



BRILL



brill.com/joc

# Prestige Does Not Affect the Cultural Transmission of Novel Controversial Arguments in an Online Transmission Chain Experiment

Ángel V. Jiménez

Human Behaviour and Cultural Evolution Group (HuBCEG), Department of Biosciences, University of Exeter Cornwall Campus, Penryn, Cornwall, UK  
[aj419@exeter.ac.uk](mailto:aj419@exeter.ac.uk)

Alex Mesoudi

Human Behaviour and Cultural Evolution Group (HuBCEG), Department of Biosciences, College of Life and Environmental Science, University of Exeter Cornwall Campus, Penryn, Cornwall, UK  
[a.mesoudi@exeter.ac.uk](mailto:a.mesoudi@exeter.ac.uk)

## Abstract

Cultural evolutionary theories define *prestige* as social rank that is freely conferred on individuals possessing superior knowledge or skill, in order to gain opportunities to learn from such individuals. Consequently, information provided by prestigious individuals should be more memorable, and hence more likely to be culturally transmitted, than information from non-prestigious sources, particularly for novel, controversial arguments about which preexisting opinions are absent or weak. It has also been argued that this effect extends beyond the prestigious individual's relevant domain of expertise. We tested whether the prestige and relevance of the sources of novel, controversial arguments affected the transmission of those arguments, independently of their content. In a four-generation linear transmission chain experiment, British participants ( $N = 192$ ) recruited online read two conflicting arguments in favour of or against the replacement of textbooks by computer tablets in schools. Each of the two conflicting arguments was associated with one of three sources with different levels of prestige and relevance (high prestige, high relevance; high prestige, low relevance; low prestige, low relevance). Participants recalled the pro-tablets and anti-tablets arguments associated with each source and their recall was then passed to the next participant within their chain. Contrary to our predictions, we did not find a reliable effect of either the prestige or relevance of the sources of information on transmission fidelity. We discuss whether

the lack of a reliable effect of prestige on recall might be a consequence of differences between how prestige operates in this experiment and in everyday life.

## Keywords

prestige bias – social status – transmission biases – cultural evolution – transmission chain – social psychology

## 1 Introduction

The diffusion of knowledge, skills, attitudes, norms and other forms of information within societies via cultural transmission/social learning depends not only on the content of that information, but also on the characteristics of the source of the information. For example, when Hollywood actor Angelina Jolie wrote a newspaper article about her decision to undergo a preventive double mastectomy due to testing positive for a gene associated with breast cancer (Jolie, 2013), there followed an increase in online searches for information about breast cancer (Juthe, Zaharchuk, & Wang, 2015), increased demand for genetic screening of this disease (Desai & Jena, 2016), and an increase in the number of referrals to undergo similar preventive operations (Evans et al., 2015). While the content of Jolie's article surely played some role, it seems plausible that her prominence and success as an actor also influenced the subsequent spread of knowledge regarding breast cancer screening.

This distinction between the content and the source of culturally transmitted information has received much attention in the field of cultural evolution, which seeks to provide general explanations for cultural change that are rooted in both psychology and evolutionary theory (Henrich, 2015; Mesoudi, 2011). Our focus here is on the latter: the influence of the source. Cultural evolution researchers call this *model-based transmission bias* (Henrich & McElreath, 2003; Wood, Kendal, & Flynn, 2013), referring to the transmission advantage of information provided by models with specific characteristics such as being successful, older, genetically related, or prestigious. Such biases are typically evolutionarily adaptive. For example, learning from models who are successful (e.g. knowledgeable or skillful) within a valued domain generally leads to the acquisition of fitness-enhancing knowledge/skills. Therefore, preferentially copying successful models (*success bias*) is typically an adaptive strategy when the knowledge or skill cannot easily be acquired on one's own, via asocial learning.

However, assessing the knowledge or skill of potential models within a domain is often difficult. The learner may not have enough expertise to correctly infer the quality of others' information. Even if they do, this inference may be costly, for example requiring lengthy observation of the model to ensure their success is not due to luck. To reduce these costs, Henrich and Gil-White (2001) suggested that social learners use short-cuts to infer the success of models from whom to learn. These shortcuts might involve using fixed characteristics of models such as job titles (*first-order cues*) or behaviours displayed by other individuals towards the model, such as the amount of attention paid by others to the model (*second-order cues*) (Jiménez & Mesoudi, 2019). The use of these shortcuts is known as *prestige bias* or *prestige-biased social learning* (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019). Prestige bias is adaptive (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019) because prestige (e.g. amount of attention received by a model) is associated with high knowledge or skill within valued domains (Garfield, Hubbard, & Hagen, 2019; Henrich & Gil-White, 2001; StibbardHawkes, Attenborough, & Marlowe, 2018; Von Rueden, 2014; although see Reyes-Garcia et al., 2008). If there is a positive relationship between the success and prestige of a model, then using prestige cues to select models from whom to learn (prestige bias) is a cheaper way to acquire mostly valuable knowledge/skills than assessing directly the competence of the model (success bias).

Our aim in this study is to experimentally test whether information attributed to prestigious sources is culturally transmitted with greater fidelity than information attributed to nonprestigious sources, as predicted by Henrich and Gil-White (2001). We use the transmission chain design (Bartlett, 1932; Mesoudi, 2007) in which participants are asked to recall written information, and the resulting recall is passed to the next participant to read and recall and so on down the chain. This can reveal cumulative effects of hypothesized transmission biases. The information transmitted along the chains consists of opposing arguments about a novel and controversial issue: the replacement of textbooks by computer tablets at primary school. We were not interested in the issue itself (i.e. the content), only in the effects of prestige on its transmission (i.e. the source). We chose this issue because, although it engenders some debate (e.g. Kleeman, 2014; Rowan, 2014), it is an issue that is not widely discussed and about which people lack polarized attitudes or detailed prior knowledge. Therefore, people's attitudes towards tablets are unlikely to strongly affect the transmission of these arguments, minimising the influence of both content biases and individual judgement.

Previous research provides some evidence that prestige affects both recall in a single individual, and cultural transmission from one individual to

another. Holtgraves, Srull, and Socall (1989) had participants read conversations between two people. One group of participants were told that one of the speakers was higher status than the other. Another group were told that the two speakers had equal status. As predicted, statements were more likely to be recalled when they were thought to come from a high status speaker, compared to the same remarks made in the equal status condition. However, this effect only occurred when the status information was provided before reading the conversations; no effect was found when status information was provided afterwards. There was also no effect of status when the conversations were acted out, rather than read. Ratcliff, Hugenberg, Shriver, and Bernstein (2011), meanwhile, found that participants were better at recognising faces associated with prestigious professions (e.g. CEOs or doctors) than faces associated with non-prestigious professions (e.g. mechanics or plumbers). However, this study did not address whether information provided by prestigious individuals was more memorable or how prestige affected the cultural transmission of information.

Other studies have found evidence for prestige-biased cultural transmission. Atkisson, O'Brien, and Mesoudi (2012) found that participants preferentially copied virtual artifact designs that purported to come from models who received more attention from other participants. These prestige cues were even used as equally often as direct success information. Chudek, Heller, Birch, and Henrich (2012) similarly found that children preferentially copied the choice of toy or food from an adult model to whom others had attended to, compared to models who were ignored by others. While suggestive, these studies did not examine the long-term effect of prestige along chains of participants, as afforded by the transmission chain method, nor did they examine the transmission of controversial arguments, as we did here. The latter is particularly important because of cases such as Angelina Jolie's endorsement of breast cancer screening, or (more troublingly) the endorsement of anti-vaccination arguments by celebrities such as Jenny McCarthy, where prestigious individuals appear to influence the adoption of controversial or unfamiliar arguments.

This leads to our secondary aim, which is to explore whether the effect of prestige on cultural transmission extends beyond the domain of expertise of the information source. This seems to occur when, for example, an actor has an effect in a medical domain like breast screening or vaccination. Although prestige bias should be more beneficial within the domain of expertise of the prestigious model (*within-domain prestige bias*), the theory (Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019) assumes that people should copy prestigious individuals beyond their domain of expertise (*cross-domain prestige bias*):

(i) when success information in the domain of interest is unavailable and asocial learning in that domain is costly or difficult (general copying bias) and (ii) when there is a positive correlation in skill between domains due to general characteristics such as IQ or perseverance that lead the model to be successful across multiple domains (cross-domain general ability). The first of these, at least, would apply to controversial arguments for which people have little pre-existing knowledge.

However, the evidence regarding cross-domain prestige bias is mixed. In the aforementioned study by Chudek et al. (2012), the effect of prestige did not extend beyond the domain within which the model expressed a choice. For example, a prestigious model's choice of food did not affect whether their choice of toy was also copied. Cases such as the 'Angelina Jolie effect' remain anecdotal and ambiguous, given the difficulty of using observational data to determine cause-effect relationships. When Jolie wrote her article, there was already growing interest in genetic screening for cancer, and guidelines about detecting breast cancer had just been published by public health organisations (Acerbi, 2019). Moreover, it is difficult to separate the effect of the content of the information she provided from the effect of her prestige. Given this ambiguous but limited evidence, there is a need for more experimental tests of the domain generality of prestige.

Consequently, we sought to test whether the influence of prestige on cultural transmission only occurs in the relevant domain of expertise of the prestigious model, or whether it extends beyond the relevant domain. We therefore used three sources that varied in prestige and relevance: a relevant high prestige source (Head of the Department of Education of a leading university), an irrelevant high prestige source (aircraft pilot), and an irrelevant low prestige source (cleaner). Prestige was manipulated using job titles (a first-order cue), following previous non-transmission chain experiments (Dalmaso, Galfano, Coricelli, & Castelli, 2014; Dalmaso, Pavan, Castelli, & Galfano, 2012; Ratcliff et al., 2011). This contrasts with typical manipulations of prestige in the cultural evolution literature, which usually entails a second-order cue: the amount of attention displayed by other individuals to the model (Atkisson et al., 2012; Chudek, Baron, & Birch, 2016; Chudek et al., 2012). Nevertheless, both types of cues should be positively correlated given that individuals with prestigious job titles (e.g. doctors) tend to be paid more attention by others (Dalmaso et al., 2014; Dalmaso et al., 2012; Ratcliff et al., 2011), due to both cues being used as proxies for competence.

In light of the above theory and evidence, we preregistered (<https://osf.io/pk2rz/>) the following two hypotheses:

H1: The arguments provided by high prestige sources will be better recalled over generations than arguments provided by low prestige sources.

H2: The arguments provided by high relevance sources will be better recalled over generations than the arguments provided by low relevance sources.

## 2 Methods

### 2.1 *Ethical Statement*

The study was approved by the Biosciences Ethical Committee at the University of Exeter Cornwall Campus on the 6th June 2017 (Ref 2017/1963).

### 2.2 *Participants*

Participants were recruited online through Prolific ([www.prolific.ac](http://www.prolific.ac)) (Peer, Samat, Brandimarte, & Acquisti, 2016) following the procedure stated in the preregistration (<https://osf.io/pk2rz/>). We used pre-screening filters to select participants who had not previously participated in any of our studies on Prolific, had an approval rating of 90% or above, were aged 18–60 years, spoke English as a first language and had British nationality.

The data was collected at four different times (one for each “cultural generation”), with 48 participants per generation across 48 separate, parallel chains. 48 chains were necessary to provide two replications of each of the 24 counter-balanced versions of the experimental materials (see Section 2.4). Participants were paid at a rate of £5.25/hour for an estimated time of completion of 20 minutes for generations 1 and 2 and 15 minutes for generations 3 and 4.

We excluded from the dataset participants who (i) failed the attentional check (“If you are carefully reading the questions, select ‘Green’”), or (ii) failed the manipulation check to identify the professions (i.e. manipulation of prestige) of the sources of the information, or (iii) read both sets of arguments (pro- and anti-tablets) in less than 10 seconds (1080 words/minute). The latter exclusion criterion was changed compared to the preregistered protocol, which originally specified a cutoff of 27 seconds (400 words/minute). After seeing the data for the first wave of participants it became apparent that the original cutoff was too strict and unnecessarily excluded large numbers of participants who had provided valid data. The new cutoff was derived from inspection of the data: three participants who read the material in less than 10 seconds (the new cutoff) were unable to recall anything, while amongst the

remaining participants there was no relationship between time spent reading the materials and the number of propositions correctly recalled (Bayesian multilevel Poisson logistic model predicting the number of propositions correctly recalled by the time spent reading each of the materials associated with each source and intercepts varying by participant:  $B = 0.001$ ,  $SE = 0.003$ , 89% CI  $[-0.003, 0.005]$ ), confirming that these participants provided valid data.

Excluded participants were replaced by new participants. Overall, we collected responses from 288 participants (203 females, 85 males) aged 18–61<sup>1</sup> ( $M = 34.57$ ,  $SD = 10.12$ ), with 192 of these participants (143 females, 49 males) aged 18–61 ( $M = 35.10$ ,  $SD = 10.11$ ) included in the study.

### 2.3 Materials

To select the sources of information, we asked an independent sample of 10 Prolific participants (5 females, 5 males) aged 19–42 ( $M = 28.8$ ,  $SD = 7.16$ ) to rate 24 professions on their prestige within society and their relevance for the debate over whether tablets should replace books in schools. Ratings for prestige were given in a 5-point Likert scale from “non-prestigious at all” (1) to “extremely prestigious”. Ratings for relevance were given in a 7-point Likert scale from “extremely irrelevant” (-3) to “extremely relevant” (+3), respectively. We selected three sources who differed in their level of prestige and relevance: the Head of Education of a leading university (high prestige:  $M = 1.6$ ,  $SD = 1.1$ ; high relevance:  $M = 2.4$ ,  $SD = 0.8$ ), an aircraft pilot (high prestige:  $M = 1.5$ ,  $SD = 1.2$ ; low relevance:  $M = -2.1$ ,  $SD = 1.3$ ) and a cleaner (low prestige:  $M = -2.6$ ,  $SD = 1$ ; and low relevance:  $M = -2.2$ ,  $SD = 1.2$ ). See SM1A for a full list of professions and their associated ratings.<sup>2</sup>

In the experimental materials, the sources of information were associated with one of two sets of arguments in favour of and against the replacement of textbooks by computer tablets at primary school. Two different names

1 Although in the pre-screening we selected an age range of 18–60, 1 participant was 61 years old in the experiment. This was probably a natural consequence of the pass of time, since participants reported their age for the pre-screening prior the participation in the study.

2 This procedure to select professions to use in the experiment on the basis of the means of ratings on a Likert scale is a standard practice in psychology. However, this practice contradicts what students are taught about ordinal variables in statistical courses. Ordinal variables are categorical variables that are ordered but the distances between adjacent categories might be different. Consequently, this precludes carrying out mathematical operations such as addition, subtraction, multiplication, and divisions with them to obtain means and standard deviations. As we became more aware of the consequences of treating ordinal variables as continuous (Liddell & Kruschke, 2018) after data collection, we report descriptive statistics in SM1B that are more adequate for ordinal variables, such as medians.



(“William Healey”<sup>3</sup> and “Daniel Bryanson”) were used for the sources, as each participant read the information about two sources and their associated arguments. Their names and the order of presentation were fully counterbalanced. The sources were described as follows:

**Head of Education of a leading university (high prestige, high relevance)**

*William Healey [Daniel Bryanson] is the Head of the Department of Education of a leading university. Outside of his job, he [also] volunteers for the Australian Learning Trust. As a volunteer, Professor Healey [Bryanson] visits schools once a fortnight and teaches the children the importance of his job for society. He firmly supports the replacement of textbooks by computer tablets at primary schools for the following reasons: [He is firmly against the replacement of textbooks by computer tablets at primary school for the following reasons:]*

**Aircraft pilot (high prestige, low relevance)**

*William Healey [Daniel Bryanson] is a commercial Aircraft pilot who regularly flies between Auckland and Madrid, Paris, Amsterdam and other European destinations. Outside of his job, he [also] volunteers for the Australian Learning Trust. As a volunteer, Mr Healey [Bryanson] visits schools once a fortnight and teaches the children the importance of his job for society. He firmly supports the replacement of textbooks by computer tablets at primary schools for the following reasons: [He is firmly against the replacement of textbooks by computer tablets at primary school for the following reasons:]*

**Cleaner (low prestige, low relevance)**

*William Healey [Daniel Bryanson] works as a cleaner in a telemarketing company. Outside of his job, he volunteers for the Australian Learning Trust. As a volunteer, Mr. Healey [Mr Bryanson] visits schools once a fortnight and teaches the children about the importance of his job for society. He firmly supports the replacement of textbooks by computer tablets at primary school for the following reasons: [He is firmly against the replacement of textbooks by computer tablets at primary school for the following reasons:]*

Table 1 shows the full list of arguments in favour of (‘pro-tablets’) and against (‘anti-tablets’) the replacement of textbooks by computer tablets at primary school. ‘Pro-tablets’ arguments pointed out both the advantages of tablets (e.g.

3 In the preregistration, we used the name “William Yellowfield”. When we were piloting the experiment, a colleague commented that this name was obviously made-up. Therefore, we decided to change it to William Healey for the experiment.



TABLE 1 Arguments in favour ('pro-tablets') and against ('anti-tablets') the replacement of textbooks by computer tablets at primary schools used as materials to transmit along the transmission chains

'Pro-tablets' arguments	'Anti-tablets' arguments
<i>The continuous carrying of textbooks from home to school gives long-term back pain to our children.</i>	<i>The continuous use of devices with screens such as computer tablets gives long-term vision problems to our children.</i>
<i>Children using computer tablets learn much better as they are more engaged and understand the material better.</i>	<i>Children using textbooks learn much better as they are more engaged and understand the material better.</i>
<i>Teachers have less control over children's learning when using textbooks because they cannot effectively manage what children are doing during the class.</i>	<i>Children are more easily distracted when using computer tablets because they can play games instead of attending to the lesson.</i>
<i>The production of textbooks for our schools requires the consumption of tons of paper each year, which contributes to the problem of deforestation.</i>	<i>The production of computer tablets requires the emission of a considerable amount of pollutants to the air, which contributes to the problem of the greenhouse effect.</i>
<i>Children need less support from their parents when they use computer tablets than when they use textbooks because tablets offer personalized lessons.</i>	<i>Children receive more support from their parents when they use textbooks than when they use computer tablets because parents offer personalized help.</i>
<i>Computer tablets permit the storage of hundreds of books and the instant access to those books from everywhere.</i>	<i>Textbooks can last hundreds of years and they don't require electricity or batteries to work.</i>

“Computer tablets permit the storage of hundreds of books and the instant access to those books from everywhere”) and the disadvantages of books (e.g. “The continuous carrying of textbooks from home to school gives long-term back pain to our children”). ‘Anti-tablets’ arguments pointed out both the advantages of books (e.g. “Textbooks can last hundreds of years and they don’t require electricity or batteries to work”) and the disadvantages of tablets (e.g. “The continuous use of devices with screens such as computer tablets gives long-term vision problems to our children”). Both sets of arguments were created with the explicit intention of being balanced and covering similar topics. Both sets included arguments about the impact of computer tablets and textbooks on health, the environment, children’s learning, student distraction,

control by the teacher, parental support and technical aspects such as storage, accessibility, durability and battery life. Both sets of arguments were inspired by the pros and cons arguments compiled by ProCon.org (<https://tablets-textbooks.procon.org/>).

#### 2.4 *Design*

A 3 (pair of sources: Cleaner vs Head of the Department of Education, Cleaner vs Aircraft Pilot, Head of the Department of Education vs Aircraft Pilot)  $\times$  2 (opinions associated with sources: 'pro-tablets' vs 'antitables', 'antitables vs protables') factorial design was used for this experiment. The order of presentation and the names associated with the sources ('William Healey' and 'Daniel Bryanson') were fully counterbalanced, which resulted in 24 different versions of the experimental materials. As we were only interested in the influence of the sources of the information on transmission, we analyse three experimental conditions: Condition 1 (Head of the Department of Education vs Cleaner), Condition 2 (Aircraft Pilot vs Cleaner) and Condition 3 (Head of the Department of Education vs Aircraft Pilot).

We used the transmission chain method (Bartlett, 1932; Mesoudi, 2007) to experimentally simulate this controversy. This method is similar to the children's game "Chinese Whispers" or "Broken Telephone". A first "generation" of participants read the original materials given above in Section 2.3. They then recalled this material from memory. Their recall is then transmitted to the next participant in their chain (second generation). This procedure is repeated for four generations in total, and across 48 separate, parallel chains to provide independent replications of transmission effects. This method has typically been used to study *content transmission biases*, which entail a transmission advantage to certain types of information due to their intrinsic characteristics (Henrich & McElreath, 2003) such as having a more emotional (Eriksson & Coultas, 2014; Stubbersfield, Tehrani, & Flynn, 2017), social (Mesoudi, Whiten, & Dunbar, 2006; Stubbersfield, Tehrani, & Flynn, 2014), negative (Bebbington, MacLeod, Ellison, & Fay, 2017) or stereotypical (Bangerter, 2000; Kashima, 2000) content. The novelty of the present study is to use this method to study a model-based transmission bias: prestige bias. Transmission chain experiments permit the study of the consequences of experimental conditions over multiple transmission events. Single-generation experiments sometimes cannot detect certain transmission biases, as the effects of the experimental conditions are only revealed after multiple transmission events (e.g. Jiménez, Stubbersfield, & Tehrani, 2018), or they are detected in the first generation and then reversed in latter generations (e.g. Kashima, 2000).

In our experiment, we kept the description of the sources of information (i.e. names and job titles) constant along the chains, i.e. these were not subject

to participant recall. This ensured that our manipulation was applied across all generations and all chains, and simulated the fact that prestigious individuals tend to be acknowledged and recognised by most members of a society or social group. The content of the information (the arguments in favour of and against the replacement of textbooks by computer tablets at primary school) was allowed to vary as the information provided to participants in generations 2–4 was the information recalled by the participant in the previous generation within their chain.

### 2.5 *Procedure*

Prior to the presentation of the experimental materials, participants were asked their opinion about the replacement of textbooks by computer tablets at primary school ('pre-test' opinion) by rating their agreement with the statement "If the decision were in my hands, I would replace all the textbooks with computer tablets in primary schools" on a 7-point Likert scale from "totally disagree" (-3) to "totally agree" (+3). They were also asked about their familiarity with computer tablets using a 7-point Likert scale from "very unfamiliar" (-3) to "very familiar" (+3).

Participants were then randomly assigned to one of 48 transmission chains and provided the following instructions:

In many schools across the world, computer tablets have started to replace traditional textbooks. This recent change has given rise to a heated debate about the benefits and risks of computer tablets and textbooks for children's education. We have asked a number of volunteers at schools in Australia about their views on this issue. In the following, you will learn about two of these volunteers and about their opinions about the use of textbooks and computer tablets at primary school. It is very important that you read the information at a pace that allows you full comprehension as you will be asked some questions about this information later.

Participants then read information about one of the sources (e.g. cleaner) and one set of arguments (e.g. 'pro-tablets') and, immediately afterwards, information about another source (e.g. Head of the Department of Education) and the other set of arguments (e.g. 'anti-tablets'). Participants in generation 1 read the original arguments created by us. Participants in generations 2–4 read the arguments as they were recalled by the participant in the previous generation within their chain. Spelling and grammar mistakes were corrected before transmitting the information from one participant to the next.

Participants were then asked to identify the source professions they had read from a list of six professions (Head of the Department of Infectious Diseases of

a leading university, cleaner, Head of the Department of Education of a leading university, writer, warehouseman, aircraft pilot and taxi driver) and to rate the prestige within society of the two sources they had read about on a 5-point Likert scale from “not prestigious at all” (1) to “incredibly prestigious” (5) and their relevance for the debate about the benefits and risks of tablets and books at schools on a 7-point Likert scale from “very irrelevant” (-3) to “very relevant” (+3).

There was then a surprise free recall test in which participants had to recall the arguments provided by each source. We originally intended to force participants to do this for each source for at least 2.5 minutes and no more than 5 minutes. However, a technical problem disabled the feature that forced participants to stay on the task at least 2.5 minutes. Therefore, participants submitted their responses when they considered they had finished, which is the standard procedure for transmission chain experiments. The feature that prevented participants to complete the recall task beyond 5 minutes worked correctly.

Finally, participants provided demographic details (age, gender, nationality, first language, nearest city, and profession) and were asked again about their opinion about whether computer tablets should replace books at primary schools by rating their agreement with the statement “If I were an education policy maker I would replace textbooks by computer tablets all over the country” on a 7-point Likert scale from “totally disagree” (-3) to “totally agree” (+3). They were also asked to provide reasons to support one or the other side by writing their own opinion within a textbox. At the end, participants were informed about the goal of the experiment and our hypotheses.

## 2.6 *Coding and Data Analyses*

Participants’ recall accuracy was assessed by comparing their correct recall with a preregistered table (<https://osf.io/6d5ga/>) containing twelve central propositions, i.e. verbs, adjectives or other relational terms followed by complementary nouns, which capture the core meaning of the sentence (Kintsch, 1974). For instance, the core meaning of “the continuous carrying of textbooks from home to school gives long-term back pain to our children” is “textbooks give back pain (to children)”. This table specified which elements were important (central propositions) to recall within each sentence. The recall was coded by the first author. A second coder, who was blind to the hypothesis, coded 12.5% of the chains (6 chains). We found a high inter-coder reliability between both coders (Cohen’s Kappa = 0.84).

All statistical analyses were conducted with the Bayesian package *brms* (Bürkner, 2017) in R 3.5.3 (R Core Team, 2019). This was a change from the preregistered script (<https://osf.io/dt2uq/>), in which we did the analyses with

dummy data with the package *lme4* (Bates, Maechler, Bolker, & Walker, 2015). The change was due to the advantage of *brms* in handling ordinal predictors (Bürkner & Charpentier, 2018). This change did not affect the qualitative conclusions derived from the results (see SM2 and SM3).

### 3 Results

#### 3.1 Manipulation Checks: Prestige and Relevance of the Sources

As the ratings of prestige and relevance were measured in an ordinal scale, we used the median, range, and frequency of each rating to describe the central tendency and the distribution of the ratings for each source (Figures 1 and 2 and SM4). For the same reason, we conducted Bayesian multilevel ordinal logistic models with source as a predictor. Because each participant rated the prestige and relevance of two sources of information, the intercepts were allowed to vary by participant. For these analyses, we used the default priors in *brms* (Bürkner, 2017).

We expected that the Head of the Department of Education (henceforth ‘educator’) and the pilot would be rated similarly prestigious, and both rated more prestigious than the cleaner.

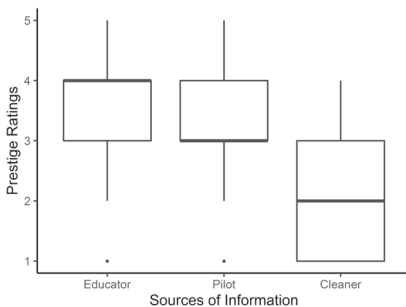


FIGURE 1

The distribution of ratings of prestige for the three sources of information: Head of the Department of Education (“educator”), pilot and cleaner. Prestige was measured on a 5-point scale from “Not prestigious at all” (1) to “incredibly prestigious” (5). Thick lines represent the median, the box the middle 50% of ratings or inter-quantile range, the lower whisker ranges from the 25th percentile to the smallest value no smaller than 1.5 times the inter-quantile range and outliers are ratings outside 1.5 times the inter-quantile range.

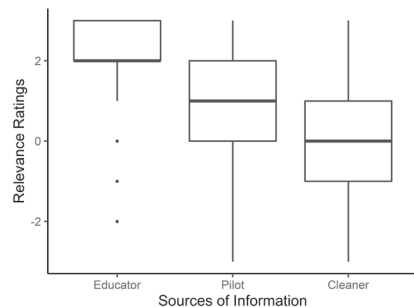


FIGURE 2

The distribution of ratings of relevance for the three sources of information: Head of the Department of Education (“educator”), pilot and cleaner. Relevance was measured on a 7-point Likert scale from “very irrelevant” (-3) to “very relevant” (+3). The thick line represents the median, the box the middle 50% of ratings or inter-quantile range, the lower whisker ranges from the 25th percentile to the smallest value no smaller than 1.5 times the inter-quantile range and outliers are ratings outside 1.5 times the inter-quantile range.

Prestige was rated on a 5-point Likert scale from “not prestigious at all” (1) to “incredibly prestigious” (5). The educator was considered the most prestigious (Median = 4), being mostly rated as “(4) very prestigious” (50%), followed by “(3) somewhat prestigious” (37%). The pilot was seen as less prestigious than the educator (Median = 3), being mostly rated as “(3) somewhat prestigious” (58%) followed by “(4) very prestigious” (32%). The cleaner was seen as the least prestigious (Median = 2), being rated mostly as “(2) hardly prestigious” (40%), followed by “(3) somewhat prestigious” (30%) and “(1) not prestigious at all” (27%). As expected, the regression model confirmed that the educator ( $B = 5.26$ ,  $SE = 0.51$ , 89% CI: [4.46, 6.11]) and the pilot ( $B = 3.95$ ,  $SE = 0.42$ , 89% CI: [3.29, 4.62]) were rated as more prestigious than the cleaner.

Contrary to expectations, however, the educator was rated as more prestigious than the pilot ( $B = 1.31$ ,  $SE = 0.32$ , 89% CI: [0.81, 1.82]), although the magnitude of this difference was much smaller than that between the educator and cleaner, and between pilot and cleaner.

We also expected that the educator would be rated more relevant to the issue of tablets in schools than both the pilot and cleaner, who would be rated similarly (non) relevant. Relevance was rated on a 7-point Likert scale from “very irrelevant” (-3) to “very relevant” (+3). The educator was considered most relevant (Median = 2), being mainly rated as +3 (46%). The pilot was considered the next most relevant (Median = 1), being mostly rated as +1 (33%). The cleaner was considered the least relevant (Median = 0), being mainly rated as +1 (30%). As expected, the regression model confirmed that the pilot ( $B = -3.08$ ,  $SE = 0.36$ , 89% CI: [-3.65, -2.51]) and the cleaner ( $B = -3.40$ ,  $SE = 0.35$ , -89% CI: [-3.96, -2.85]) were considered less relevant sources of information than the educator, while there was little difference between the pilot and the cleaner ( $B = 0.31$ ,  $SE = 0.25$ , 89% CI: [-0.08, 0.71]).

### 3.2 Cumulative Recall

Figure 3 shows the recall of correct central propositions across cultural generations by opinion (‘pro-tablets’ vs ‘anti-tablets’, Figure 3A) and source of the information (educator, pilot and cleaner, Figure 3B). As in similar transmission chain experiments, Figure 3 shows that recall decreased over cultural generations, with a larger decrease from generation 1 to generation 2 than for subsequent generations. Unexpectedly, the ‘anti-tablets’ opinion seems to have been better transmitted than the ‘pro-tablets’ opinion (Figure 3A). Contrary to our hypotheses, Figure 3B suggests little difference in cumulative recall between the three sources of information.

To statistically analyse these trends, we produced several Bayesian multilevel Poisson regression models with intercepts varying by chain and compared

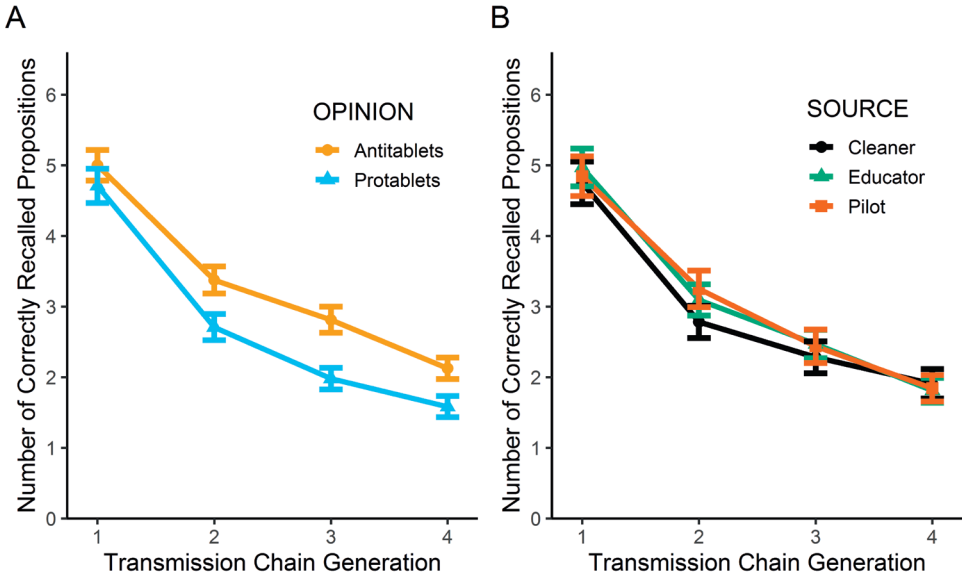


FIGURE 3 Raw means for the recall of information with 1.6σ standard error bars (corresponding to 89% CI as given in the text) plotted against generation. A: the ‘anti-tablets’ arguments were better transmitted than the ‘protatables’ arguments. B: contrary to H<sub>1</sub> and H<sub>2</sub>, the information provided by the three sources (Head of the Department of Education, Pilot and Cleaner) was similarly transmitted.

their model fit (Table 2). Model fit was compared using leave-one-out cross validation information criterion (LOOIC; Vehtari, Gelman, & Gabry, 2017), which can be interpreted similarly to Akaike information criterion (AIC) or Watanabe-Akaike information criterion (WAIC): a lower LOOIC indicates better fit to the data.

We first ran a generation model with generation (one to four) as the sole predictor of recall (number of correctly recalled propositions). Generation was treated as a monotonic variable (Bürkner & Charpentier, 2018) as recall tends to decrease across generations but the amount of the decrease varies between adjacent generations. This model was compared with an opinion model, which included opinion (‘pro-tablets’ vs ‘anti-tablets’) and generation as predictors. As the opinion model (LOOIC = 1301.2, SE = 14.4) had a better fit to the data than the generation model (LOOIC = 1311.1, SE = 14.3), the opinion model was selected as the control model to compare against subsequent models. (Allowing an interaction between generation and opinion did not improve the model fit by much, LOOIC = 1298.5, SE = 14.3. Including participants’ degree of agreement with the statement “I would replace all the textbooks with computer tablets in



TABLE 2 Unstandardized Coefficients (B) and their standard errors (in brackets) for each of the main regression models. Square brackets indicate reference categories. Variables modelled as monotonic effects are labelled mo(variable). More regression models and further details can be found in the SM3. LOOIC = leave-one-out cross validation information criterion (lower values indicate better fit to the data; see text for details). Model weights were calculated using pseudo-Bayesian model average weights with Bayesian bootstrap (Vehtari & Gabry, 2019; Yao, Vehtari, Simpson, & Gelman, 2018) with the loo package (Vehtari, Gabry, & Gelman, 2019).

Unstandardized Coefficients	Main Regression Models					
	Generation B (SE)	Opinion B (SE)	Source B (SE)	Prestige B (SE)	Relevance B (SE)	Prestige Relevance B (SE)
Intercept	1.56 (0.06)	1.65 (0.06)	1.60 (0.08)	1.62 (0.09)	1.63 (0.12)	1.61 (0.12)
mo(Generation)	-0.95 (0.09)	-0.95 (0.09)	-0.95 (0.09)	-0.95 (0.09)	-0.95 (0.09)	-0.95 (0.09)
Opinion [Pro-tablets]		-0.19 (0.06)	-0.19 (0.06)	-0.19 (0.06)	-0.19 (0.06)	-0.19 (0.06)
Source [Educator]			0.06 (0.08)			
Source [Pilot]			0.06 (0.08)			
mo(Prestige)				0.05 (0.14)		0.05 (0.15)
mo(Relevance)					0.03 (0.13)	0.01 (0.13)
<b>LOOIC</b>	1311.1	1301.2	1302.5	1302.2	1302.6	1303.8
<b>model weights</b>	0.002	0.329	0.208	0.205	0.167	0.089

primary schools" prior to being exposed to the experimental materials also did not improve the model fit. See SM3 for details).

To test H<sub>1</sub>, which stated that the information provided by high prestige sources of information (educator and pilot) would be better recalled than the information provided by the low prestige source (cleaner), we ran a model ('source model') which included source of information as a predictor of recall (with cleaner as reference category) in addition to generation and opinion.

Consistent with hypothesis H<sub>1</sub>, the information provided by the prestigious sources, the educator ( $M = 3.09$ ,  $SD = 1.70$ ,  $B = 0.06$ ,  $SE = 0.08$ , 89% CI  $[-0.07, 0.18]$ ) and the pilot ( $M = 3.09$ ,  $SD = 1.77$ ,  $B = 0.06$ ,  $SE = 0.08$ , 89% CI  $[-0.06, 0.19]$ ), were better recalled than the less prestigious cleaner ( $M = 2.93$ ,  $SD = 1.75$ ). However, the credible intervals for both comparisons include 0, indicating an unreliable effect of source of information on recall. Furthermore, the fit of the source model (LOOIC = 1302.5, SE = 14.5) was worse than the fit of the control model (LOOIC = 1301.2, SE = 14.4). Allowing an interaction between source and generation did not improve model fit (LOOIC = 1305.5, SE = 14.5). Consequently, the results did not support H<sub>1</sub>.

To test H<sub>2</sub>, which stated that the information provided by the high relevance source (educator) would be better recalled than the information provided by the low relevance sources (pilot and cleaner), we refitted the source model with educator as the reference category. Contrary to H<sub>2</sub>, the information provided by the educator was not better recalled than the information provided by the pilot ( $B = 0.01$ ,  $SE = 0.07$ , 89% CI  $[-0.11, 0.12]$ ) and the cleaner ( $B = -0.06$ ,  $SE = 0.08$ , 89% CI  $[-0.18, 0.07]$ ). Consequently, the results did not support H<sub>2</sub>.

An alternative way to test H<sub>1</sub> and H<sub>2</sub> is to use the participants' own ratings of prestige and relevance as predictors of recall, instead of assuming based on the independent raters' judgements (see Section 2.3) that both the pilot and the educator are high prestige sources (with the educator also being a relevant high prestige source) and the cleaner a low prestige source.

As prestige and relevance were ordinal measures, we modelled them as monotonic variables (Bürkner & Charpentier, 2018): the effects of prestige and relevance on recall should increase or decrease with higher ratings but the effect might vary between adjacent ratings. In addition to generation and opinion, these models included ratings of prestige (prestige model: LOOIC = 1302.2, SE = 14.5), ratings of relevance (relevance model: LOOIC = 1302.6, SE = 14.5) or ratings of both prestige and relevance (prestige-relevance model: LOOIC = 1303.8, SE = 14.5). Marginal effects of the latter prestige-relevance model are shown in Figure 4. In these models, the effects of prestige and relevance were in the expected direction: prestige (prestige model:  $B = 0.05$ ,  $SE = 0.14$ , 89% CI  $[-0.17, 0.27]$ ) and relevance (relevance model:  $B = 0.03$ ,  $SE = 0.13$ , 89% CI  $[-0.17, 0.24]$ ) both positively predicted recall in the prestige model and relevance model respectively, while for the prestige-relevance model, the effect of relevance ( $B = 0.01$ ,  $SE = 0.13$ , 89% CI  $[-0.19, 0.23]$ ) diminished after accounting for prestige ( $B = 0.05$ ,  $SE = 0.15$ , 89% CI  $[0.17, 0.28]$ ). However, the credible intervals for both variables in all models include 0 indicating unreliable effects of prestige and relevance ratings on recall. Furthermore, none of these models had better

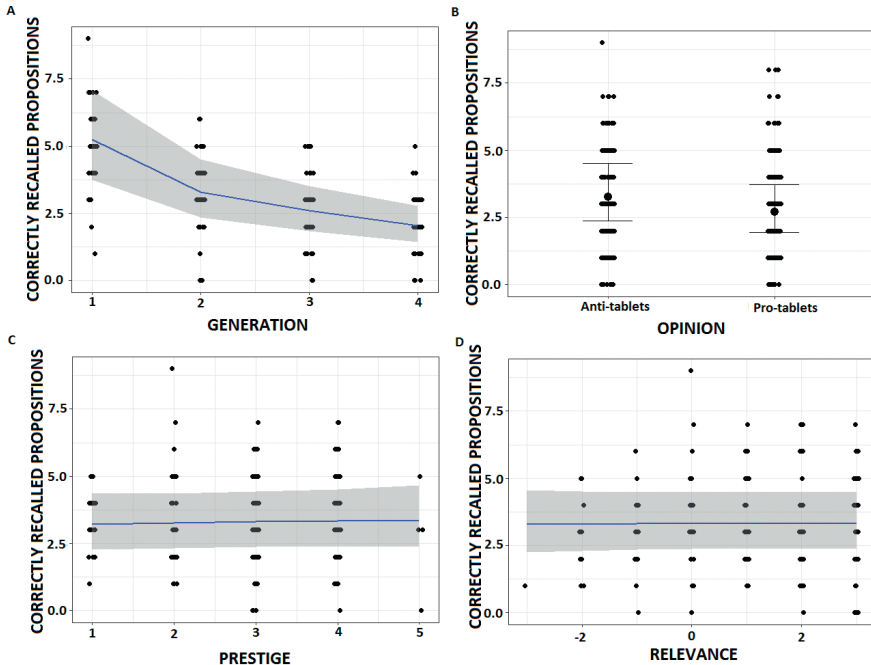


FIGURE 4 Marginal effects plots of the prestige-relevance model. A: The correct recall of propositions decreased monotonically with generation ( $B = -0.95$ ,  $SE = 0.09$ , 89% CI  $[-1.09, -0.81]$ ), the greatest decrease being between generation 1 to generation 2. B: The pro-tablets view was transmitted worse than the anti-tablets view ( $B = -0.19$ ,  $SE = 0.06$ , 89% CI  $[-0.29, -0.10]$ ). C: Prestige is positively related to recall but this effect is very weak and the wide standard error shows that it is unreliable ( $B = 0.05$ ,  $SE = 0.15$ , 89% CI  $[-0.19, 0.28]$ ). D: Relevance is positively related to recall but this effect is very weak and the wide standard error shows that is unreliable ( $B = 0.01$ ,  $SE = 0.13$ , 89% CI  $[-0.19, 0.23]$ ).

fit than the control model ( $LOOIC = 1301.2$ ). Consequently, these alternative analyses did not provide support for either  $H_1$  or  $H_2$ .

#### 4 Discussion

In this experiment, we studied whether the prestige and the relevance of sources of information affect the transmission of arguments in favour (“pro-tables”) and against (“anti-tablets”) the replacement of textbooks with computer tablets in schools. Contrary to our predictions, we failed to find a reliable advantage of the prestige ( $H_1$ ) or relevance ( $H_2$ ) of the sources of information on

the transmission of their arguments using two different analytical procedures: (i) assuming that the sources differ in prestige and relevance in a similar way for all participants in our sample (Head of the Department of Education: high prestige, high relevance; airline pilot: high prestige, low relevance; cleaner: low prestige, low relevance) and (ii) using the participants' own ratings of prestige and relevance of the sources as predictors of recall.

Although the effect of prestige on recall was statistically unreliable, it was in the expected direction. This suggests that either prestige has a weak effect on the recall of information or the results in the expected direction were due to chance. If the effects of prestige on recall are too small to have multi-generational effects in our experiment, this might indicate that the prestige of the source is not important in reality, where information is transmitted repeatedly from person to person. This would mean that, contrary to Henrich and Gil-White's (2001) hypothesis, there is no intrinsic memory bias to recall better what prestigious individuals say or argue. It would also mean that purported cases of prestige bias in real life (e.g. the 'Angelina Jolie effect', see Introduction) are exaggerated or spurious.

Alternatively, people might recall better what prestigious individuals say due to *repeated or redundant transmission* (Morin, 2016). That is, outside the experimental setting, people are exposed to the opinions of prestigious individuals multiple times (e.g. on TV, radio, and conversations) and this multiple exposure is what makes them more memorable. Future transmission chain studies could manipulate both the prestige of the sources and the number of exposures or number of sources per argument within each generation (e.g. Eriksson & Coultas, 2012; Kempe & Mesoudi, 2014), rather just one source per argument, to test this alternative.

Another possibility is that our manipulation of prestige (i.e. ascribing opinions to fictitious sources of information with different levels of prestige) does not adequately capture the way prestige works in everyday life. In large scale societies, especially in the digital era, people encounter individuals who are prestigious at different societal levels. *Prestigious individuals at the group level* are people who are admired and respected within a face-to-face group united by a common task or activity (e.g. a basketball team or a student class). Here, prestigious individuals and the people who admire them know each other and interact. *Prestigious individuals at the society level* (e.g. Hollywood actors/actresses, pop singers, and professional footballers) do not normally know personally their admirers. Rather, their admirers have the illusion of familiarity and personally knowing the prestigious individuals, giving rise to a type of relationship known as *para-social interaction* (Horton & Wohl, 1956). At both levels, social learners already know and admire the prestigious individual,

while participants in our study were for the first time exposed to the (fictional) sources. It is possible that, if participants already knew and admired the sources of information, the effects of prestige on recall might have been stronger. For example, if fans of a famous singer such as Beyoncé listens to her arguing in favour or against the replacement of computer tablets at school, these fans might recall better these arguments than similar arguments provided by other sources that they know but they do not admire. Nevertheless, similar manipulations to the one used in our study have been effective in demonstrating a positive effect of prestige on attention (Dalmaso et al., 2014; Dalmaso et al., 2012; Ratcliff et al., 2011). This suggests that prestige-biased social learning might have different effects on different measures of influence such as recall of information, attention towards the sources, opinion formation, etc. (Jiménez & Mesoudi, 2019). Future research will need to determine whether prestige-biased social learning only occurs at certain levels (e.g. decision-making under uncertainty, attention towards sources) but not others (e.g. memory).

The recall advantage of the “anti-tablets” arguments over the “pro-tablets” arguments was unexpected. It is possible that the anti-tablets advantage was caused by a *confirmation bias effect*, i.e. better recall of information that is congruent with pre-existing attitudes (Frost et al., 2015). Supporting this, the full sample of participants was initially slightly against the replacement of textbooks by computer tablets at schools (Median = -1). However, participants’ attitudes towards tablets did not predict recall, which contradicts the confirmation bias explanation (see also Jiménez, Mesoudi, & Tehrani, 2020).

In conclusion, this experiment does not provide convincing evidence in support of a prestige bias in the recall of information. To our knowledge, this is the first experiment studying this question. Therefore, it is difficult to know whether the prestige of the source of information genuinely does not affect recall, or whether our study suffers from methodological limitations (e.g. our manipulations of prestige). Given the potentially important consequences of prestige-driven information diffusion, such as the spread of attitudes towards breast screening or vaccination, we encourage further tests of whether information attributed to prestigious sources is preferentially recalled and transmitted through society.

### Acknowledgements

This research was supported by the Leverhulme Trust (grant RPG-2016-122658 awarded to AM). We thank Rosie Layfield for the second coding of participants’

recall and Joseph M. Stubbersfield, Charlotte Brand, and Shakti Lamba for useful comments on earlier versions of the manuscript.

## References

- Acerbi, A. (2019). *Cultural Evolution in the Digital Age*: Oxford University Press.
- Atkisson, C., O'Brien, M. J., & Mesoudi, A. (2012). Adult learners in a novel environment use prestige-biased social learning. *Evolutionary Psychology*, 10(3), 519–537.
- Bangerter, A. (2000). Transformation between scientific and social representations of conception: The method of serial reproduction. *British Journal of Social Psychology*, 39, 521–535. doi:10.1348/014466600164615.
- Bartlett, F. C. (1932). *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01.
- Bebbington, K., MacLeod, C., Ellison, T. M., & Fay, N. (2017). The sky is falling: evidence of a negativity bias in the social transmission of information. *Evolution and Human Behavior*, 38(1), 92–101. doi:10.1016/j.evolhumbehav.2016.07.004.
- Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 28. doi:10.18637/jss.v080.i01.
- Bürkner, P.-C., & Charpentier, E. (2018). Monotonic Effects: A Principled Approach for Including Ordinal Predictors in Regression Models. *PsyArXiv*. doi:10.31234/osf.io/9qkhj.
- Chudek, M., Baron, A. S., & Birch, S. (2016). Unselective Overimitators: The Evolutionary Implications of Children's Indiscriminate Copying of Successful and Prestigious Models. *Child Development*, 87(3), 782–794. doi:10.1111/cdev.12529.
- Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, 33, 46–56.
- Dalmaso, M., Galfano, G., Coricelli, C., & Castelli, L. (2014). Temporal Dynamics Underlying the Modulation of Social Status on Social Attention. *PLOS ONE*, 9(3). doi:10.1371/journal.pone.0093139.
- Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biological Letters*, 8(3), 450–452. doi:10.1098/rsbl.2011.0881.
- Desai, S., & Jena, A. B. (2016). Do celebrity endorsements matter? Observational study of BRCA gene testing and mastectomy rates after Angelina Jolie's New York Times editorial. *bmj*, 355, i6357.

- Eriksson, K., & Coultas, J. C. (2012). The advantage of multiple cultural parents in the cultural transmission of stories. *Evolution and Human Behavior*, 33(4), 251–259. doi:10.1016/j.evolhumbehav.2011.10.002.
- Eriksson, K., & Coultas, J. C. (2014). Corpses, Maggots, Poodles and Rats: Emotional Selection Operating in Three Phases of Cultural Transmission of Urban Legends. *Journal of Cognition and Culture*, 14(1–2), 1–26. doi:10.1163/15685373-12342107.
- Evans, D. G., Wisely, J., Clancy, T., Lalloo, F., Wilson, M., Johnson, R., ... Howell, A. (2015). Longer term effects of the Angelina Jolie effect: increased risk-reducing mastectomy rates in BRCA carriers and other high-risk women. *Breast Cancer Research*, 17(1), 143.
- Frost, P., Casey, B., Griffin, K., Raymundo, L., Farrell, C., & Carrigan, R. (2015). The Influence of Confirmation Bias on Memory and Source Monitoring. *The Journal of General Psychology*, 142(4), 238–252. doi:10.1080/00221309.2015.1084987.
- Garfield, Z. H., Hubbard, R. L., & Hagen, E. H. (2019). Evolutionary Models of Leadership. *Human Nature*. doi:10.1007/s12110-019-09338-4.
- Henrich, J. (2015). *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, New Jersey : Princeton University Press.
- Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22(3), 165–196. doi:10.1016/S1090-5138(00)00071-4.
- Henrich, J., & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, 12(3), 123–135. doi:10.1002/evan.10110.
- Holtgraves, T., Srull, T. K., & Socall, D. (1989). Conversation memory: the effects of speaker status on memory for the assertiveness of conversation remarks. *Journal of Personality & Social Psychology*, 56(2), 149–160.
- Horton, D., & Wohl, R. (1956). Mass Communication and Para-Social Interaction. *Psychiatry*, 19(3), 215–229. doi:10.1080/00332747.1956.11023049.
- Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, 5(1), 20. doi:10.1057/s41599-019-0228-7.
- Jiménez, Á. V., Mesoudi, A., & Tehrani, J. J. (2020). No evidence that omission and congruity biases affect the perception and recall of vaccine-related information. *PLOS ONE* 5(3): doi: 10.1371/journal.pone.0228898
- Jiménez, Á. V., Stubbersfield, J. M., & Tehrani, J. J. (2018). An experimental investigation into the transmission of antivax attitudes using a fictional health controversy. *Social Science & Medicine*, 215, 23–27. doi:10.1016/j.socscimed.2018.08.032.
- Jolie, A. (2013). My medical choice. Retrieved from <https://www.nytimes.com/2013/05/14/opinion/my-medical-choice.html>.



- Juthe, R. H., Zaharchuk, A., & Wang, C. (2015). Celebrity disclosures and information seeking: the case of Angelina Jolie. *Genetics in Medicine*, 17(7), 545.
- Kashima, Y. (2000). Maintaining Cultural Stereotypes in the Serial Reproduction of Narratives. *Personality and Social Psychology Bulletin*, 26(5), 594–604. doi:10.1177/0146167200267007.
- Kempe, M., & Mesoudi, A. (2014). An experimental demonstration of the effect of group size on cultural accumulation. *Evolution and Human Behavior*, 35(4), 285–290. doi:https://doi.org/10.1016/j.evolhumbehav.2014.02.009.
- Kintsch, W. (1974). *The representation of meaning in memory*. Oxford, England: Lawrence Erlbaum.
- Kleeman, D. (2014). 10 reasons why we need research literacy, not scare columns. Retrieved from [https://www.huffpost.com/entry/10-reasons-why-we-need-re\\_n\\_4940987?ncid=fbcklnkushpimg00000042](https://www.huffpost.com/entry/10-reasons-why-we-need-re_n_4940987?ncid=fbcklnkushpimg00000042).
- Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, 79, 328–348. doi:https://doi.org/10.1016/j.jesp.2018.08.009.
- Mesoudi, A. (2007). Using the methods of experimental social psychology to study cultural evolution. *Journal of Social, Evolutionary, and Cultural Psychology*, 1(2), 35.
- Mesoudi, A. (2011). *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*: University of Chicago Press.
- Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, 97(3), 405–423. doi:10.1348/000712605x85871.
- Morin, O. (2016). *How Traditions Live and Die*. New York: Oxford University Press.
- Peer, E., Samat, S., Brandimarte, L., & Acquisti, A. (2016). Beyond the Turk: an empirical comparison of alternative platforms for crowdsourcing online behavioral research. *SSRN*. doi:10.2139/ssrn.2594183.
- Ratcliff, N. J., Hugenberg, K., Shriver, E. R., & Bernstein, M. J. (2011). The allure of status: high status targets are privileged in face processing and memory. *Personality and Social Psychology Bulletin*, 37(8), 1003–1015. doi:10.1177/0146167211407210.
- Reyes-Garcia, v., Molina, J. L., Broesch, J., Calvet, L., Huanca, T., Saus, J., ... McDade, T. W. (2008). Do the aged and knowledgeable men enjoy more prestige? A test of predictions from the prestige-bias model of cultural transmission. *Evolution and Human Behavior*, 29(4), 275–281. doi:10.1016/j.evolhumbehav.2008.02.002.
- Rowan, C. (2014). 10 reasons why handheld devices should be banned for children under the age of 12. Retrieved from [https://www.huffpost.com/entry/10-reasons-whyhandheld-devices-should-be-banned\\_n\\_4899218](https://www.huffpost.com/entry/10-reasons-whyhandheld-devices-should-be-banned_n_4899218).
- Stibbard-Hawkes, D. N. E., Attenborough, R. D., & Marlowe, F. W. (2018). A noisy signal: To what extent are Hadza hunting reputations predictive of actual

- hunting skills? *Evolution and Human Behavior*, 39(6), 639–651. doi:10.1016/j.evolhumbehav.2018.06.005.
- Stubbersfield, J. M., Tehrani, J. J., & Flynn, E. G. (2014). Serial killers, spiders and cybersex: Social and survival information bias in the transmission of urban legends. *British Journal of Psychology*, 106(2), 288–307. doi:10.1111/bjop.12073.
- Stubbersfield, J. M., Tehrani, J. J., & Flynn, E. G. (2017). Chicken Tumours and a Fishy Revenge: Evidence for Emotional Content Bias in the Cumulative Recall of Urban Legends. *Journal of Cognition and Culture*, 17(1–2), 12–26. doi:10.1163/15685373-12342189.
- Team, R. C. (2019). R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>.
- Vehtari, A., & Gabry, J. (2019). *Bayesian Stacking and Pseudo-BMA weights using the loo package*. Retrieved from <https://cran.r-project.org/web/packages/loo/vignettes/loo2-weights.html>.
- Vehtari, A., Gabry, J., & Gelman, A. (2019). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. Retrieved from <https://mc-stan.org/loo>.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leaveone-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. doi:10.1007/s11222-016-9696-4.
- Von Rueden, C. (2014). The roots and fruits of social status in small-scale human societies. In *The psychology of social status* (pp. 179–200): Springer.
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, 33(4), 341–356. doi:10.1016/j.dr.2013.08.002.
- Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2018). Using stacking to average Bayesian predictive distributions (with discussion). *Bayesian Analysis*, 13(3), 917–1003.