nature human behaviour

LETTERS

https://doi.org/10.1038/s41562-018-0489-y

# Even arbitrary norms influence moral decision-making

Campbell Pryor, Amy Perfors and Piers D. L. Howe*

**It is well known that individuals tend to copy behaviours that are common among other people—a phenomenon known as the descriptive norm effect[1–3]. This effect has been successfully used to encourage a range of real-world prosocial decisions[4–7], such as increasing organ donor registrations[8]. However, it is still unclear why it occurs. Here, we show that people conform to social norms, even when they understand that the norms in question are arbitrary and do not reflect the actual preferences of other people. These results hold across multiple contexts and when controlling for confounds such as anchoring or mere-exposure effects. Moreover, we demonstrate that the degree to which participants conform to an arbitrary norm is determined by the degree to which they self-identify with the group that exhibits the norm. Two prominent explanations of norm adherence—the informational and social sanction accounts[2,9–11]—cannot explain these results, suggesting that these theories need to be supplemented by an additional mechanism that takes into account self-identity.**

Previous work on the descriptive norm effect has focused on norms that arise through people's choices. However, there are often situations in which a lack of available alternatives or a lack of incentive to explore these alternatives can lead to a norm occurring. For example, when a default option is available, people may choose the default simply because they do not care enough about the decision to act differently, such as subscribing to a company's default pension plan[12]. This raises the question of whether people will follow norms that have arisen arbitrarily and do not reflect other people's actual preferences.

Two of the most prominent explanations of the descriptive norm effect focus on objectively rational reasons for following descriptive norms. According to the informational account, people might choose to follow the descriptive norm because other people's choices are informative about what is likely to be "effective and adaptive action"[2]. Under this account, the descriptive norm is followed because it provides information as to what is likely to be an appropriate behaviour. Alternatively, the social sanction account proposes that when people prefer a particular behaviour they may respond negatively to people who fail to act in line with this preference. People might therefore conform to descriptive norms to reduce the possibility of social sanctions[9,10]. These rationality-based mechanisms were first delineated by Deutsch and Gerard[11]. From a rationality-based perspective, people would not be expected to follow obviously arbitrary norms because such norms provide no useful information about the value of the options or about potential social sanctions.

Self-categorization theory provides a very different explanation of the descriptive norm effect. This theory proposes that when an individual identifies with a social group, that person will strive to engage in "behaviors that optimally minimize in-group differences and maximize intergroup differences"[13]. Consequently, if their in-group displays a prominent behaviour that is different from the behaviour displayed by the out-group, self-categorization theory predicts that the individual will tend to conform to this behaviour to reinforce their identity as a member of the in-group. Crucially, there is nothing in self-categorization theory that requires the behaviour to reflect the actual preferences of the in-group. Rather, self-categorization theory predicts that conformity will occur whenever the behaviour is salient and characteristic of the in-group[14]. As explained by Terry and Hogg[13], individuals will conform to the norms of a particular group whenever membership of that group is "the contextual basis for [their] self-definition". Thus, self-categorization theory predicts that people should follow salient norms, even if they are arbitrary and do not reflect people's actual preferences, providing they are characteristic of the group to which people perceive they belong. The aim of our study was to test this prediction.

In experiment 1a, participants were presented with a single moral dilemma. Specifically, they were asked to imagine that they had seen a man rob a bank but then give the money to a run-down orphanage that would benefit greatly from the money. These participants were then informed of an arbitrary norm. Half of them were told that in a previous study the majority of participants who were similar to them in terms of age, gender and personality profile (based on a personality questionnaire distributed at the start of the experiment) had been randomly allocated to imagine reporting the robber to the police and then asked to rate how they would feel about doing so (the report condition). The other half of the participants were told the opposite: that the majority of participants in the previous study who were similar to them had been randomly allocated to imagine not reporting the robber to the police and then asked to rate how they would feel about that (the not report condition). Having been told the arbitrary norm, both sets of participants in our experiment were then asked to indicate on a six-point Likert scale how likely they would be to report the robber to the police.

The results are shown in Fig. 1. The dark grey dashed line shows the average response in the report condition. The light grey dashed line shows the average response in the not report condition. Participants tended to conform to the arbitrary norm they were given. In particular, they were more likely to favour reporting the robber when they were told that the majority of their in-group had been randomly allocated to reporting the robber. An ordinal logistic regression run on participants who answered an understanding check correctly (thereby showing that they realized that the norm was arbitrary and did not reflect the previous participants' choices or preferences) found a significant effect of the randomly allocated norm on responses (odds ratio for 68 participants (OR(68)) = 3.64; $P = 0.003$; 95% confidence interval (CI): 1.58–8.65): people shifted their own responses towards the arbitrary norm. The observed odds ratio shows that the odds of choosing an option more in line with

School of Psychological Sciences, University of Melbourne, Melbourne, Victoria, Australia. *e-mail: pdhowe@unimelb.edu.au

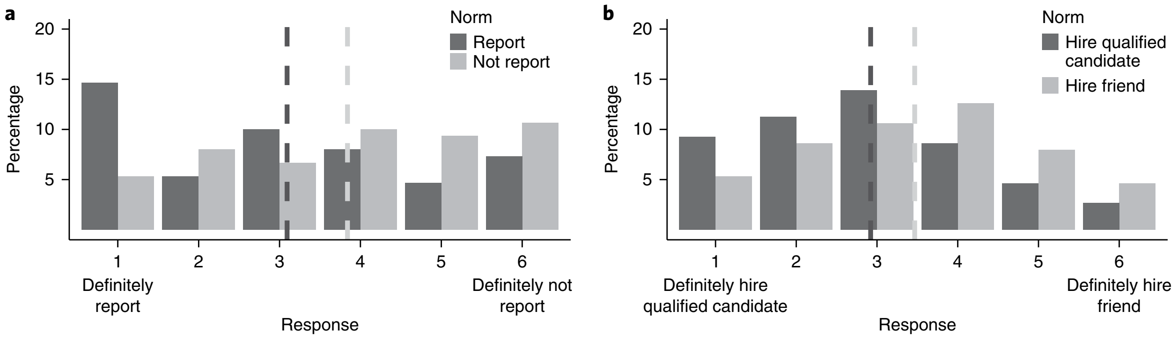


**Fig. 1 | Effect of arbitrary norms in experiments 1a and 1b. a**, Proportion of responses in experiment 1a for both the report (dark grey) and not report (light grey) norm conditions. People shifted their responses in the direction of the arbitrary norm they were given (OR(144) = 2.16; $P$ = 0.009; 95% CI: 1.22–3.85). **b**, Proportion of responses in experiment 1b for when the norm favoured hiring the more qualified candidate (dark grey) or hiring the friend (light grey). As in experiment 1a, participants tended to follow the arbitrary norm (OR(145) = 1.98; $P$ = 0.019; 95% CI: 1.12–3.53). The dashed lines represent the mean of responses in each condition.

reporting the robber were approximately four times higher in the report condition than in the not report condition. These results remain qualitatively the same if we include all participants, regardless of their response on the understanding check (OR(144) = 2.16; $P$ = 0.009; 95% CI: 1.22–3.85), or if we only exclude those participants who chose the understanding check option corresponding to believing that the norm reflected other participants' choices (OR(110) = 2.30; $P$ = 0.013; 95% CI: 1.20–4.46). In this, and all subsequent analyses reported in this paper, we failed to find sufficient evidence ($P > 0.05$) to reject the proportional odds assumption of the ordinal logistic regression (that is, that the effect of the norm was the same across response thresholds).

Experiment 1b conceptually replicated the results of experiment 1a with a single, different moral dilemma that involved participants choosing whether to hypothetically hire a friend or a more qualified candidate for a job at their firm. Half of the participants were told that in a previous study the majority of participants who were similar to them in terms of age, gender and personality profile had been randomly assigned to imagine hiring the more qualified candidate and then asked to rate how they would feel about doing so (the hire qualified candidate condition). The remaining half of the participants were told the opposite: that in a previous study the majority of participants who were similar to them had been randomly assigned to imagine hiring their friend and then asked to rate how they would feel about doing so (the hire friend condition). As shown in Fig. 1, participants tended to conform with the arbitrary norm they were presented with. After excluding participants who failed an understanding check, an ordinal logistic regression found that the odds of showing stronger preference towards hiring the more qualified candidate were significantly higher for participants in the hire qualified candidate condition than in the hire friend condition (OR(87) = 2.84; $P$ = 0.006; 95% CI: 1.35–6.08). As before, we obtained qualitatively the same result if we included all participants (OR(145) = 1.98; $P$ = 0.019; 95% CI: 1.12 to 3.53) or excluded only those participants who indicated a belief that the norm arose through participant choices (OR(99) = 2.88; $P$ = 0.003; 95% CI: 1.43–5.93).

Mere-exposure studies have found that simply increasing exposure to an item can increase preference for that item[15]. Thus, it may have been that the increased preference for the arbitrarily popular option observed in the previous studies arose simply due to its increased exposure in our instructions to the participants. Experiment 2 controlled for this by conceptually replicating experiment 1a, but presenting the descriptive norms in terms of which option was less common. If participants were previously following the norm due to mere-exposure effects, they should have shown a preference for this randomly unpopular action. In contrast, if they were following the actual norm, they should still have preferred the randomly popular action (that is, the action not described in the normative statement).

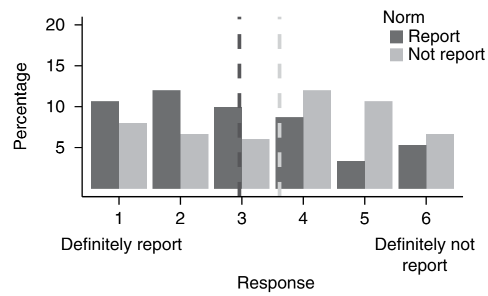


**Fig. 2 | Effect of arbitrary norms that refer to which option is unpopular.** Proportion of responses for both the report (dark grey) and not report (light grey) norm conditions in experiment 2. As in experiments 1a and 1b, people continued to shift their responses towards the arbitrary norm (OR(144) = 2.08; $P$ = 0.0127; 95% CI: 1.17–3.72). This result suggests that mere-exposure effects were not driving norm conformity. The dashed lines represent the average response in each condition.

As Fig. 2 shows, people still followed the arbitrary norm, even when it was expressed in a different way. An ordinal logistic regression found that the odds of choosing an option that more strongly favours reporting the robber were significantly higher for participants in the report condition (who were told that relatively few participants were allocated to imagine not reporting the robber) than those in the not report condition, when excluding participants who failed the understanding check (OR(66) = 4.68; $P < 0.001$; 95% CI: 1.95–11.72). As before, we draw the same conclusions if we include all participants regardless of their response on the understanding check (OR(144) = 2.08; $P$ = 0.0127; 95% CI: 1.17–3.72) or if we exclude only those participants who chose the understanding check option corresponding to believing that the norm arose through participant choices (OR(116) = 2.39; $P$ = 0.008; 95% CI: 1.26–4.59). These results are not consistent with a mere-exposure explanation, instead suggesting that people like to do what others are doing and are not simply choosing the option that they have been exposed to more.

There were still two further potential confounds that needed to be addressed. First, participants may have been systematically misreading the normative statement. The fact that we found descriptive norm effects even when we excluded participants who failed the

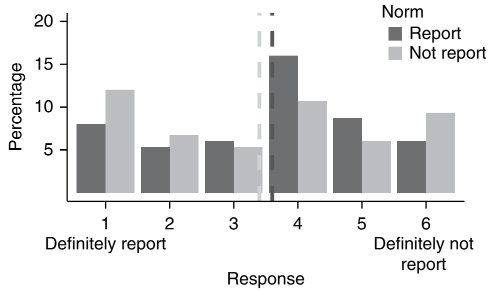


**Fig. 3 | Effect of unrelated arbitrary norms.** Proportion of responses in experiment 3 for both the report (dark grey) and not report (light grey) norm conditions. Participants did not significantly shift their responses to more closely match an unrelated arbitrary norm (OR(144) = 0.83; $P$ = 0.512; 95% CI: 0.47–1.46). This suggests that the results from the previous experiment did not simply reflect an anchoring effect or a failure to read the norm statement properly. Dashed lines represent the average response in each condition.

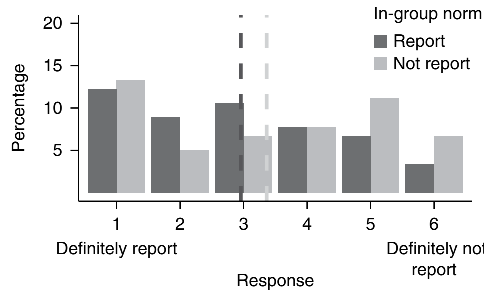


**Fig. 4 | Effect of arbitrary norms when both an in-group and out-group norm are presented.** Proportion of responses in experiment 4 as a function of whether the in-group norm favours reporting or not reporting the robber (with the out-group norm favouring the opposite in each case). Responses tended to shift towards the in-group norm rather than the out-group norm (OR(78) = 2.64; $P$ = 0.0155; 95% CI: 1.21–5.85). Although not apparent from this figure, this effect was strongest for participants who more strongly identified with the in-group over the out-group (OR(113) = 1.30; $P$ = 0.037; 95% CI: 1.02–1.67). Dashed lines represent the average response in each condition.

understanding check provides evidence against such an explanation. However, it is still possible that participants focused primarily on the percentage information (for example, "75%") and the corresponding action (for example "call the police and report the robber"), skipping over the rest of the normative statement during the trial. Thus, at the time of making a choice, they may have implicitly treated the norm as reflecting preferences, ignoring the fact that it was random. Second, an anchoring process may have been driving people's conformity to the arbitrary norm. Anchoring is the phenomenon that presenting information at one end of a scale biases judgements towards that end of the scale[16]. For example, Tversky and Kahneman[16] found that people's estimates of various percentages were systematically higher when they had previously spun a high number on a wheel with numbers ranging from 0 to 100. In the case of our study, it could be that participants were anchoring to the norm, thereby shifting their preferences towards the option that was arbitrarily more popular.

Experiment 3 tested these potential explanations by re-performing experiment 1a, where participants were presented with the same, single moral dilemma, but in this case the normative statement was altered such that it no longer related to the current moral dilemma. If participants were systematically misreading the normative statement, they should not be sensitive to such changes. Similarly, anchoring occurs even when the information is clearly irrelevant to the current task[17,18], so if their behaviour is being driven by anchoring, we would expect participants to continue to follow this now unrelated norm. In contrast, if participants fully comprehend the normative statement, they should recognize that it does not relate to the current moral dilemma, so should not follow it.

The proportions of responses in each condition are shown in Fig. 3. After excluding any participant who failed the comprehension check, we found that participants did not significantly conform to the norm when the norm was derived from an unrelated moral dilemma (OR(85) = 0.79; $P$ = 0.524; 95% CI: 0.38–1.63). We draw the same conclusions if we include all participants regardless of their response on the understanding check (OR(144) = 0.83; $P$ = 0.512; 95% CI: 0.47–1.46) or if we only exclude those participants who chose the understanding check option corresponding to believing that the norm arose through participant choices (OR(103) = 0.75; $P$ = 0.395; 95% CI: 0.38–1.45).

In fact, this effect is significantly weaker than the effect of the norm observed in experiment 1a, as evidenced by an interaction between which action was randomly popular and the experiment (OR(157) = 0.21; $P$ = 0.005; 95% CI: 0.07–0.62). Given that the only difference between these experiments was that experiment 3 stated that the norm was not related to the current moral dilemma, this suggests that the conformity to arbitrary norms observed in previous experiments was not simply the result of anchoring or failing to read the norm statement carefully.

The previous experiments found that people follow arbitrary norms, suggesting that a mechanism that does not require the norms to reflect actual preferences must contribute to the descriptive norm effect. A possible mechanism is provided by self-categorization theory, which predicts that people who more strongly identify with a group are more likely to follow the norms of that group. Experiment 4 tested this prediction. In this experiment, participants were presented with a single moral dilemma along with opposing norms from an intended in-group (people the participant was expected to identify with) and an out-group (people the participant was expected not to identify with). If identity does not drive conformity to arbitrary descriptive norms, these opposite norms should cancel out and we should not see a bias in either direction. In contrast, self-categorization theory predicts that participants who more strongly identify with the in-group over the out-group will more strongly conform to the arbitrary in-group norm over the arbitrary out-group norm.

Figure 4 shows the distribution of responses to the moral dilemma for the 121 participants who answered the understanding check correctly, as a function of whether the in-group norm favoured reporting the robber and the out-group norm favoured not reporting, or vice versa. An ordinal logistic regression did not find a significant main effect of whether the in-group or out-group norm favoured reporting the robber (OR(113) = 0.90; $P$ = 0.811; 95% CI: 0.38–2.11) or a main effect of the extent to which the in-group was identified with over the out-group (OR(113) = 1.04; $P$ = 0.679; 95% CI: 0.87–1.24). Crucially, there was a significant interaction between these two variables, wherein participants who more strongly identified with the in-group over the out-group conformed more strongly to the in-group norm rather than the out-group norm (OR(113) = 1.30; $P$ = 0.037; 95% CI: 1.02–1.67). We also found that there was a significant simple effect of the norms for participants who identified with the in-group more than the out-group, wherein participants tended to conform to the arbitrary in-group norm rather than the arbitrary out-group norm (OR(78) = 2.64; $P$ = 0.0155; 95% CI: 1.21–5.85). These results show that the degree to which participants identify with a group determines the extent to which they follow random, arbitrary norms associated with that group, as predicted by self-categorization theory.

In summary, we found that, across multiple experiments, people conform to random, arbitrary norms that do not reflect others' actual preferences. Experiment 1a found that arbitrary norms influenced participants' preferences. Experiment 1b replicated this finding with a different moral dilemma. Experiment 2 conceptually replicated this finding again and showed that the result continued to hold even when the wording was altered to focus attention on the least popular option, ruling out the possibility that people followed the stated norms merely because of increased exposure to that option.

Of course, these results would not be interesting if they arose from our participants being confused and believing that the reported norms actually did convey information about the preferences of previous participants. We addressed this concern in two ways. First, at the end of each experiment, we ran a comprehension check to identify any participants who held this belief. Excluding these participants did not alter the pattern of our results. Second, experiment 3 presented a norm that referred to a different task, so was not relevant to the task at hand. If participants were blindly following whatever norm was presented to them, this norm should have also affected their behaviour. Similarly, if our previous results were due to anchoring, we would have expected people to continue to conform to the norm, even in experiment 3, since anchoring can even be caused by irrelevant information. As participants did not conform to the norm in experiment 3, we can exclude both possibilities: our previous results were not due to either anchoring or our participants misreading the norm.

Finally, experiment 4 found that participants more strongly conform to arbitrary norms of groups they identify with over those they do not identify with. This was demonstrated by presenting participants with two opposing norms: participants who identified with one group over the other tended to conform to the norm of the group they identified with rather than the group they did not identify with. This result is consistent with previous findings that the degree to which participants' self-identified with an in-group determined the degree of their conformity to the norm of the in-group[19–21]. However, those previous studies presented non-arbitrary norms and, thus, offered rational reasons for participants to more strongly conform when self-identification with the in-group was high. For example, individuals may be more likely to conform to the norms of a group when they more strongly identify with the group because, in such circumstances, it is more likely that their goals and values are similar to those of the in-group. Experiment 4 provided evidence that increasing self-identification with the in-group increases in-group norm conformity, even when there is no objectively rational reason for it to do so—a finding that the informational and social sanction accounts cannot explain. This role of self-identification was successfully predicted by self-categorization theory, which brings us to the theoretical implications of our findings.

Two of the most prominent explanations for the descriptive norm effect—the informational account and the social sanction account—both assume that people follow norms because norms reflect something useful and informative about other people's preferences[11]. Our finding that people follow arbitrary norms shows that norms do not have to provide useful information about people's preferences for norm conformity to occur. This result cannot be explained by either the informational or social sanction accounts as both accounts predict that people should not conform to arbitrary norms. This shows that these prominent theories cannot, on their own, provide a complete account of the descriptive norm effect; an additional mechanism—one involving self-identification—is needed. Our paradigm allows this additional mechanism to be studied, independent of any informational or social sanction effects.

One of the key assumptions of self-categorization theory is that people tend to internalize the characteristics of salient social groups with which they personally identify[14]. In typical studies of self-categorization, the relevant social group is a well-defined group of people who the individual would probably interact with and potentially consider part of their self-identity outside the confines of the experiment, such as people from their university[13,22,23]. In contrast, in our experiments, the relevant social group was not one a current participant would identify with outside the experiment. It consisted of people who happened to previously participate in a similar study and were similar to the current participant, but with whom the current participant would probably never interact. Despite this, we found that participants tended to follow the norms of these impersonal, unfamiliar groups. This finding is consistent with the claim of self-categorization theory that people's identity is not stable but is instead driven by context[14]. Although these groups may not have been how participants typically defined themselves, in the context of our experiments, these groups became salient, and so formed the basis for self-categorization.

In conclusion, our work suggests that people conform to descriptive norms even when they are entirely arbitrary. These results are consistent with the self-categorization account of norm adherence and cannot be explained by either the informational account or the social sanction account. Our results generalize to different dilemmas and cannot be attributed to participants misunderstanding or misreading the arbitrary norms, or to anchoring effects. Given the impact descriptive norms can have on decisions, and their increasing application in nudging strategies, the better we understand the mechanisms underlying this normative influence, the better we can design normative nudges to encourage prosocial behaviour.

## Methods

The experiments were run using Amazon's Mechanical Turk (MTurk) because MTurk participants tend to cover a wider range of ages and ethnicities than the traditional university participant pool[24] and match the general population closely in terms of the distribution of occupations. Although they are slightly younger than the general population, they are much closer to that standard than university undergraduates[25]. From a quality-of-data point of view, MTurk participants are similar to a traditional university participant pool. Hauser and Schwarz[26] found that MTurk participants actually outperformed laboratory participants in their comprehension of instructions. Findings from laboratory studies tend to be replicated in MTurk samples[27–29] and the test–retest reliability of MTurk data is generally quite high[30]. Kittur et al.[31] reported that failing to include understanding checks can substantially reduce the quality of data. Consequently, understanding checks were included in all of our experiments.

In all experiments, participants completed the experiment in a web browser and were paid US$0.65 for participating. Each experiment took around 2–3 min to complete. Participants were not allowed to participate in more than one experiment. Ethics approval for this research was granted by the University of Melbourne Human Research Ethics Committee, and all participants gave informed written consent. Data collection and analysis were not performed blind to the conditions of the experiments.

Experiment 1a recruited 150 English-speaking participants (mean age = 36 years; 43% female). This sample size was based on a power analysis using 10,000 bootstrapped samples from a pilot experiment we ran on 105 participants (see Supplementary Note 1), which was largely the same as experiment 1a. Power was calculated as the proportion of bootstrapped samples that had a significant result ($P < 0.05$) on the ordinal logistic regression outlined below. This power analysis suggested that a sample size of 72 would provide over 80% power to observe a significant result. The (unreported) pilot experiment did not include an understanding check but, as outlined below, all of the experiments we ran included a preplanned understanding check as an exclusion criterion. To ensure adequate power, even in the face of losing some participants who failed the understanding check, we therefore increased the sample size to 150 for each of our studies.

In experiment 1a, on agreeing to participate, participants provided some demographic information and completed a ten-item personality questionnaire taken from Gosling and colleagues[32]. On the basis of their responses, people were presented with a 'personality profile' that informed them of which two 'big five' personality traits (openness, agreeableness, emotional stability, extraversion and conscientiousness) they scored highest on.

As a cover story to justify the arbitrary norm presented later in the experiment, participants were then informed that we were following up on a previous paper that looked at how people feel during moral dilemmas. Crucially, they were told that participants in the previous study were not asked to choose how they would act in the moral dilemma, but instead were randomly allocated to imagine performing a specific action and to then rate how good or bad they would feel about performing that action. After reading the instructions, participants in experiment 1a proceeded to the experimental trial, where they were presented with the following moral dilemma:

"Imagine you have witnessed a man rob a bank. However, you then saw him do something unexpected with the money. He donated it all to a run-down orphanage that would benefit greatly from the money. You must decide whether to call the police and report the robber or do nothing and leave the robber alone."

Immediately below this moral dilemma was a normative statement. Importantly, this statement made it clear that the norm had arisen not through choices, but instead through an error in the random allocation process, which meant that it provided no information about the actual beliefs or values of the previous participants. The specific wording of the normative statement was:

"In the previous study, approximately 75% of [gender] participants aged [age range] that rated high in [personality description] were allocated to [randomly popular action] and then rate how they would feel, due to an error in their random allocation."

When presented to participants, [gender] was replaced with the participant's gender, [age range] was replaced with the participant's five-year age range (for example, 30–35 years) and [personality description] was replaced with the two personality traits that the participant had rated high on (for example, extraversion and agreeableness). Also, [randomly popular action] was replaced with one of the available actions, varied randomly between participants. Thus, half of the participants were informed that the [randomly popular action] was to "call the police and report the robber" (the report condition) while the remaining half were informed that it was to "do nothing and leave the robber alone" (the not report condition).

Participants were then asked about what action they would choose by responding on the following six-point Likert scale:

(1) Definitely call the police and report the robber.
(2) Very likely call the police and report the robber.
(3) Probably call the police and report the robber.
(4) Probably do nothing and leave the robber alone.
(5) Very likely do nothing and leave the robber alone.
(6) Definitely do nothing and leave the robber alone.

After completing this main experimental trial, participants rated how good or bad they would feel about performing their chosen action. This was included purely for the sake of consistency with the cover story offered in the introduction and was not analysed. Finally, participants completed an understanding check, asking them to identify the problem with the previous study that we described in the instructions. This was included to ensure that participants believed that the norm they saw really was arbitrary and did not mistakenly think that it reflected the previous participants' preferences. The response options in the understanding check, presented in a random order, were as follows (where option 1 is correct):

(1) An error meant that more participants were allocated to rate one action than the other.
(2) One action was preferred by most participants, while very few participants said they were likely to perform the other action.
(3) No data were saved during the experiment.
(4) The participants completed the experiment with their eyes closed.

A one-shot between-subjects design was used in experiment 1a, where participants were only presented with a single moral dilemma. The independent variable, which was randomly allocated, was whether the arbitrary norm said that more people had been randomly allocated to rate reporting the robber (report condition; $n=75$) or not reporting the robber (not report condition; $n=75$). The dependent variable was participants' responses on the Likert scale rating the certainty with which they would act a certain way.

An ordinal logistic regression was run to assess the extent to which preferences towards reporting the robber increased for participants in the report condition compared with those in the not report condition. All analyses reported in the paper were two tailed. Obviously, it would not be interesting if people followed the norm because they were under the misapprehension that it reflected the true beliefs and values of people like themselves. Consequently, we excluded from the analysis any participant who answered the understanding check incorrectly; however, as we describe in our results, our conclusions remain the same when all participants are included or even when we only exclude participants who chose option 2 above. The code for all analyses reported in this paper is available in the Supplementary Software file.

Experiment 1b was exactly the same as experiment 1a, except it was run on a new sample of 151 participants (mean age = 35 years; 40% female) who responded to a single moral dilemma, but the moral dilemma was changed in experiment 1b compared with experiment 1a:

"Imagine you are hiring someone for a job at your firm. Your friend has applied for the position and is qualified. However, another applicant seems to be even more qualified."

This time, participants were presented with a norm that stated either that more people had been allocated to imagine "hiring the more qualified candidate" ($n=76$) or that more people had been allocated to imagine "hiring their friend" ($n=75$), and the response options were updated accordingly.

Experiment 2 was exactly the same as experiment 1a, where participants were shown a single moral dilemma about reporting a robber who donated the money they stole to an orphanage, except it was run on a new sample of 150 participants (mean age = 36 years; 54% female) and the phrase "approximately 75%" in the normative statement was replaced with "approximately 25%". For simplicity, we will continue to refer to variables in terms of which action is randomly popular ($n=75$ per condition). For example, participants in the report condition were shown a normative statement that the option to "do nothing and leave the robber alone" had "approximately 25%" of similar participants allocated to it, implying that most similar participants were allocated to imagine and rate reporting the robber.

Experiment 3 was, again, the same as experiment 1a except that it was run on a new sample of 150 participants (mean age = 35 years; 44% female; $n=75$ per condition) and the normative statement informed participants that the supposed previous study presented participants with an unrelated moral dilemma. As in the other experiments reported here, experiment 3 was a one-shot design where participants responded to a single moral dilemma. The specific wording of the normative statement was as follows:

"The previous study used a completely different moral dilemma that happened to also involve a robber. In the previous study, approximately 75% of [gender] participants aged [age range] that rated high in [personality description] were allocated to randomly popular action] and then rate how they would feel, due to an error in their random allocation."

Experiment 4 was preregistered at http://aspredicted.org/blind.php?x=hr7cs4. It was the same as experiment 1a except for two important changes. First, participants were still presented with a single moral dilemma, but this dilemma was now accompanied by two norms instead of one. One of these norms was the same in-group norm that was presented in experiment 1a. The other was a norm from a supposed out-group that favoured the opposite option to the in-group norm. For example, if a female participant rated high in extraversion and agreeableness and was allocated to an in-group norm favouring reporting the robber, she would be shown the following norms:

"In the previous study:

- approximately 75% of female participants that rated high in extraversion and agreeableness were allocated to imagine calling the police and reporting the robber and then rate how they would feel, due to an error in their random allocation.
- approximately 75% of male participants that rated low in extraversion and agreeableness were allocated to imagine doing nothing and leaving the robber alone and then rate how they would feel, due to an error in their random allocation."

To avoid any potential confound due to participants focusing more on the first or second piece of information, we also randomly varied whether the in-group norm was presented before or after the out-group norm.

Second, we included two questions at the end of the experiment to measure the extent to which participants actually identified with the alleged in-group (for example, females who rated high in extraversion and agreeableness) and out-group (for example, males who rated low in extraversion and agreeableness). Specifically, we used the single-item social identification measure developed by Postmes et al.[33], which asked participants to rate their agreement with the statements "I identify with [in-group description]" and "I identify with [out-group description]", on a seven-point Likert scale ranging from "fully disagree" to "fully agree".

Experiment 4 was run on a new sample of 180 participants (mean age = 34 years; 38.3% female). A bootstrap power analysis based on a pilot sample (see Supplementary Note 2) of 75 participants (who were not included in any of the main analyses for experiment 4) suggested that a sample of 180 participants would give over 95% power to observe a significant result for the effect of interest (the interaction between which actions the in-group and out-group norms favoured, and how much the participant identified with the in-group over the out-group).

Experiment 4 used a 2 (in-group norm favours reporting the robber and out-group norm favours not reporting the robber ($n=89$) or vice versa ($n=91$)) × 2 (in-group norm presented first ($n=94$) or second ($n=86$)) between-subjects design. Given that we only varied whether the in-group norm was presented first or second to control for a potentially confounding bias, we collapsed across this when analysing the data. Participants who failed the understanding check were excluded from the analysis.

The data were analysed using a preregistered ordinal logistic regression, which predicted responses to the moral dilemma as a function of the following two variables and their interaction:

(1) Whether the in-group norm favoured reporting the robber and the out-group norm favoured not reporting the robber, or vice versa.
(2) The difference between how much the participant identified with the supposed in-group compared with the out-group. This was calculated by subtracting the participant's rating of identification with the out-group from their rating of identification with the in-group.

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

**Code availability.** The code used for all analyses reported in this paper is available in the Supplementary Software file, and a live version is available in the following OSF repository: https://osf.io/n6uz5/?view_only=7dc67fcc0c1f4fdea8fe1dfe5d492480.

## Data availability

The response data for all pilot experiments and experiments in this paper are available in the following OSF repository: https://osf.io/n6uz5/?view_only=7dc67fcc0c1f4fdea8fe1dfe5d492480.

Received: 1 May 2018; Accepted: 2 November 2018;
Published online: 17 December 2018

## References

1. Burger, J. M. et al. Nutritious or delicious? The effect of descriptive norm information on food choice. *J. Soc. Clin. Psychol.* **29**, 228–242 (2010).
2. Cialdini, R. B., Reno, R. R. & Kallgren, C. A. A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* **58**, 1015–1026 (1990).
3. Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J. & Griskevicius, V. Normative social influence is underdetected. *Pers. Soc. Psychol. Bull.* **34**, 913–923 (2008).
4. Hallsworth, M. et al. Stating appointment costs in SMS reminders reduces missed hospital appointments: findings from two randomised controlled trials. *PLoS ONE* **10**, 1–14 (2015).
5. Wenzel, M. Misperceptions of social norms about tax compliance: from theory to intervention. *J. Econ. Psychol.* **26**, 862–883 (2005).
6. Goldstein, N. J., Cialdini, R. B. & Griskevicius, V. A room with a viewpoint: using social norms to motivate environmental conservation in hotels. *J. Consum. Res.* **35**, 472–482 (2008).
7. Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J. & Griskevicius, V. The constructive, destructive, and reconstructive power of social norms. *Psychol. Sci.* **18**, 429–434 (2007).
8. *Applying Behavioural Insights to Organ Donation: Preliminary Results From a Randomised Controlled Trial* (Cabinet Office Behavioural Insights Team, NHS Blood and Transplant, Department of Health, Government Digital Service & Driver and Vehicle Licensing Agency, 2013).
9. Ajzen, I. & Fishbein, M. *Understanding Attitudes and Predicting Social Behaviour* (Prentice Hall, Englewood Cliffs, 1980).
10. Bendor, J. & Swistak, P. The evolution of norms. *Am. J. Sociol.* **106**, 1493–1545 (2001).
11. Deutsch, M. & Gerard, H. B. A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* **51**, 629–636 (1955).
12. Choi, J. J., Laibson, D., Madrian, B. C. & Metrick, A. in *Analyses in the Economics of Aging* (ed. Wise, D. A.) 59–78 (Univ. Chicago Press, Chicago, 2005).
13. Terry, D. J. & Hogg, M. A. Group norms and the attitude–behavior relationship: a role for group identification. *Pers. Soc. Psychol. Bull.* **22**, 776–793 (1996).
14. Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. & Wetherell, M. S. *Rediscovering The Social Group: A Self-Categorization Theory* (Basil Blackwell, Oxford, 1987).
15. Zajonc, R. B. Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* **9**, 1–27 (1968).
16. Tversky, A. & Kahneman, D. Judgment under uncertainty: heuristics and biases. *Science* **185**, 1124–1131 (1974).
17. Englich, B. & Mussweiler, T. Sentencing under uncertainty: anchoring effects in the courtroom. *J. Appl. Soc. Psychol.* **31**, 1535–1551 (2001).
18. Englich, B., Mussweiler, T. & Strack, F. Playing dice with criminal sentences: the influence of irrelevant anchors on experts' judicial decision making. *Pers. Soc. Psychol. Bull.* **32**, 188–200 (2006).
19. Neighbors, C. et al. Group identification as a moderator of the relationship between perceived social norms and alcohol consumption. *Psychol. Addict. Behav.* **24**, 522–528 (2010).
20. Smith, J. R. & Terry, D. J. Attitude–behaviour consistency: the role of group norms, attitude accessibility, and mode of behavioural decision-making. *Eur. J. Soc. Psychol.* **33**, 591–608 (2003).
21. Rimal, R. N. Modeling the relationship between descriptive norms and behaviors: a test and extension of the theory of normative social behavior (TNSB). *Health Commun.* **23**, 103–116 (2008).
22. Platow, M. J. et al. "It's not funny if they're laughing": self-categorization, social influence, and responses to canned laughter. *J. Exp. Soc. Psychol.* **41**, 542–550 (2005).
23. Cruwys, T. et al. Modeling of food intake is moderated by salient psychological group membership. *Appetite* **58**, 754–757 (2012).
24. Behrend, T. S., Sharek, D. J., Meade, A. W. & Wiebe, E. N. The viability of crowdsourcing for survey research. *Behav. Res. Methods* **43**, 800–813 (2011).
25. Huff, C. & Tingley, D. "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Res. Polit.* **2**, 1–12 (2015).
26. Hauser, D. J. & Schwarz, N. Attentive turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behav. Res. Methods* **48**, 400–407 (2016).
27. Horton, J. J., Rand, D. G. & Zeckhauser, R. J. The online laboratory: conducting experiments in a real labor market. *Exp. Econ.* **14**, 399–425 (2011).
28. Huh, Y. E., Vosgerau, J. & Morewedge, C. K. Social defaults: observed choices become choice defaults. *J. Consum. Res.* **41**, 746–760 (2014).
29. Paolacci, G., Chandler, J. & Ipeirotis, P. G. Running experiments on Amazon Mechanical Turk. *Judgem. Decis. Mak.* **5**, 411–419 (2010).
30. Buhrmester, M., Kwang, T. & Gosling, S. D. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspect. Psychol. Sci.* **6**, 3–5 (2011).
31. Kittur, A., Chi, E. H. & Suh, B. Crowdsourcing user studies with Mechanical Turk. In *Proc. SIGCHI Conference on Human Factors in Computing Systems* 453–456 (Association for Computing Machinery, 2008).
32. Gosling, S. D., Rentfrow, P. J. & Swann, W. B. A very brief measure of the big-five personality domains. *J. Res. Pers.* **37**, 504–528 (2003).
33. Postmes, T., Haslam, S. A. & Jans, L. A single-item measure of social identification: reliability, validity, and utility. *Br. J. Soc. Psychol.* **52**, 597–617 (2013).

## Acknowledgements

This research was supported by an Australian Government Research Training Program Scholarship. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

All authors contributed to the study design. Testing, data collection and analysis were performed by C.P. under the supervision of P.D.L.H. and A.P. C.P. drafted the manuscript, and P.D.L.H. and A.P. provided critical revisions. All authors approved the final version of the manuscript for submission.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41562-018-0489-y.

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

**Correspondence and requests for materials** should be addressed to P.D.L.H.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Piers Howe

# 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.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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 |
| ☐ | ☒ | An indication of 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 statistics 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 |
| ☒ | ☐ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | All data was collected using custom web experiments coded using HTML, Javascript, CSS and PHP |
| Data analysis | Data was all analysed using R v3.3.2 with the "ordinal" package |

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 upon request. 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

The response data collected for this paper are available in the following OSF repository: https://osf.io/n6uz5/?view_only=7dc67fcc0c1f4fdea8fe1dfe5d492480



April 2018

# Field-specific reporting

Please select 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/authors/policies/ReportingSummary-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 | Quantitative experimental |
| Research sample | English-speaking Mechanical Turk sample (Mean age = 35, 44% females). This sample is fairly typical of a Mechanical Turk study and existing research suggests Mechanical Turk samples are relatively representative of populations outside of Mechanical Turk participants. |
| Sampling strategy | Sample were collected using convenience samples on the Mechanical Turk platform. We conducted a power analysis using 10,000 bootstrapped samples from a pilot study we ran on 105 participants. Power was calculated as the proportion of bootstrapped samples that had a 95%CI excluding OR=1 using the analysis outlined in the paper. This power analysis suggested that a sample size of 72 would provide over 80% power. This pilot study did not include an understanding check whereas, in the paper, we excluded participants that failed an understanding check. To be safe, we therefore increased the sample size to 150 for each of our experiments.<br>The final experiment utilized a different analysis and so a power analysis was run on a pilot of Experiment 4 which suggested a sample of 180 participants would provide over 95% power to detect the effect of interest. |
| Data collection | All data was collected online. Participants completed the experiment by accessing a website. The researchers were not present when any participants were completing the experiment and thus, had no influence over any specific responses. |
| Timing | Data collection begun on 1st November 2017 and concluded on 14th July 2018 |
| Data exclusions | Data from participants who failed an understanding check were excluded from our analysis. Across our experiments, this totaled 330 participants. This exclusion rule was planned before analyzing the data, though for completeness we also report the results when such participants were not excluded. The pattern of results remains the same regardless. For the final experiment, this exclusion rule was pre-registered. |
| Non-participation | No participants dropped out from the study. We cannot determine how many people declined to participate given recruitment was based on a public ad posted on Mechanical Turk |
| Randomization | Participants were randomly allocated to the experimental conditions. |

# 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

### Materials & experimental systems

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

### Methods

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

## Unique biological materials

Policy information about availability of materials

| | |
|---|---|
| Obtaining unique materials | *Describe any restrictions on the availability of unique materials OR confirm that all unique materials used are readily available from the authors or from standard commercial sources (and specify these sources).* |

## 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.* |

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | *Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."* |
| Recruitment | *Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.* |

## 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 *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.*