of group size on cultural complexity. *Nature* **503**, 389–391 (2013).
25. Muthukrishna, M., Shulman, B. W., Vasilescu, V. & Henrich, J. Sociality influences cultural complexity. *Proc. R. Soc. B* **281**, 20132511 (2014).
26. Hill, K. R., Wood, B. M., Baggio, J., Hurtado, A. M. & Boyd, R. T. Hunter-gatherer inter-band interaction rates: implications for cumulative culture. *PLoS One* **9**, e102806 (2014).
27. Derex, M. & Boyd, R. Partial connectivity increases cultural accumulation within groups. *Proc. Natl Acad. Sci. USA* **113**, 2982–2987 (2016).
28. Creanza, N., Kolodny, O. & Feldman, M. W. Greater than the sum of its parts? Modelling population contact and interaction of cultural repertoires. *J. R. Soc. Interface* **14**, 20170171 (2017).
29. Derex, M., Perreault, C. & Boyd, R. Divide and conquer: intermediate levels of population fragmentation maximize cultural accumulation. *Phil. Trans. R. Soc. B* **373**, 20170062 (2018).

30. Caldwell, C. A. & Millen, A. E. Studying cumulative cultural evolution in the laboratory. *Phil. Trans. R. Soc. B* **363**, 3529–3539 (2008).
31. Zwirner, E. & Thornton, A. Cognitive requirements of cumulative culture: teaching is useful but not essential. *Sci. Rep.* **5**, 16781 (2015).
32. R Development Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2011).
33. McElreath, R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (CRC Press, 2016).

## Acknowledgements
We thank J. Clewett for valuable advice about the building of the wheel, F. Gosselin and A. Deymier for organizing the experimental sessions, and members of the Laboratory for Experimental Anthropology for helpful discussions during the development of the experimental protocol. This project has received funding from the European Union's Horizon 2020 research and innovation programme under Marie Sklodowska-Curie grant agreement number 748310. Support from the ANR-Labex Institute for Advanced Study in Toulouse is acknowledged. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions
R.B. and M.D. developed the research question. M.D. conceived the experimental task and protocol with input from A.M. and J.-F.B. M.D. performed the experiment, analysed the data and wrote the manuscript with input from all authors.

## Competing interests
The authors declare no competing interests.

## Additional information
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41562-019-0567-9.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to M.D.

nature research

Corresponding author(s): Masxime Derex

Last updated by author(s): 30/01/2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☐ | ☒ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | The software that has been developed to collect data is provided as supplementary material (https://osf.io/afwmr/) |
| Data analysis | Srcipts are available as supplementary material (https://osf.io/afwmr/) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data are available as supplementary material (https://osf.io/afwmr/)

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | *Describe how sample size was determined, detailing any statistical methods used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.* |
| Data exclusions | *Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.* |
| Replication | *Describe the measures taken to verify the reproducibility of the experimental findings. If all attempts at replication were successful, confirm this OR if there are any findings that were not replicated or cannot be reproduced, note this and describe why.* |
| Randomization | *Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.* |
| Blinding | *Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.* |

# Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Study description | Experimental study. Quantitative data were collected |
| Research sample | 140 participants took part in the study (70 women and 70 men). Participants were randomly selected from a database managed by Catholic University of Lille. The subjects ranged in age from 18 to 38 y (mean of 20.5, SD of 3.4). |
| Sampling strategy | The number of participants (140) was based on time and monetary constraints. |
| Data collection | The experiment took place in an experimental room at the Laboratory for Experimental Anthropology at Catholic University of Lille. For each session (around 20 minutes long), a single individual was recruited and sat at a computer that was placed parallel to and at 2 meters from the experimental apparatus. Participants were randomly assigned to one condition of the experiment and one sex-segregated chain. Before starting the experiment, participants were asked to complete a consent form and were asked their age. At the end of the experiment, participants indicated whether they have an academic background in physics or engineering. Participants entered and left the room by two different doors to prevent any form of direct interactions between participants. |
| Timing | from 19/02/2018 to 07/03/2018 |
| Data exclusions | No data were excluded from the analysis |
| Non-participation | No participants dropped out |
| Randomization | Participants were randomly assigned to one of two treatmets. |

# Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Study description | *Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.* |
| Research sample | *Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.* |
| Sampling strategy | *Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.* |
| Data collection | *Describe the data collection procedure, including who recorded the data and how.* |
| Timing and spatial scale | *Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken* |

| | |
|---|---|
| Data exclusions | *If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.* |
| Reproducibility | *Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.* |
| Randomization | *Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.* |
| Blinding | *Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.* |

Did the study involve field work? ☐ Yes ☐ No

## Field work, collection and transport

| | |
|---|---|
| Field conditions | *Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).* |
| Location | *State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).* |
| Access and import/export | *Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).* |
| Disturbance | *Describe any disturbance caused by the study and how it was minimized.* |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study | |
|---|---|---|
| ☐ | ☐ | Antibodies |
| ☐ | ☐ | Eukaryotic cell lines |
| ☐ | ☐ | Palaeontology |
| ☐ | ☐ | Animals and other organisms |
| ☐ | ☒ | Human research participants |
| ☐ | ☐ | Clinical data |

### Methods

| n/a | Involved in the study | |
|---|---|---|
| ☐ | ☐ | ChIP-seq |
| ☐ | ☐ | Flow cytometry |
| ☐ | ☐ | MRI-based neuroimaging |

## Antibodies

| | |
|---|---|
| Antibodies used | *Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.* |
| Validation | *Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.* |

## Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | *State the source of each cell line used.* |
| Authentication | *Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.* |
| Mycoplasma contamination | *Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.* |
| Commonly misidentified lines (See ICLAC register) | *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.* |

## Palaeontology

| | |
|---|---|
| Specimen provenance | *Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).* |
| Specimen deposition | *Indicate where the specimens have been deposited to permit free access by other researchers.* |
| Dating methods | *If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.* |

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

| | |
|---|---|
| Laboratory animals | *For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.* |
| Wild animals | *Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.* |
| Field-collected samples | *For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.* |
| Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.* |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | French University Student |
| Recruitment | Participants were randomly selected from a database managed by the Catholic University of Lille and recruited by email from various universities in Lille, France |
| Ethics oversight | University of Exeter Biosciences Research Ethics Committee (2018/2310) and Catholic University of Lille Research Ethics Committee (2018-01-31-E) |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

| | |
|---|---|
| Clinical trial registration | *Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.* |
| Study protocol | *Note where the full trial protocol can be accessed OR if not available, explain why.* |
| Data collection | *Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.* |
| Outcomes | *Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.* |

## ChIP-seq

### Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

| | |
|---|---|
| Data access links<br>*May remain private before publication.* | *For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.* |
| Files in database submission | *Provide a list of all files available in the database submission.* |

Genome browser session
(e.g. UCSC)

*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

## Methodology

Replicates

*Describe the experimental replicates, specifying number, type and replicate agreement.*

Sequencing depth

*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*

Antibodies

*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*

Peak calling parameters

*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*

Data quality

*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*

Software

*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

# Flow Cytometry

## Plots

Confirm that:

- [ ] The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- [ ] The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- [ ] All plots are contour plots with outliers or pseudocolor plots.
- [ ] A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

Sample preparation

*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*

Instrument

*Identify the instrument used for data collection, specifying make and model number.*

Software

*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*

Cell population abundance

*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*

Gating strategy

*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

- [ ] Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

# Magnetic resonance imaging

## Experimental design

Design type

*Indicate task or resting state; event-related or block design.*

Design specifications

*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*

Behavioral performance measures

*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

## Acquisition

Imaging type(s) | Specify: functional, structural, diffusion, perfusion.

Field strength | Specify in Tesla

Sequence & imaging parameters | Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition | State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.

Diffusion MRI ☐ Used ☐ Not used

## Preprocessing

Preprocessing software | Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).

Normalization | If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.

Normalization template | Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal | Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring | Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

## Statistical modeling & inference

Model type and settings | Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested | Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016) | Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction | Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

## Models & analysis

| n/a | Involved in the study |
|---|---|
| ☐ | ☐ Functional and/or effective connectivity |
| ☐ | ☐ Graph analysis |
| ☐ | ☐ Multivariate modeling or predictive analysis |

Functional and/or effective connectivity | Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis | Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis | Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.