

## Sex differences in confidence influence patterns of conformity

Catharine P. Cross<sup>1,2</sup>, Gillian R. Brown<sup>1\*</sup>, Thomas J. H. Morgan<sup>2,3</sup>, and Kevin N. Laland<sup>2</sup>

1. School of Psychology & Neuroscience, University of St Andrews, U.K.

2. Centre for Social Learning and Cognitive Evolution, School of Biology, University of St Andrews, U.K.

3. Department of Psychology, University of California, Berkeley, USA.

In press: *British Journal of Psychology*

\* Correspondence to Dr Gillian R. Brown, School of Psychology & Neuroscience, University of St Andrews, South Street, St Andrews, Fife, KY16 9JP, U.K (e-mail: [grb4@st-andrews.ac.uk](mailto:grb4@st-andrews.ac.uk))

### Acknowledgements

We are grateful to Alex Mesoudi, Volker Franz and two anonymous reviewers for comments on the manuscript. The research was supported in part by an ERC Advanced Grant (EVOCULTURE, Ref. 232823) awarded to KNL.

## **Abstract**

Lack of confidence in one's own ability can increase the likelihood of relying on social information. Sex differences in confidence have been extensively investigated in cognitive tasks, but implications for conformity have not been directly tested. Here, we tested the hypothesis that, in a task that shows sex differences in confidence, an indirect effect of sex on social information use will also be evident. Participants (N=168) were administered a mental rotation (MR) task or a letter transformation (LT) task. After providing an answer, participants reported their confidence before seeing the responses of demonstrators and being allowed to change their initial answer. In the MR, but not the LT, task, women showed lower levels of confidence than men, and confidence mediated an indirect effect of sex on the likelihood of switching answers. These results provide novel, experimental evidence that confidence is a general explanatory mechanism underpinning susceptibility to social influences. Our results have implications for the interpretation of the wider literature on sex differences in conformity.

## Introduction

When individuals lack certainty in their own ability, copying others can provide a quick and low-cost solution to a broad range of problems and challenges (Laland, 2004; Hoppitt & Laland, 2013). Mathematical models indicate that social learning is an effective route to finding a solution when an individual's own experience or information appears to be insufficient (Boyd & Richerson, 1988). Experimental studies have confirmed that human beings are more likely to use social information when lacking confidence (defined as the subjective probability of correctness; Baranski & Petrusic, 1994). For example, Morgan *et al.* (2012) asked participants to judge which of two virtual 'food patches' was more productive, or to complete a two-alternative forced-choice mental rotation task, and then report their confidence in their answers. Participants with low confidence were more likely than participants with high confidence to change their answer if they subsequently found that a majority disagreed with them (Morgan *et al.*, 2012). Similarly, studies of non-human animals show that individuals follow the responses of others when personal information is either lacking or unreliable (Galef, Dudley & Whiskin, 2008; van Bergen, Coolen & Laland, 2004). Levels of certainty among individuals are therefore likely to influence the dynamics of social learning and the spread of information through populations.

In human beings, self-reported confidence can be influenced by factors other than ability or performance. These factors include gender stereotypes, which can lead to lower levels of confidence in the sex that is perceived to be less competent (Oswald, 2008). The psychological literature on confidence strongly suggests that women have lower confidence than men in a variety of contexts. Among university students, for example, women under-rate their academic performance in areas as diverse as economics (Jakobsson, 2012), medicine (Blanch, Hall, Roter, & Frankel, 2008), and biology (Langan *et al.*, 2008). In psychological experiments, too, women under-rate their performance on simple navigation tasks on a computer screen (Picucci, Caffò, & Bosco, 2011), assessments of creativity (Lenney, 1981), and a wide range of other tasks (Maccoby &

Jacklin, 1978). Given that confidence is a strong predictor of using social information, participant sex could have an indirect effect on informational conformity via confidence.

Meta-analyses of the social psychology literature have shown that, on average, women conform more than men (Bond & Smith, 1996; Cooper, 1979; Eagly & Carli, 1981). However, some studies have also explored variation in sex differences in conformity, showing that the use of social information by men and women varies according to the gender-stereotyped content of the task (Goldberg, 1975; Karabenick, 1983); for example, women are more likely than men to conform to the majority response on questionnaire items considered to be ‘masculine’ (e.g., related to fishing), while men are more likely than women to conform on ‘feminine’ items (e.g., related to knitting) (Goldberg, 1975). Given that these topics were selected to invoke beliefs and expectations about the relative competences of the sexes, it is possible that sex differences in confidence underpinned the variable sex differences in conformity.

Here, we tested the prediction that, where sex differences in confidence exist on a task, an indirect effect of sex on conformity will also exist. We used a mental rotation (MR) task (Shepard & Metzler, 1971) that is commonly assumed to favour men (Hausmann *et al.*, 2009) and on which sex differences in confidence are clearly evident (Cooke-Simpson & Voyer, 2007; Estes & Felker, 2012). We also used a letter transformation (LT) task (Majeres, 2007) that has not been widely used and is less likely to be viewed as gender stereotyped. Our aim was to test for a consistent relationship between confidence and conformity across two different tasks that were similar in format but different in content. Each task used a two-alternative forced-choice format, and each presented two stimuli which were either accurate ‘transformations’ of each other or not. The difference lay in the types of information being presented: in the MR task, participants stated whether two 3-D rotated images matched, while in the LT task, participants stated whether a one set of letters has been correctly generated from another. In both tasks, participants provided a response and then rated their confidence in their answer. Participants were then shown the responses of

previous participants ('demonstrators') and provided with the opportunity to change their answer. For both tasks, social information was administered according to a conditional information lottery, which enables social information to be manipulated while avoiding the use of deception.

We predicted that i) low confidence in the initial answer would positively correlate with the likelihood of using social information in both sexes, ii) the likelihood of switching answers in response to social information would correlate positively with the proportion of demonstrators that disagreed with the initial answer in both tasks, and iii) men would have higher average scores, and higher average confidence, than women on the MR task. We tentatively predicted no sex difference in performance or confidence on the LT task. Crucially, we predicted that iv) the sex difference in confidence in the MR task would not be explained entirely by the sex difference in accuracy, v) there would be an indirect effect of sex on use of social information in the MR task, mediated by confidence. We modelled the probability that participants would switch their answers, using participant sex, initial accuracy, confidence, and proportion of demonstrators disagreeing with the initial answer as predictors. All data were analysed using binomial and Bernoulli regression models with Bayesian estimation.

## Methods

**Participants and procedure.** A total of 168 adult participants completed either the MR task (54 women, 33 men) or the LT task (51 women, 28 men). Participants were tested in groups of 2-10 in sessions lasting 20-30 minutes. Each participant had access to a computer, and opaque barriers were used to occlude sight lines to other participants' screens. Participants were briefed upon entering the testing room and provided with instructions that were both read aloud by the experimenter and presented on the computer screen. Participants were informed that they would be reimbursed for their time (£5) and would earn a bonus payment (£0-£5) based on the accuracy of their final

answers. The experimenter sat in the adjacent room during testing. Participants reported their sex at the end of the task and were asked to describe the strategy used when choosing whether or not to use social information (see ‘Social information’ section), before being debriefed. Ethical approval was provided by the local ethics committee, and informed consent was obtained from all participants.

**Tasks.** Both tasks were administered as two-alternative forced-choice tasks. In the MR task, participants were shown an image of two 3-D objects viewed from different orientations, which were either identically shaped or mirror-reversed. After 4 seconds, the images disappeared, and the participant was asked to indicate whether the objects were the same or not by clicking one of two answer buttons (‘match’/‘no match’) with no time limit. Twenty-four pairs of images (i.e. 24 trials) were presented to each participant in a randomised order. In the LT task, participants were shown two sets of letters (e.g., ‘AE’ and ‘DH’), plus a transformation rule (e.g., ‘move the letters three spaces forward in the alphabet’), and were asked whether the second set of letters was a correct transformation of the first. The pair of letter strings and the transformation rule were displayed for 3.2 seconds, before another, randomly selected, set of letters was flashed up to signal the end of the viewing time. Participants indicated whether the transformation had been applied correctly by clicking one of two answer buttons (‘correct’/‘not correct’). Although participants in the LT task could take as long as they wanted to answer, they were informed that they could only earn bonus points by answering within 4 seconds of the end of the viewing time. This time limit prevented a ceiling effect in performance and confidence, by discouraging participants from committing the pairs of letter strings to memory and working out the answer after the viewing time had finished. Answers that were given after 4 seconds were excluded from the analyses (i.e., 171 out of 1817 trials; mean answer time of excluded trials = 8.9 seconds). Twenty-three sets of letter strings and transformation rules (i.e. 23 trials) were presented to each participant in a randomised order.

**Confidence measure and switching answers.** In both tasks, participants were asked ‘How confident are you in your answer?’ after answering and selected a confidence rating on a scale from 1 (low, ‘not at all confident’) to 7 (high, ‘very confident’). Participants were then provided with social information (see ‘Social information’), alongside their initial answer, and were asked, ‘Now do you think the shapes matched?’ in the MR task or ‘Now do you think the transformation was correct?’ in the LT task. Participants then answered again. Participants did not receive any feedback about the accuracy of either the initial answer or the revised answer. In both the MR and LT tasks, the programme only moved onto the next trial once an answer had been given to each of the questions.

**Social information.** Participants saw the answers of a randomly selected number ( $N= 4, 8, \text{ or } 12$ ) of demonstrators. The demonstrators were 12 participants recruited prior to the main experiment, and participants in the main experiment were informed of this. Each answer button on the screen flashed red once for each demonstrator who had made that decision. A conditional information lottery procedure was used to manipulate the social information (Bardsley, 2000). In a conditional information lottery, participants are made aware that some information is real and some is not real. This method avoids the need for deception, which can cause suspicion in participants and create problems for future experiments (Hertwig & Ortmann, 2008). The social information was real in five trials for each participant, and the actual answers of a group of demonstrators were provided. On the remaining trials, participants received experimenter-generated social information, chosen randomly from a battery that spanned the full range from ‘majority correct’ to ‘majority incorrect’. Therefore, the majority response within the manipulated social information was sometimes aligned with the participant’s initial answer and sometimes contradicted the initial answer.

Participants were made aware during the briefing that the social information they received would be real in some trials ('We gathered decisions from pilot participants earlier this year') and manipulated in other trials ('Some of the information is generated by us'). Participants were also informed that their answers would only contribute towards their bonus in cases where the social information was real, based on a previously published protocol (Bardsley, 2000). Participants were informed that the bonus would be between £0-5 but were unaware how many trials contained real information. Participants were thus encouraged to treat all social information as genuine. Previous studies have shown that empirical results obtained using conditional information lottery procedures are very similar to those obtained using traditional methods that involve deception, but with the advantage of not corrupting the subject pool (Bardsley, 2000; Bardsley & Sausgruber, 2005). As in a previous study (Morgan *et al.*, 2012), a free-response text box at the end of the session asked, 'Please tell us how you decided to use the social information', and an *a priori* selection criterion was applied, as follows. Where a participant's comments indicated that the conditional information lottery had not produced the intended effect of motivating participants to treat the information as real (e.g. they reported that they did not 'trust' the information because some of it was 'fake'), their data were discarded (MR: N= 6 participants; LT: N= 10 participants). The final sample size (N=168) excludes these participants, and this was the only exclusion criterion. Data from participants who simply stated that they ignored social information without specifically mentioning a distrust owing to the experimental manipulation were retained in the dataset. Exclusion was decided solely on the basis of the free-response comments, which were assessed (by CPC and TJHM) without viewing actual switching rates.

**Statistical analyses.** All data were analysed with binomial and Bernoulli regression models using Bayesian estimation. This approach allows a model to be constructed that incorporates several simultaneously varying parameters (e.g. sex, confidence, accuracy) and allows the trial, rather than



the participant, to be used as the unit of analysis by specifying random participant-level effects (Bolker *et al.*, 2009). We used Markov Chain Monte Carlo methods to fit the models in OpenBUGS (Lunn *et al.*, 2009) and to generate 95% credible intervals (CI) for each parameter. Strong evidence of a relationship between variables was inferred when CIs for a parameter value did not include zero. Parameter values were estimated using a sample of >3000 iterations, after a suitable burn-in period and thinning to remove autocorrelation. We used minimally informative priors (diffuse normal distributions centred on zero for coefficients, and half-normal distributions for variance parameters, see Ntzoufras, 2011).

In the first set of analyses, all trials from the MR (N= 2088) and LT tasks (N= 1646) were analysed. We modelled the probability that the participant's initial answer would switch following social information on each trial, with participant's sex, accuracy of the initial answer, confidence in the initial answer, and proportion of demonstrators disagreeing with the participant as predictors. The effects of participant sex and accuracy on confidence were estimated using a binomial model (because confidence could only take integer values within a fixed range), while the other paths were estimated using a Bernoulli model. The shape of the relationship between the proportion of demonstrators disagreeing with the initial answer and switching was predicted to be s-shaped, because conformist transmission entails a bias towards copying the majority (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Morgan *et al.*, 2012; Rendell *et al.*, 2011), but was also allowed to be inverse s-shaped (i.e., anti-conformist) or a straight line (i.e., no bias). The models included a random participant effect to allow for the non-independence of trials within participants. The number of demonstrators was not significantly related to any other measure, so this variable was excluded from the final models.

In the second set of analyses, we performed mediation analyses (again, using Bayesian estimation in OpenBUGS) on a sub-set of trials. The effect of interest was the indirect effect of sex on switching via confidence. Our *a priori* decision to test the strength of this indirect effect was not

contingent upon finding an overall total effect (see, e.g., Hayes, 2013). Because we were interested in switching, we analysed only those trials in which the majority of demonstrators disagreed with the participant (hereafter ‘critical trials’), irrespective of the size of this majority and regardless of whether the social information was real or generated (MR: N= 1001 trials; LT: N= 722 trials). We modelled the probability that participants would switch their decisions on these critical trials using confidence as a direct predictor, accuracy of the initial answer as an indirect predictor (via confidence) and sex as both a direct predictor and indirect predictor (via the other two variables). Accuracy could not influence switching directly because participants did not receive feedback on their performance. The number of demonstrators was not included as a predictor variable, because this variable was not associated with switching in the full dataset. Indirect effects were estimated by multiplying the standardised regression coefficients of their constituent components (see MacKinnon, 2009) to create posterior distributions for the size of the indirect effects.

**Figure 1** presents the median of posterior samples from the fitted models, as estimates of parameter values, and the error bars represent CIs. Plots of posterior parameter estimates are preferred to raw data because these directly illustrate the marginal effect sizes associated with a given predictor while controlling for the effects of other predictors. **Figure 2** shows direct and indirect effects (via accuracy and confidence) of sex on switching, and positive coefficients indicate that men had higher values than women. When confidence intervals do not include zero, the standardised coefficients (i.e.,  $\beta$ ) can be interpreted in a similar way to standardised coefficients in frequentist models and represent the number of standard deviations by which the outcome variable will change if there is a 1 standard deviation increase in the predictor variable. We also present descriptive statistics in **Table 1**, and our raw data are available online in the data repository system at the University of St Andrews (<http://www.st-andrews.ac.uk/staff/research/pure/>).

## Results

Across both sexes and task types, participants were more confident when their initial answer was correct than when incorrect (MR:  $\beta = 0.84$ , CI [0.66, 1.07]; LT:  $\beta = 0.99$ , CI [0.77, 1.34]). In addition, participants deployed a copy-when-uncertain strategy: the likelihood of switching an answer following exposure to social information decreased with higher reported confidence in the initial answer, in both the MR ( $\beta = 0.69$ , CI [0.53, 0.85]; **Figure 1a**) and LT tasks ( $\beta = 0.54$ , CI [0.44, 0.64]; **Figure 1b**). Switching was also strongly and positively predicted by the proportion of demonstrators disagreeing with the participant in both the MR ( $\beta = 6.35$ , CI [5.43, 7.36]; **Figure 1a**) and LT task ( $\beta = 5.34$ , CI [4.45, 6.32]; **Figure 1b**). There was an s-shaped relationship between proportion of demonstrators disagreeing and likelihood of switching in both tasks, meaning that majorities exerted a disproportionately strong effect on the likelihood of participants switching their answers.

In the MR task, men gave a higher proportion of correct initial answers than women ( $\beta = 0.38$ , CI [0.13, 0.63]) and reported higher levels of confidence in their answers than women ( $\beta = 0.55$ , CI [0.17, 0.93]). In the LT task, there was no strong evidence that men and women differed in either their initial accuracy ( $\beta = 0.14$ , CI [-0.09, 0.46]) or confidence in their answers ( $\beta = 0.16$ , CI [-0.33, 0.72]). Participants of both sexes were more likely to switch when more of the demonstrators contradicted their initial answer in the MR task (**Figure 1c**) and LT task (**Figure 1d**). Because all participants were highly unlikely to switch their answers on trials when the majority agreed with the initial answer, overall rates of switching across all trials did not differ between the sexes in either the MR ( $\beta = -0.12$ , CI [-0.32, 0.08]) or LT task ( $\beta = -0.04$ , CI [-0.19, 0.11]). However, when the social information contradicted a participant's initial answer, women were more likely than men to switch their decision in the MR task (**Figure 1c**), and the 95% credible intervals for the sex difference in probability of switching excluded zero for this task (**Figure 1e**). There was no such divergence in the LT task (**Figure 1d**), although women were slightly less likely than men to switch when a non-total majority of demonstrators disagreed with the initial answer (**Figure 1f**).

Having found effects of sex on accuracy and confidence in the MR task, we investigated the extent to which each of these sex differences mediated indirect effects of sex on conformity on trials in which the majority of demonstrators disagreed with the participant's initial answer. For these *critical trials*, the analyses confirmed the sex difference in initial answer accuracy ( $\beta= 0.09$ , CI [0.02, 0.17]) and confidence controlling for accuracy ( $\beta= 0.17$ , CI [0.08, 0.26]) in the MR task (**Figure 2a**). The effects of accuracy on confidence ( $\beta= 0.10$ , CI [0.07, 0.14]) and confidence on switching ( $\beta= -0.77$ , CI [-0.96, -0.60]) were also upheld in this sub-set of trials (**Figure 2a**). Crucially, the analyses revealed a small indirect effect of sex on switching mediated by confidence ( $\beta= -0.14$ , CI [-0.22, -0.06]; not shown in figure); thus, lower average levels of confidence in women than men increased the likelihood of switching in female participants independent of the modelled relationship between sex and accuracy. The indirect effect of sex on switching via accuracy was negligible ( $\beta= -0.01$ , CI [-0.01, -0.01]; not shown in figure), and there was no direct effect of sex on switching ( $\beta= -0.01$ , CI [-0.24, 0.23]; **Figure 2a**). Therefore, the principal effect of participants' sex on switching in the MR task was exerted through confidence and not through accuracy.

In the LT task, the effects of accuracy of the initial answer on confidence ( $\beta= 0.19$ , CI [0.15, 0.22]) and confidence on switching ( $\beta= -0.64$ , CI [-0.78, -0.51]) were upheld in the subset of critical trials (**Figure 2b**). There was no strong evidence of a direct effect of sex on switching ( $\beta= -0.10$ , CI [-0.08, 0.29]; **Figure 2b**), nor was there strong evidence of a sex difference in accuracy ( $\beta= 0.07$ , CI [-0.01, 0.16]) or confidence ( $\beta= 0.12$ , CI [-0.01, 0.25]) (**Figure 2b**). The estimate of the indirect effects of sex on switching via confidence had a 95% credible interval that included zero ( $\beta=0.08$ , CI [-0.18/0.01]).

## Discussion

In both of our tasks, participants were more likely to change their answer in response to social information when they expressed low confidence in their initial answer. Participants were also more likely to switch answers when a high proportion of demonstrators disagreed with their initial answer than when a low proportion disagreed, consistent with previous theoretical and empirical research on conformist transmission (Morgan *et al.*, 2012). In the MR task, men had higher average scores and higher average confidence than women, which also supports previous studies (Cooke-Simpson & Voyer, 2007; Voyer, Voyer & Bryden, 1995). Importantly, the sex difference in confidence in the MR task existed independently of the modelled relationship between sex and accuracy, and underpinned an indirect effect of sex on use of social information. In the LT task, the evidence for a sex difference in confidence was weak, as was the evidence for an indirect effect of sex on conformity via confidence. These analyses provide clear evidence supporting an effect of sex on conformity mediated by confidence in a gender-stereotyped task, which suggests that sex differences in conformity could be partly explained by sex differences in confidence.

The significance of our results is that confidence is implicated as a strong predictor of susceptibility to social influence, and therefore, wherever there is evidence for a sex difference in confidence, an indirect effect of sex on social influence via confidence is likely. Although the indirect effect was relatively small, this finding has potentially wide-ranging implications given that women greatly underestimate their performance on a number of other laboratory tasks (Syzmanowicz & Furnham, 2011) and are reported to have considerably lower confidence than men in a variety of real-world settings (Jostl *et al.*, 2012). Sex differences in confidence could therefore contribute to producing sex difference in susceptibility to social influences. Based on our results, we hypothesise that where sex differences in the use of social information are found, this effect will be at least partly mediated by sex differences in confidence. The strongest test of this hypothesis would come from experimentally lowering confidence (see, e.g. Estes & Felker, 2012), for example by

providing negative feedback about prior performance, in order to produce an increased reliance on social information.

While we did not examine why women have lower confidence than men in the MR task, nor directly test for an interaction between task type and sex when predicting confidence, the MR task is strongly gender stereotyped, with people of both sexes expecting that men will outperform women (see, e.g., Hirnstein, Andrews, & Hausmann, 2014). Whether a stereotyped task that is assumed to favour women would induce lower confidence, and hence a greater reliance on social information, in men than women remains unclear. Indeed, the findings might differ from those in the current study, given that men are reported to be less susceptible than women to stereotype threat manipulations (Hirnstein, Freund & Hausmann, 2012) and more willing than women to guess answers in experimental settings (Baldiga, 2014). Verbal fluency is a frequently used cognitive task that shows a consistent female advantage (e.g., Hausmann, 2014), but the free-response formats of these tasks are less amenable to presenting participants with contradictory social information. Sex differences in conformity on female-typed tasks therefore require further investigation.

The fact that social information was under-used by both sexes also deserves further investigation. In our study, participants who consistently copied the majority would have obtained the maximum bonus, yet social information was used in fewer than half of the trials in which the majority disagreed with the initial answer. This finding is consistent with a number of studies showing that social information is commonly underused in laboratory experiments (e.g., Mesoudi, 2011; Morgan *et al.*, 2012), and over-confidence in these settings could thus contribute to this under-use of social information. Previous research has also indicated that self-estimates of accuracy tend to be inflated (e.g., Pallier, 2003) and are related to socially transmitted information about expected performance. For instance, a study of sex differences in children's confidence indicated that the gender-stereotyped beliefs held by parents partly explained boys' tendency to over-rate their intelligence (Steinmayr & Spinath, 2009). Taken together, these findings suggest that the

relationship between stereotypes and confidence, and the use of social information, is likely to be complex and multi-directional (Hausmann, 2014).

Our results suggest that, wherever sex differences in confidence exist, this will have implications for patterns of conformity. This finding is important because there is a large body of experimental data reporting that women are more likely than men to conform to the decisions and opinions of others (Bond & Smith, 1996; Cooper, 1979; Eagly & Carli, 1981). This sex difference in conformity has been attributed to evolved psychological mechanisms (Sundi *et al.*, 2006) or social norms (Eagly & Chrvala, 1986), yet conformity has been shown to be highly situation-dependent in both sexes (Griskevicius, 2006). As argued previously for other traits (e.g., Cross & Campbell, *in press*; Cross, Cyrenne, & Brown, 2013), explanations for sex differences that resort solely either to evolved dispositions or cultural norms are unlikely to be satisfactory. We therefore propose that sex differences in conformity – where they exist – can be characterized as the complex product of a generally adaptive social learning strategy and culturally mediated inputs to this strategy (such as gender stereotypes) that differ according to sex.

## References

- Baldiga, K. (2014). Gender differences in willingness to guess. *Management Science*, **60**, 434–448.  
DOI: 10.1287/mnsc.2013.1776
- Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics* **55**, 412–428. DOI: 10.3758/BF03205299
- Bardsley, N. (2000). Control without deception: individual behaviour in free-riding experiments revisited. *Experimental Economics*, **3**, 215–240. DOI: 10.1007/BF01669773

- Bardsley, N., & Sausgruber, R. (2005). Conformity and reciprocity in public good provision. *Journal of Economic Psychology*, **26**, 664–681. DOI:10.1016/j.joep.2005.02.001
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, **24**, 127–135. DOI:10.1016/j.tree.2008.10.008
- Bond, R., & Smith, P. B. (1996). Culture and conformity: a meta-analysis of studies using Asch's line judgement task. *Psychological Bulletin*, **119**, 111–137. DOI: 10.1037/0033-2909.119.1.111
- Boyd, R., & Richerson, P. J. (1988). An evolutionary model of social learning: the effects of spatial and temporal variation. In T. R. Zentall & B. G. Galef (Eds.), *Social learning: psychological and biological perspectives* (pp. 29–40). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cooke-Simpson, A., & Voyer, D. (2007). Confidence and gender differences on the mental rotations test. *Learning & Individual Differences*, **17**, 181–186. DOI: 10.1016/j.lindif.2007.03.009
- Cross, C. P., & Campbell, A. C. (in press) Integrating approaches requires more than a division of labour: Comment on Wolfer & Hewstone (2015). *Psychological Science*.
- Cross, C. P., Cyrenne, D. L. M., & Brown, G. R. (2013). Sex differences in sensation-seeking: a meta-analysis. *Scientific Reports*, **3**: 2486. DOI: 10.1038/srep02486
- Eagly, A. H., & Carli, L. L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: a meta-analysis of social influence studies. *Psychological Bulletin*, **90**, 1–20. DOI: 10.1037/0033-2909.90.1.1
- Eagly, A. H., & Chivala, C. (1986). Sex differences in conformity: status and gender role interpretations. *Psychology of Women Quarterly*, **10**, 203–220. DOI: 10.1111/j.1471-6402.1986.tb00747.x
- Estes, Z., & Felker, S. (2012). Confidence mediates the sex difference in mental rotation performance. *Archives of Sexual Behavior*, **41**, 557–570. DOI: 10.1007/s10508-011-9875-5



- Galef, B. G., Dudley, K. E., & Whiskin, E. E. (2008). Social learning of food preferences in 'dissatisfied' and 'uncertain' Norway rats. *Animal Behaviour*, **75**, 631–637.  
DOI:10.1016/j.anbehav.2007.06.024
- Goldberg, C. (1975). Conformity to majority type as a function of task and acceptance of sex-related stereotypes. *Journal of Psychology*, **89**, 25–37. DOI:  
10.1080/00223980.1975.9923903
- Griskevicius, V., Goldstein, N. J., Mortensen, C. R., Cialdini, R. B., & Kenrick, D. T. (2006). Going along versus going alone: when fundamental motives facilitate strategic (non)conformity. *Journal of Personality & Social Psychology*, **91**, 281–294. DOI:  
10.1037/0022-3514.91.2.281
- Hausmann, M. (2014). Arts versus science - academic background implicitly activates gender stereotypes on cognitive abilities with threat raising men's (but lowering women's) performance. *Intelligence*, **46**, 235-245. DOI:10.1016/j.intell.2014.07.004
- Hausmann, M., Schoofs, D., Rosenthal, H. E., & Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences – a psychobiosocial approach. *Psychoneuroendocrinology*, **34**, 389–401. DOI:10.1016/j.psyneuen.2008.09.019
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Hertwig, R., & Ortmann, A. (2008). Deception in experiments: revisiting the arguments in its defense. *Ethics*, **18**, 59-92. DOI: 10.1080/10508420701712990
- Hirstein, M., Andrews, L. C., & Hausmann, M. (2014). Gender-stereotyping and cognitive sex differences in mixed-and same-sex groups. *Archives of Sexual Behavior*, **43**, 1663-1673.  
DOI: 10.1007/s10508-014-0311-5
- Hirstein, M., Freund, N., & Hausmann, M. (2012). Gender stereotyping enhances verbal fluency performance in men (and women). *Zeitschrift für Psychologie*, **220**, 70–77. DOI:  
10.1027/2151-2604/a000098

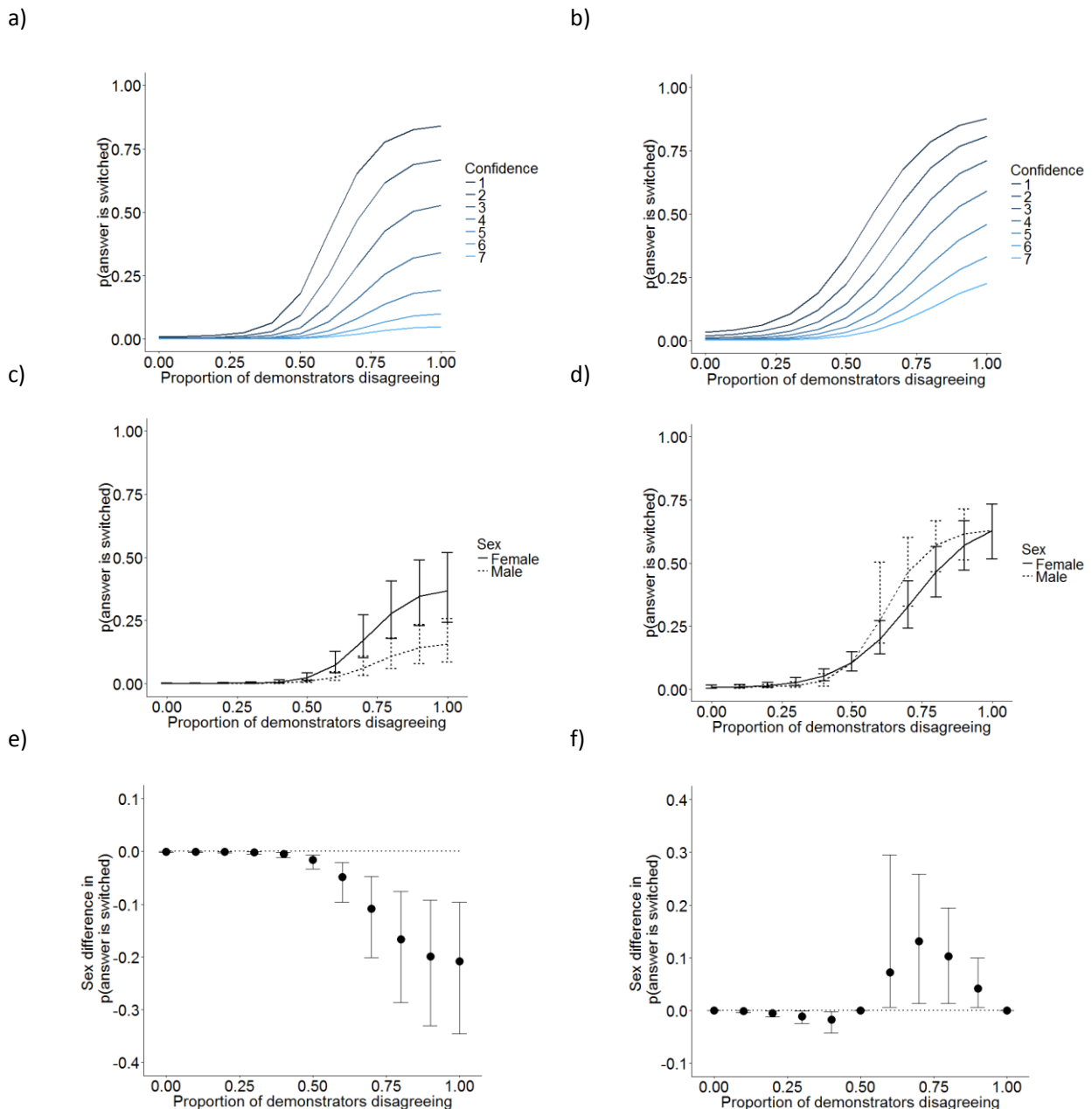
- Hoppitt, W., & Laland, K. N. (2013). *Social learning: an introduction to mechanisms, methods and models*. Princeton University Press.
- Jöstl, G., Bergsmann, E., Lüftenegger, M., Schober, B., & Spiel, C. (2012). When will they blow my cover? The impostor phenomenon among Austrian doctoral students. *Zeitschrift für Psychologie*, **220**, 109–120. DOI: 10.1027/2151-2604/a000102
- Karabenick, S. A. (1983). Sex-relevance of content and influenceability: Sistrunk and McDavid revisited. *Personality & Social Psychology Bulletin*, **9**, 243–252. DOI: 10.1177/0146167283092008
- Laland, K. N. (2004). Social learning strategies. *Learning & Behavior*, **32**, 4-14. DOI: 10.3758/BF03196002
- Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: evolution, critique and future directions. *Statistics in Medicine*, **28**, 3049–3067. DOI: 10.1002/sim.3680
- MacKinnon, D. P. (2009). *Statistical Mediation Analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Majeres, R. L. (2007). Sex differences in phonological coding: alphabet transformation speed. *Intelligence*, **35**, 335–346. DOI: 10.1016/j.intell.2006.08.005
- Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334-342. DOI: 10.1016/j.evolhumbehav.2010.12.001
- Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society, B*, **279**, 653–662. DOI: 10.1098/rspb.2011.1172
- Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS* (Vol. 698). John Wiley & Sons.
- Oswald, D. L. (2008). Gender stereotypes and women's reports of liking and ability in traditionally masculine and feminine occupations. *Psychology of Women Quarterly*, **32**, 196–203. DOI: 10.1111/j.1471-6402.2008.00424.x

- Pallier, G. (2003). Gender differences in the self-assessment of accuracy on cognitive tasks. *Sex Roles* **48**: 265-276. DOI: 10.1023/A:1022877405718
- Rendell, L., Fogarty, L., Hoppitt, W. J., Morgan, T. J., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, **15**, 68-76. DOI: 10.1016/j.tics.2010.12.002
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of 3-dimensional objects. *Science*, **171**, 701-703. DOI: 10.1126/science.171.3972.701
- Steinmayr, R. & Spinath, B. (2009). What explains boys' stronger confidence in their intelligence? *Sex Roles*, **61**: 736-749. DOI: 10.1007/s11199-009-9675-8
- Sundie, J. M., Cialdini, R. B., Griskevicius, V., & Kenrick, D. T. (2006). Evolutionary social influence. In M. Schaller, J. A. Simpson & D. T. Kenrick (Eds), *Evolution and social psychology* (pp. 287–316). New York, NY: Psychology Press.
- Syzmanowicz, A., & Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: four meta analyses. *Learning & Individual Differences*, **21**, 493-504. DOI: 10.1016/j.lindif.2011.07.001
- van Bergen, Y., Coolen, I., & Laland, K. N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proceedings of the Royal Society, B*, **271**, 957-962. DOI: 10.1098/rspb.2004.2684
- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin*, **117**, 250-270. DOI: 10.1037/0033-2909.117.2.250

**Table 1** Summary statistics from the raw data (means with SDs in brackets) for proportion of initial answers correct, confidence, proportion of answers switched in response to social information, and proportion of final answers correct for men and women in the MR and LT tasks. a = all trials; b = critical trials where the majority of the social information conflicted with the initial answer.

Task	Measure	Men	Women
MR task	Proportion of initial answers correct <sup>a</sup>	0.71 (0.14)	0.63 (0.11)
	Confidence <sup>a</sup>	4.68 (1.10)	3.90 (0.99)
	Proportion of answers switched <sup>a</sup>	0.10 (0.10)	0.12 (0.10)
	Proportion of answers switched <sup>b</sup>	0.22 (0.23)	0.27 (0.24)
	Proportion of final answers correct <sup>a</sup>	0.70 (0.09)	0.68 (0.13)
LT task	Proportion of initial answers correct <sup>a</sup>	0.67 (0.11)	0.62 (0.15)
	Confidence <sup>a</sup>	4.25 (1.13)	3.82 (1.28)
	Proportion of answers switched <sup>a</sup>	0.23 (0.12)	0.23 (0.15)
	Proportion of answers switched <sup>b</sup>	0.47 (0.28)	0.47 (0.28)
	Proportion of final answers correct <sup>a</sup>	0.69 (0.14)	0.63 (0.11)

**Figure 1** Probability that an answer is switched ( $p(\text{answer is switched})$ ) in response to social information as a function of the proportion of demonstrators disagreeing and the confidence rating (1 = low confidence; 7 = high confidence) in: a) the MR, and b) the LT task. Average probability that an answer is switched as a function of the proportion of demonstrators disagreeing for men and women separately in: c) the MR, and d) the LT task. Posterior estimates of the sex difference in probability of switching (male  $p(\text{answer is switched})$  minus female  $p(\text{answer is switched})$ ) as a function of proportion of demonstrators disagreeing in: e) the MR, and f) the LT task. Error bars indicate 95% credible intervals.



**Figure 2** Standardised regression coefficients for the relationships between participant sex, accuracy of initial answer, confidence rating and switching answers in response to social information in: a) MR, and b) LT task. Numbers in brackets are upper and lower bounds of 95% credible intervals. Solid red lines indicate standardised regression coefficients for which the CIs exclude zero (i.e., the parameter is very unlikely to be zero), and dashed grey lines indicate coefficients for which the CIs include zero.

