

# Intentional transmission of knowledge optimises cumulative cultural evolution

Gemma Mackintosh · Mark Atkinson · Elizabeth Renner · Christine A. Caldwell 

Received: 29 June 2025 / Accepted: 14 July 2026

© The Author(s) 2026

## Abstract

Human evolutionary success has been attributed to the capacity to accumulate knowledge and skills over generations, via cultural transmission. But explanations for the distinctiveness of human culture remain heavily debated. Accounts appealing to human cognitive sophistication must explain users' limited mechanistic understanding of many cultural traits, whereas accounts that emphasise cultural adaptations as emergent outcomes of selective retention over generations must explain what precludes this in other species. In humans the transmission process itself may implicate cognitive capacities not available to other species, but to date there has been little direct evidence linking differences in cognitive capacity with differences in social transmission behaviour that impact the accumulation of cultural information. Using an experimental task, here we show that adults, but not young children, exhibit strategic preferences in the information they transmit to potential learners. This strategic selection promoted the cultural accumulation of beneficial discoveries, relative to randomly selected task information. Children, despite task competence, did not show such strong strategic transmission preferences, and accordingly obtained no additional benefit from intentional, over inadvertent, knowledge transmission until 9–10 years old. This late development supports the view that human cultural transmission may implicate cognitive abilities dependent on cultural experience, unavailable to other species.

## Introduction

The accumulation of beneficial knowledge and skills over multiple generations via social transmission, sometimes referred to as “cumulative cultural evolution” (or just “cumulative culture”), appears to be ubiquitous in human populations<sup>1,2</sup> (although see also<sup>3</sup>). Several recent high-profile accounts of human evolution, while disagreeing over the origins and cognitive underpinnings of cumulative culture, converge in regarding this process as key to humans' success as a species<sup>4–9</sup>. Many other distinctively human traits (e.g. technology, language and other systems of symbolic representation) are widely understood to be products of cumulative culture<sup>4,10</sup>, and some cultural practices have had such profound evolutionary consequences that their effects can be detected in the human genome<sup>11</sup>.

Cumulative culture has thus provided a compelling and parsimonious explanation for human distinctiveness. But this highlights a more fundamental question, of why cumulative culture itself appears restricted to humans.

---

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.



Examples of similar processes in nonhuman animals<sup>12,13</sup> are widely acknowledged to be rare, and lacking some features typical of human cumulative culture<sup>14</sup>. However, there is still no consensus regarding the explanation for this striking disparity between humans and other species.

For example, although some theorists have argued that the complexity of human culture is attributable to human capacities for cognitive reasoning (e.g. causal understanding)<sup>15,16</sup>, others maintain that cultural adaptations occur through the selective retention of beneficial changes, often in the absence of understanding<sup>17,18</sup>. Here we investigate an alternative to both views, specifically the idea that the effectiveness of human social transmission does depend on peculiarly human cognitive skills, but that these involve reasoning about states of knowledge (both one's own, and those of others).

We report an experiment investigating the accumulation of beneficial information over learner generations, comparing the effects of inadvertent social information with intentional information sharing. This comparison may be particularly relevant to understanding differences between human and non-human cultural evolution, as humans readily engage in intentional knowledge sharing, whereas most social information use in non-humans depends on inadvertent social information. *Inadvertent social information* refers to information available incidentally as a result of others' activities<sup>19</sup>. As such, it is neither selected for, nor intended to have, any communicative function for observers. In spite of this, observers may nonetheless be able to use such "public information" to their advantage. For example, others' foraging success may provide valuable information about patch quality<sup>20,21</sup>. Learning from such inadvertent social information may be the main mechanism supporting cases of even relatively enduring behavioural traditions in non-humans. For example, chimpanzees learn the foraging and social behaviours particular to their social group<sup>22</sup>, but there is very limited evidence of experienced individuals modifying their behaviour in ways that would function to facilitate others' learning<sup>23,24</sup>. The chimpanzees' learning therefore appears to occur primarily through exposure to information which is available purely incidentally as a result of others' activities.

Nonetheless, behaviours that function to teach have been documented in some non-humans. However, these examples represent fairly narrowly focused adaptations that facilitate learning only in the specific context of particular species-typical behaviours<sup>25,26</sup>. And critically, such non-human teaching does not appear to be accompanied by any intent to inform, likely due to limited capacities for mental state understanding<sup>23,27</sup>.

Distinctively human teaching, which does involve representation of states such as knowledge and ignorance, has been identified by a number of authors as potentially a key mechanism that could explain the evolutionary anomaly of human cumulative culture<sup>7,27–29</sup>. Despite the abundance of speculative arguments, there is little evidence demonstrating that intentional knowledge sharing can in fact facilitate the accumulation of information, over and above inadvertent social information. Experimental studies investigating cumulative culture under laboratory conditions (which control the transmission of information between participants over multiple learner generations) have sometimes involved manipulations of learning conditions<sup>30–33</sup>. However, in these previous studies, any advantage found in "teaching" conditions might simply be a consequence of learners receiving more information in these conditions compared with the relevant comparison groups. Indeed, in the "teaching" condition of one of these studies<sup>32</sup>, information from different sources (including teaching) was provided additively, with the trend in performance across conditions directly corresponding to the richness of information in terms of the number of sources available (note also that this study investigated retention of expertise rather than its cumulative development). These studies therefore fail to test the key question (as we see it) of whether information acquired through teaching is actually *better*, in the sense that it provides a more efficient means of transmitting expertise, compared with other potential sources of social information. In the absence of such evidence, it cannot reasonably be claimed that it is teaching per se that facilitates cultural learning, over and above what is possible from alternative sources of social information.

It is our view that distinctively human teaching as we have described it above (i.e., intentional knowledge transmission) permits exactly the kind of efficiencies likely to allow expertise to accumulate even under conditions that severely constrain learning opportunities, in ways that are not possible for other sources of social

information. A small snapshot of inadvertent social information may or may not be useful to a potential learner. In contrast, a human teacher intentionally transmitting their knowledge has control over the selection of information and would be expected to choose strategically to maximise the value of the information they provided (however limited) to a potential learner.

In the current study therefore, we aimed to directly compare conditions which differed only with respect to whether the transmitted information had been intentionally shared or was acquired incidentally from another's activity. We test the hypothesis that when transmission occurs through a learning bottleneck (i.e., only limited information is transmitted), intentional knowledge sharing will allow for greater accumulation of beneficial information, as compared with inadvertent social information, due to the potential for highlighting information likely to hold the greatest value for the learner.

We go on to investigate the development of such information-sharing skills in children, studying three different age groups (5–6 years, 7–8 years, and 9–10 years) to establish which, if any, of these groups perform in a similar way to human adults. This part of the investigation was included to shed light on which cognitive capacities may be necessary to support the accumulation of beneficial information via the mechanism of effective intentional information sharing. Children's cognitive development typically progresses according to a relatively predictable chronology. Comparing age groups can therefore potentially permit inferences about the specific cognitive advances implicated in adult-like competence on our experimental task, including whether these abilities are ones which are shared with other species.

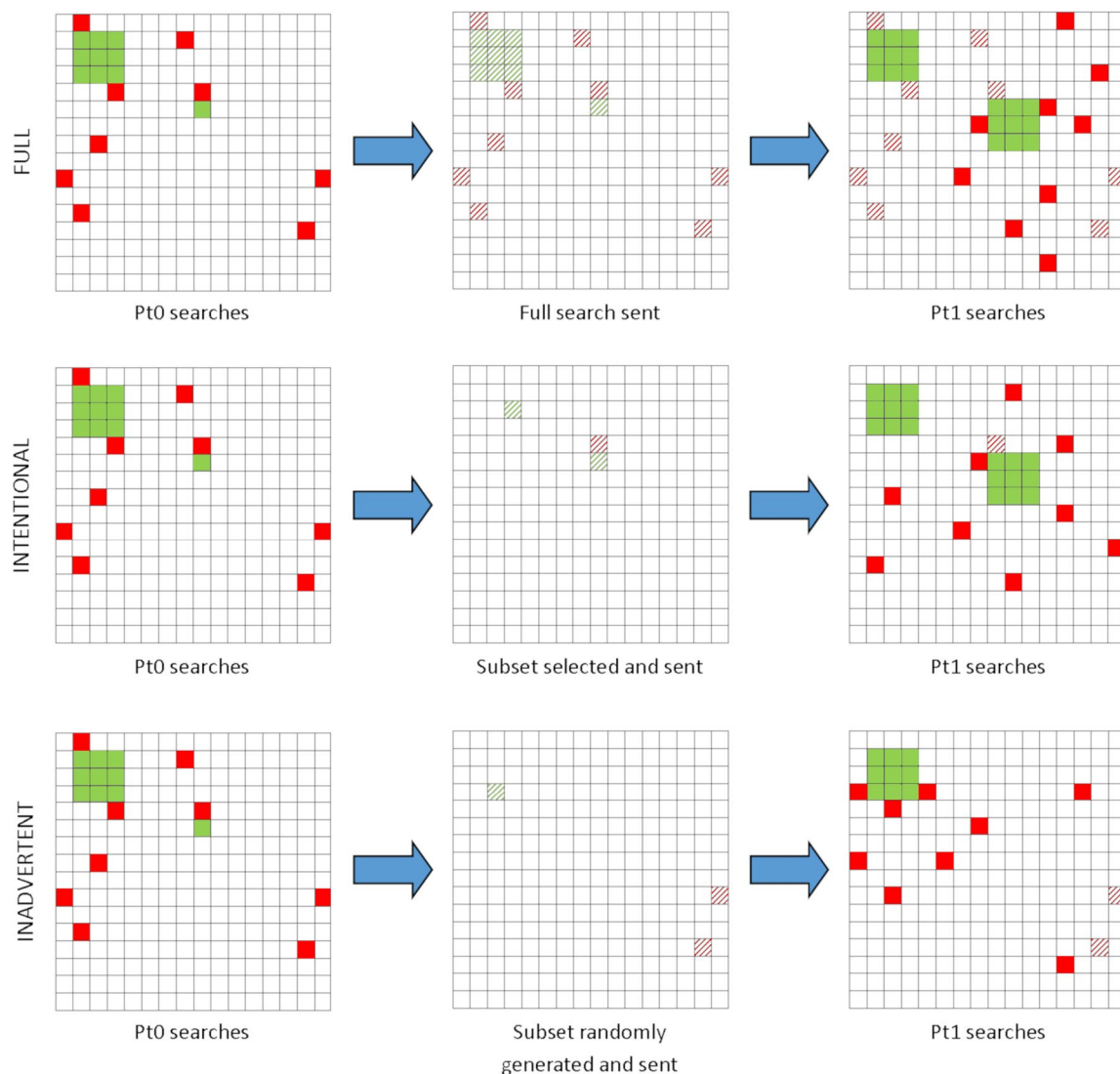
We used a transmission chain design, and each chain was assigned to one of three different transmission conditions: Full, Intentional and Inadvertent information. All participants completed the same grid search task (Fig. 1). In the Full information condition, information receivers (henceforth referred to as *cultural offsprings*) were shown the entire search outcome obtained by their predecessor in the chain (henceforth referred to as *cultural parents*). In the Intentional information condition, each participant, in their role as cultural parent (after having completed the grid search in the role of cultural offspring), was asked to select a subset of their search outcome (three grid squares) for transmission to the next cultural offspring. In the Inadvertent information condition, an equivalent subset (three grid squares) was randomly selected from the search outcome, and this formed the information transmitted to the next participant in the chain (see Methods for full details).

## Experiment 1: Adults

For adult participants, we predicted that task performance would improve over generations, reflecting the accumulation of task expertise via social transmission. We also predicted that there would be better task performance in the Intentional information condition, compared with the Inadvertent information condition, and that this difference would be more apparent in later generations, compared with earlier ones (indicative of increasing benefits due to the accumulation of information enabled by social transmission). Performance in the Full information condition was expected to be best of all, with this providing an indication of the upper limit of the benefits that could accumulate via social transmission in this task context. Our key predictions were pre-registered: <https://osf.io/jszvc/>.

## Experiment 2: Children

Our predictions for the experiment with children were less specific than for adults, as we had no definite expectations regarding the age at which children would develop general task competence, nor the age at which they might begin to show some benefit of intentional, over inadvertent, information sharing. Nonetheless, it was expected that any difference in performance between the Intentional and Inadvertent conditions would be more pronounced in older, compared with younger, children, due to advances in cognitive development.



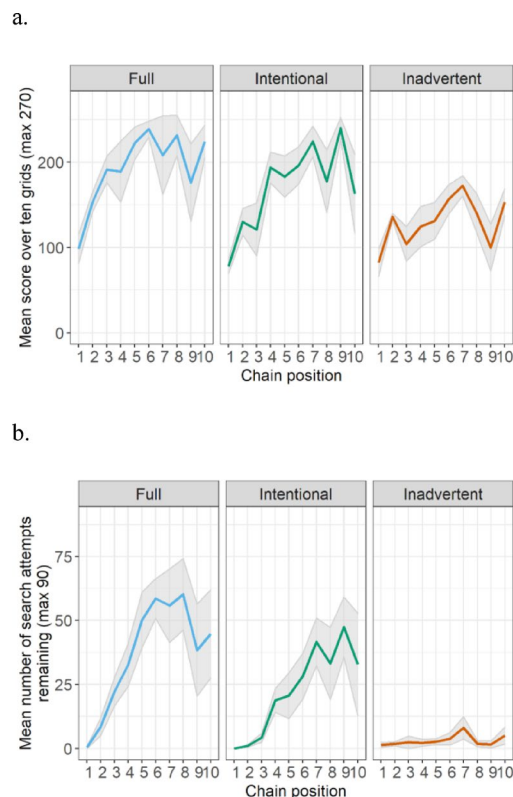
**Fig. 1** Examples illustrating intergenerational transmission in the three task conditions. Hits were arranged into three  $3 \times 3$  targets within each individual grid arrangement. Solid shading indicates the search outcome, with green indicating a hit, and red a miss. Striped shading indicates information that was transmitted to the cultural offspring (which may or may not be subsequently selected by them, see final column). Children completed the same task, in the same three conditions (top row, Full information; middle row, Intentional information; bottom row, Inadvertent information), as adults, with only minor variations in presentation (e.g. colours), and a smaller number of trials per participant (see Methods for full details).

## Results

### Adults

Our primary measure of task success was task score. For the purpose of analysis, participants' scores were calculated using their total score across all 10 completed grids. The maximum score possible for participants to achieve was 270, that is, 27 points for each of the ten grids searched. Mean total scores for each of the three conditions, across all ten chain positions, are displayed in Fig. 2a. A GLMM was used to investigate effects of condition (full details reported in Table S1). Condition and chain position (centred) were included as independent variables in the model, and chain was included as a random variable, with the Intentional condition used as the

**Fig. 2** Adult participants' mean task score (**a**) and mean search attempts remaining at end of search (**b**) over generations, by condition. Shaded areas represent  $\pm 1$  standard error.



intercept. Across all conditions, scores increased with chain position ( $p < .001$ ). Scores were higher in the Intentional condition compared to the Inadvertent condition, as predicted ( $p = .002$ ). Scores were higher in the Full condition compared to the Inadvertent condition ( $p < .001$ ) but were not significantly different between the Full and Intentional conditions.

Interactions between condition and chain position were used to identify whether any differences between conditions increased over generations, in line with our predictions. The effect of chain position showed no significant difference between conditions.

We also evaluated task success using a count of each participant's unused search attempts. This measure of search attempts remaining was effectively a measure of errorless performance (see Methods for further details). The maximum possible number of search attempts remaining was 90, that is, nine search attempts for each of the ten grids searched. This maximum could only be achieved by a participant selecting grid squares that were exclusively "hits" (parts of a target) and avoiding all "miss" (i.e. non-target) grid squares. Mean search attempts remaining for each of the three conditions, across all ten chain positions, are displayed in Fig. 2b. A GLMM equivalent to the one used on the task scores was used to investigate effects of condition on participants' total search attempts remaining, across all ten grids (full details in Table S1).

There was again a significant effect of chain position across all conditions ( $p < .001$ ), indicating that participants who took part later in the chain had more search attempts remaining compared with those who took part earlier. The number of search attempts remaining was higher in the Intentional, compared with Inadvertent, condition ( $p < .001$ ), as predicted. The number of search attempts remaining was also higher in the Full condition compared with both the Intentional ( $p = .001$ ) and Inadvertent ( $p < .001$ ) conditions.

Interactions between condition and chain position were also used for the search attempts measure. Effects of chain position on search attempts remaining were stronger in the Intentional condition compared with the Inadvertent condition ( $p = .002$ ) and were also stronger in the Full condition compared with the Inadvertent condition ( $p < .001$ ), but not when compared with the Intentional condition.

Our pre-registered predictions included a specific expectation regarding the relative rates at which the maximum score would be reached for any given grid, between the three conditions. Consistent with our general expectations regarding task performance across conditions, we predicted that all targets within a grid would typically be found in the fewest generations in the Full condition, followed by the Intentional condition, with the Inadvertent condition requiring the greatest number of generations before the maximum score was reached. We compared time-to-event across conditions using survival analysis. A log rank test showed a difference between the three groups ( $p < .001$ ). Cox Hazard Ratios between the Intentional condition and the other two showed that grids tended to be fully solved earlier for the Full condition compared with Intentional ( $p = .024$ ), and earlier in the Intentional condition compared with Inadvertent ( $p < .001$ ), in line with our predictions. See Fig. S1 for the Kaplan–Meier survival plot.

In the Intentional condition, cultural parents preferentially transmitted hits over misses, with a higher proportion of hits sent from each grid compared with the proportion available for transmission ( $t = 15.892$ ,  $df = 499$ ,  $p < .001$ , see Figure S2 for further details).

The positions transmitted within each target were also analysed to determine whether cultural parents in the Intentional condition preferentially selected particular target segments for transmission. The proportions of each target segment sent can be seen in Fig. 3. The data for this analysis only included target segments sent from targets which had been fully revealed (all nine grids squares selected) by the cultural parent.

A Pearson's Chi-Square test was used to analyse any bias in cultural parents' transmission of locations within fully revealed targets (e.g., top left, centre, bottom right). Full details of the statistical tests are provided in Table S2. In the Intentional condition, the distribution of segments chosen for transmission was significantly different from uniform, showing a bias for transmitting particular locations from fully revealed targets. In contrast, in the Inadvertent condition (where transmitted segments were selected randomly within the task script), unsurprisingly no such bias was found.

Follow-up Binomial tests were carried out to establish which segment positions were selected more or less frequently than would be expected by chance in the Intentional condition (see Table S2). The top left and central segments were selected at frequencies significantly above chance, with all other segments being selected at frequencies significantly below chance.

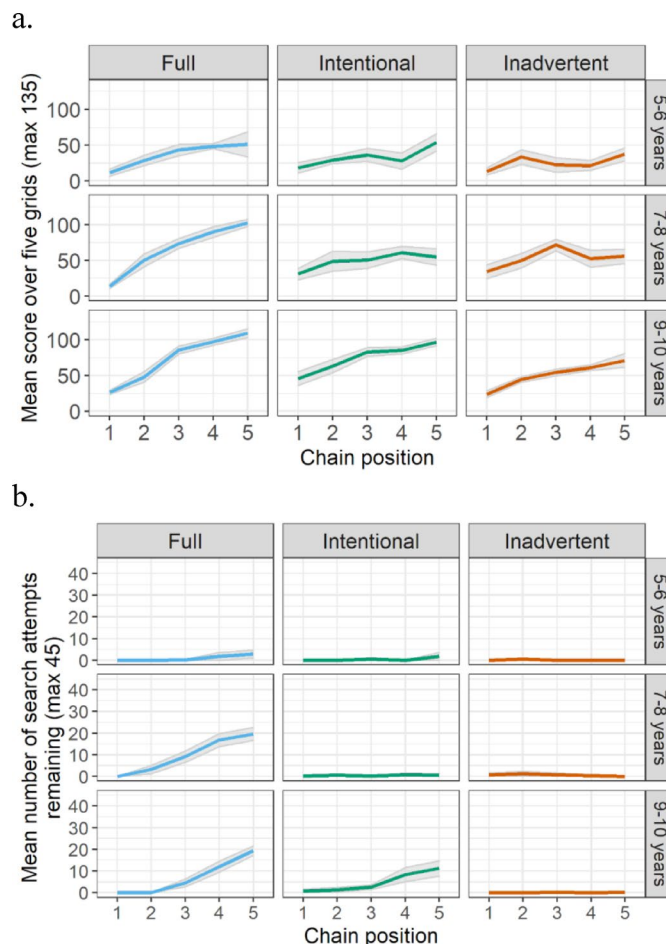
## Children

We analysed children's task performance using the same measures of task success used for the adults (total task score, and total search attempts remaining). Mean total scores for each of the three conditions, across all chain positions, for all three age groups, are displayed in Fig. 4a. Figure 4b displays equivalent plots for search attempts remaining. See Methods for details of procedural differences between Experiment 1 (adults) and Experiment 2 (children) in terms of chain length and number of grids.

**Fig. 3** Count of particular target segments (corners, sides, centre) transmitted in Experiment 1 (adults), by condition and age group. This analysis includes data only from targets with all nine segments revealed. In the Intentional condition, cultural parents showed a strong preference for transmitting the centre segment of the  $3 \times 3$  targets, with a weaker bias towards the top left segment. All other segments were sent with lower than chance probability. Full statistical details can be found in Table S2. Inadvertent and Full conditions are also shown for comparison.



**Fig. 4** Children’s mean task score (a) and mean search attempts remaining at end of search (b) over generations, by condition. Shaded areas represent  $\pm 1$  standard error.



GLMMs equivalent for those used in Experiment 1 were built separately for each of the three age groupings. Therefore, in each model, only the data from the specified age category was used to compare the differences between conditions within that age category alone. In these models, we included condition and chain position (centred) as independent variables. Chain was included as a random variable where possible.

For the 5- to 6-year-old children (Table S3), scores increased with chain position as they had in adults ( $p = .015$ ), but there were no differences between conditions, either in overall score, or as effects of interaction with chain position. For the measure of search attempts remaining, there was no effect of chain position in this age group, nor any differences between conditions, either in overall search attempts remaining, or as effects of interaction with chain position.

For the 7- to 8-year-old children (Table S4) there was a marginally non-significant effect of chain position on total score ( $p = .051$ ). Differences between conditions were restricted to differences between the Full and other conditions, with no differences found between the Intentional and Inadvertent conditions. Scores were higher in the Full condition compared with the Intentional condition ( $p = .018$ ), although only marginally non-significantly higher than the Inadvertent condition ( $p = .076$ ). Total scores showed no difference between the Intentional and Inadvertent conditions. Looking at interaction effects, chain position had a stronger effect on scores in the Full condition compared with both the Intentional ( $p = .001$ ) and Inadvertent conditions ( $p < .001$ ), but there was no difference between the Intentional and Inadvertent conditions.

For the measure of search attempts remaining in the 7- to 8-year olds, there was no effect of chain position on overall search attempts remaining. Effects of condition were (as with the total score measure) restricted to differences between the Full condition and the other two. Search attempts remaining were higher overall in the Full

condition compared with both the Intentional ( $p < .001$ ) and Inadvertent conditions ( $p < .001$ ), with no difference found between the Intentional and Inadvertent conditions. The same picture emerged when looking for interactions with chain position: chain position had a stronger effect on search attempts remaining in the Full condition compared with both the Intentional ( $p < .001$ ) and Inadvertent ( $p < .001$ ) conditions, with no difference found between the Intentional and Inadvertent conditions.

The 9- to 10-year olds (Table S5) were the only group of children in which we found significant differences in the key comparison of interest, between the Intentional and Inadvertent conditions. In this respect, the results from this group resembled the pattern found in adults. Taking the task score measure first, scores in this group increased with chain position ( $p < .001$ ). Differences in total score between conditions replicated the pattern found in adults, with higher scores in the Intentional compared with Inadvertent condition ( $p < .001$ ), and also in the Full compared with Inadvertent condition ( $p < .001$ ), and no difference between the Full and Intentional conditions. Looking at interactions with chain position, there was a stronger effect of chain position on score in the Full compared with Intentional ( $p = .004$ ) and Inadvertent ( $p < .001$ ) conditions, although there was no significant difference in the effect of chain position on score between the Intentional and Inadvertent conditions.

For the measure of search attempts remaining in the 9- to 10-year-olds, across all conditions there was an effect of chain position on search attempts remaining ( $p < .001$ ). Total search attempts remaining were higher in the Intentional compared with Inadvertent condition ( $p < .001$ ), higher in the Full compared with Inadvertent conditions ( $p < .001$ ), and marginally non-significantly higher in the Full compared with Intentional condition ( $p = .060$ ). Looking at interactions with chain position, there was a greater increase in search attempts remaining with chain position in the Intentional condition compared with the Inadvertent condition ( $p < .001$ ), and between the Full and both the Intentional ( $p = .005$ ) and Inadvertent conditions ( $p < .001$ ).

As with the adults, the positions transmitted by cultural parents in the Intentional condition were analysed to investigate selection preferences. Two participants (both in the age 5–6 age category) were excluded from this analysis due to a technical error which meant that data on the specific grid squares selected were not written to file (see Methods).

In all age groups, cultural parents preferentially transmitted hits over misses, with a higher proportion of hits sent from each grid compared with the proportion available for transmission (age 5–6:  $t = 7.660$ ,  $df = 119$ ,  $p < .001$ ; age 7–8:  $t = 4.766$ ,  $df = 124$ ,  $p < .001$ ; age 9–10:  $t = 6.933$ ,  $df = 124$ ,  $p < .001$ ; see Figure S3 for further details).

The proportions of each target segment sent can be seen in Fig. 5. The data for this analysis only included target segments sent from targets which had been fully revealed (all nine grids squares selected) by the cultural parent. Pearson's Chi-Square test was used to analyse any bias in cultural parents' transmission of locations within fully revealed targets in the Intentional condition. Full details of the statistical tests are provided in Table S6. In all age groups, the distribution of segments chosen for transmission was significantly different from uniform in the Intentional condition, showing a bias for transmitting particular locations from fully revealed targets.

Follow-up Binomial tests using an adjusted alpha criterion of  $p \leq .006$  found that the central target segment was chosen for transmission with higher-than-expected frequency in all age groups (see Table S6). In the 7- to 8-year-old chains, the bottom centre target segment was chosen with lower-than-expected frequency, and in the 9- to 10-year-old chains the centre left, centre right, and bottom centre target segments were all chosen with lower-than-expected frequency. All other tests failed to reach the adjusted alpha criterion.

All age groups therefore showed some bias for transmitting the central segment of a target which had been fully revealed. However, the extent of the bias increased with age. Amongst adult participants, 60.15% of grid locations transmitted from fully-revealed targets were the central segment position. For the oldest children (9- to 10-year-olds), this fell to 33.46%. The equivalent percentages were 24.14% for the 7- to 8-year-olds, and 20.62% for the 5- to 6-year-olds. The counts for central versus other segments were significantly different between the adults and 9- to 10-year-olds ( $X^2 = 57.65$ ,  $df = 1$ ,  $p < .001$ ), and between 9- to 10-year-olds and 7- to 8-year-olds ( $X^2 = 3.920$ ,  $df = 1$ ,  $p = .048$ ), but not between 7- to 8-year-olds and 5- to 6-year-olds ( $X^2 = 0.260$ ,  $df = 1$ ,  $p = .610$ ).

**Fig. 5** Count of particular target segments (corners, sides, centre) transmitted in Experiment 2 (children), by condition and age group. This analysis includes data only from targets within which all nine segments were potentially available for transmission (i.e. the target had been fully revealed by the cultural parent). In the Intentional condition, cultural parents from all age groups showed a preference for transmitting the centre segment of the  $3 \times 3$  targets. The strength of this bias increased with age. Full statistical details can be found in Table S6. Inadvertent and Full conditions are also shown for comparison.



## Discussion

Our results suggest that intentional knowledge sharing can – at least under some circumstances – fully compensate for the loss of accumulated expertise that can otherwise result from a tight bottleneck on transmission (e.g., due to limited exposure time). Adult participants appeared able to anticipate the needs of learners, making significantly non-random choices about the information that they transmitted, which were indeed beneficial to receivers. Specifically, cultural parents were more likely to transmit the location of a “hit” compared with a “miss”. Furthermore, they preferentially selected central squares over other possible locations within the target. By using this strategy, cultural parents who found all targets effectively afforded the possibility of errorless performance by their offspring, by communicating the location of all three with their limited transmission allowance of just three grid squares.

Our results showed that this highlighting of particularly beneficial information supported cumulative culture, permitting both the accumulation of expertise and the maintenance of near-maximal reward behaviours even when opportunities for learning from others was strictly constrained. The strategic selective behaviour we identified appears consistent with accounts of distinctively human teaching, which has been characterised as involving a much higher level of sensitivity to the needs of the learner, compared with examples from other species<sup>34</sup>, with instruction tailored for relevance<sup>27</sup>.

Our second experiment, investigating children’s behaviour, was included to shed light on the cognitive abilities implicated in the strategic transmission behaviour observed in the adults, by revealing the age at which children begin to employ similar strategies and reap the benefits of intentional, over inadvertent, knowledge transmission. Based on our own logical rationale, and in line with accounts of distinctively human teaching, we had expected that an ability to represent mental states (Theory of Mind) would be a fundamental requirement for task proficiency. Children typically pass explicit false belief tasks at around four years of age<sup>35</sup> but we did not observe competent performance on our task until much later than this. Only our oldest group (nine- and ten-year-olds) demonstrated a performance advantage for intentionally selected information over an equivalent random sample. This older group also showed transmission biases most similar to those observed in adults, which were less apparent in the younger children.

There is evidence to suggest that children’s ability to engage in more complex reasoning about mental states continues to develop later into childhood, following the initial conceptual breakthrough of understanding of

mental states as representations. For example, children begin to solve problems that require reasoning about higher-order mental states (e.g. false beliefs about false beliefs) several years later than they would typically pass the standard first-order false belief test<sup>36</sup>. Furthermore, it is only around the age of seven when children begin to appreciate that others can acquire knowledge through inference as well as perceptual exposure. Children younger than this attribute ignorance, rather than knowledge, in situations where others could reasonably be expected to be able to infer details of an unwitnessed event on the basis of other information. This occurs even when that the child has herself made that very inference using the same information<sup>37</sup>.

In our task, effective transmission strategies could involve either (or both) of these later-developing socio-cognitive skills. Higher-order mental state reasoning would allow cultural parents to represent not just the knowledge state of their offspring, but also the offspring's beliefs about the parent's communicative intention. Appropriate reasoning about the cultural offspring's ability to make inferences would allow the parent to consider how the transmitted information was likely to be used, in the context of what the offspring already knows about the task (e.g., the size, shape and number of targets).

The interpretations above are purely speculative, as the current study does not permit us to draw conclusions about the precise reasoning behind participants' responses. Nevertheless, our age group comparisons are revealing in relation to the origins of the requisite cognitive skills (whatever those might be) and whether these are likely to be shared with other species. The cognitive capacities implicated are unlikely to represent biological adaptations selected specifically for their effects on cultural transmission. The apparently late development of these capacities suggests they are likely to be largely experience-dependent (including, although probably not limited to, cultural experience)<sup>4</sup>. It is possible however that these skills build upon other more fundamental cognitive adaptations selected for wider social and/or communicative functions<sup>40,41</sup>. Both Theory of Mind and prosocial motivations have been proposed to be evolutionary adaptations which emerge early in development and set humans apart from other species<sup>42,43</sup>. Both are likely basic requirements of our experimental task, and these more fundamental abilities could represent aspects of human biological endowment. But skilled performance on our task appears to depend on a significant duration of development beyond the ages at which Theory of Mind and prosocial behaviour are established. This suggests that substantial additional learning, probably through both direct experience of social interactions and observation of others, is required in order to master the selective transmission strategies necessary for effective teaching. Further research would be needed to determine the extent to which different kinds of experience facilitate skill acquisition (e.g., involving actively teaching the strategies to children just below the age of competence established here). Nonetheless, it is reasonable to conclude that if much of human cumulative culture implicates cognitive capacities that are themselves largely acquired through human cultural experience and/or are hitherto undocumented in other species (such as higher-order mental state reasoning and attribution of knowledge via inference), this could contribute to the explanation of why cumulative culture is rarely documented in other species.

It is important to note that some of the simplifying assumptions of our experimental task may place limits on the generalisability of our findings. To give just one example, our task involves a perfectly stable environment, in which the reward landscape remains fixed across generations. Further research would be needed to determine whether intentional transmission remains as useful (or perhaps becomes even more critical, if change can be recognised or anticipated) in less stable environments.

Despite much speculation about cognitive prerequisites of human cumulative culture, to our knowledge the current study is the first to demonstrate that populations with differing levels of cognitive development exhibit correspondingly different degrees of proficiency in social transmission which are demonstrably linked to the accumulation of cultural knowledge. Furthermore, even though we propose that human cumulative culture is generally dependent on highly-developed cognitive capacities, our account is not undermined by humans' apparently poor causal understanding of some culturally acquired expertise<sup>17,18</sup>. The cognitive capacities we propose as drivers of cumulative culture would not directly function to facilitate understanding of a trait itself, but instead an understanding of the relative states of knowledge of self and other, regarding the trait.

Methods

## Methods

### Experiment 1: Adults

#### Participants

We recruited 159 adults aged between 17 and 62 years (median age=20 years; 103 females). Data from nine of these participants were excluded due to experimenter or technical error (in accordance with exclusion criteria specified in the pre-registration: <https://osf.io/jszvc/>). The final sample consisted of 150 adult participants aged between 17 and 62 years (median age=20 years, 100 females). Of the final sample, 133 participants were recruited from the University of Stirling via an online recruitment database, through which students can earn tokens required for course completion by participating in research studies, and a further 17 participants were recruited from Glasgow Science Centre ([www.glasgowsciencecentre.org](http://www.glasgowsciencecentre.org)) on a voluntary basis.

Fifty participants were assigned to each of the three conditions, with ten participants in each transmission chain, and five complete chains in each condition. Participants were assigned to a condition randomly, and chain position was constrained by the current state of completion of the chain to which they were recruited. Participants whose data were excluded were replaced in that chain by a new participant, prior to further recruitment to that particular chain. The study was approved by the University of Stirling General University Ethics Panel (approval reference number: GUEP 295), and every participant gave written consent before the experiment took place.

#### The experimental task

The task completed by the participants was a landscape searching challenge loosely based on the game ‘Battle-ships’. The task was run using PsychoPy 1.84.2<sup>44</sup>, on either a desktop computer or a Microsoft Surface tablet running Windows 10.

Participants were presented with a series of ten  $16 \times 16$  grids, in each of which three  $3 \times 3$  targets (ships) could be found. Targets were randomly placed within the full grid (see Fig. 1). Participants could search the grid for targets by clicking on (desktop), or touching (tablet), any of the grid squares. When part of a target was found, these were scored as “hits”, and selection of any of the non-target squares in the  $16 \times 16$  grid were scored as “misses”. When a selection was made, the colour of the square changed from the grey of the grid to green for a hit, and red for a miss. Selection of a hit was accompanied by an explosion sound. The participant’s primary goal was to maximise their score by finding as many hits as possible, across the ten grids.

Participants used up one of their limited search attempts for every grid square they selected which did not contain part of a target. In contrast, when participants selected a target grid square, the number of search attempts did not reduce, and they were instead awarded one point. This scoring system was intended to capture the relative payoffs typically expected from real-world subsistence activity, with fruitless search being costly, and success profitable. A maximum score of 27 points was available in each grid (nine for every full target), and participants were allotted nine search attempts in which to find them. This meant that the number of grid squares revealed varied depending on how many hits and misses were found in that grid. For example, if all 27 hits (i.e., three complete targets) were found without any error, the score for that grid would be 27, with nine search attempts remaining and exactly 27 grid squares revealed. If no targets were found, the score for that grid would be zero, with all search attempts used and only nine (miss) grid squares revealed.

The first participant in each chain received no information about the value of any of the grid squares, other than that generated by their own search. All other participants in the chain received some information about their cultural parent’s search outcome for each grid.

The same pre-specified reward landscape was kept consistent for each grid completed by a particular chain, such that the results of the cultural parent’s search held true for the cultural offspring. Depending on the condition,

participants saw some or all of the search outcome obtained by their cultural parent (including whether each revealed grid square was a hit or a miss). The nature of the information varied by condition as detailed below:

**Full Information:** With the exception of the first in each chain, each participant was shown the entire search outcome obtained by their cultural parent. Thus, cultural offsprings received information about the value of between nine and 36 grid squares for each grid.

**Intentional Information:** At the end of their search, each participant in this condition was asked to select a subset of the hits and misses they had revealed during the search stage (three grid squares). Thus, the information that each participant received comprised the value of these three squares, as selected by their cultural parent in the chain.

**Inadvertent Information:** A randomly selected subset (three) of the grid squares revealed during each participant's search was passed on to the next participant in the chain. Thus, as in the Intentional condition, participants in this condition received information about the value of three grid squares. This condition was therefore intended to be analogous to using “public information”<sup>19</sup>, available as a by-product of another's activity rather than through communicative intent.

## Procedure

Participants gave written consent before being pseudo-randomly assigned a condition and chain position according to the constraints detailed in the Participants section.

Participants were provided with thorough instructions on screen before beginning the experiment. These included picture examples of an unselected grid square, a selected hit, a selected miss, and a full 3 × 3 target. Participants were also given instructions about the constraints on their search activity for any given grid, i.e., that the program would advance to a new grid when either all nine search attempts had been used up, or when all 27 hits (i.e., three complete targets) had been selected (see description of experimental task above).

If a participant was in generation 1 (i.e., the first participant in a chain of ten), then they received no information about the location of hits or misses. Participants in all other chain positions received some information about the grid, as detailed in the description of the experimental task. In all conditions, participants were told that they would receive some information from a previous participant's performance, and that they could use this information however they wished. Further details about the source of the information were not explicitly disclosed to participants; these instructions remained the same across conditions.

After participants had completed their own search, in the Intentional condition, they were also required to make decisions about the information to be transmitted to the next participant in the chain. Participants in this condition were therefore made aware that the information was to be sent to a future participant who would be presented with the same task they had just completed. Following each grid search, they were asked to select three of the grid squares from the selections that they had made for any given grid.

After all ten grids had been completed, participants were shown their final score (total hits selected across all ten grids). Participants were then fully debriefed regarding the aims and hypotheses of the experiment, and given the opportunity to ask any questions.

## Experiment 2: Children

### Participants

Participants were recruited from six different primary schools in the north of Scotland, with the minimum number of children collected from one school being 13, and the maximum number being 46. We recruited 233 child participants aged between 5 years, 0 months and 10 years, 11 months. Eight of these participants were excluded while recruitment was ongoing, and their assigned position in the relevant chain was therefore re-assigned to a subsequent volunteer. These exclusions were made as a result of experimenter error, technical error, or incomplete/

inconsistent consent details (e.g. regarding participant age). Following recruitment, the sample consisted of 225 children, as per our original target for this study. Seventy-five participants were assigned to each of the three conditions (Intentional, Inadvertent and Full information), with 25 in each age category (5 to 6 years, 7 to 8 years, and 9 to 10 years). Five participants were assigned to each chain, and there were five chains of each age category in each condition. Participants were assigned to a condition randomly, and chain position was constrained by the current state of completion of the chain to which they were recruited.

Once analysis was under way, it became apparent that data from one of the excluded participants had not been fully deleted, and had remained in the source file used to generate the information to be transmitted to cultural offspring. It was this participant's data therefore, rather than their replacement's, which was transmitted. Consequently, this entire chain (consisting of 5- to 6-year-olds, in the Inadvertent condition) was excluded from analyses. These exclusions left 220 participants aged between 5 years, 0 months and 10 years, 11 months (median age = 8 years 0 months, 109 females). A breakdown of age categories is provided in Table S7.

Due to a technical error which was only discovered after data collection had been completed, data from two further participants (both in the 5- to 6-year age category) were not fully recorded by the computer. For both of these participants, the information about their final scores and search attempts remaining had been recorded and therefore they were included in the primary analysis. However, details of the specific grid squares they selected during their search were not recorded. These participants were therefore excluded from the relevant subsections of the analysis which required this information.

The study was approved by the University of Stirling General University Ethics Panel (approval reference number: GUEP 557). Formal consent to conduct the study was obtained from The Highland Council education manager, and all head teachers gave consent for the experimenter to enter the school. Signed parental consent was obtained prior to testing which confirmed that children were permitted to take part in the experiment, and that this could be done during class time. Participants and their teachers also gave verbal consent before the participant left their classroom to participate in the study.

## The experimental task

The task completed by the children was intended to be as similar as possible to the one presented to the adults, described previously. However, we did make several changes to the task and procedure, detailed here. Firstly, we reduced the number of grids completed by each participant, from ten (adults) to five (children), in order to limit fatigue effects, and to reduce the amount of time that children spent out of their classroom. We also reduced the length of transmission chains, informed by the results of Experiment 1 (Adults), in which we had found that scores reached a plateau from around chain position six. We therefore reduced the length of chains from ten (Adults) to five (Children) participants.

Other changes were made to the task to make it more appealing to children (framing the task as a space exploration game), and to make the instructions easier for the children to follow (e.g. "previous participant" was changed to "previous space explorer"). The task stimuli were therefore described as "spaceships" throughout, rather than "battleships", and the colour scheme of the task was changed to make it brighter, and more consistent with the space theme (the colour of hits was changed from green to purple, and the background from solid grey to a colourful space-themed scene). In addition, the sound associated with a hit was changed from an explosion noise to an alien blip. The adapted task was run using PsychoPy 1.84.2<sup>44</sup>, on either a Microsoft Surface tablet, or a Lenovo Yoga 520 touchscreen laptop in tablet mode, both running Windows 10.

The scoring system was the same as for the task used by the adult participants (i.e. three  $3 \times 3$  targets to be found, concealed within in a  $16 \times 16$  grid, and participants allotted nine search attempts in which to find them). The three conditions were exactly as per Experiment 1 (Adults), i.e. Full Information; Intentional Information; and Inadvertent Information.

## Procedure

Following the consent procedure, participants were pseudo-randomly assigned a condition and chain position according to the constraints detailed in the Participants section. Before completing the experimental task, participants were given a brief (5–10 min) test of mental state understanding adapted from the ‘Yoni task’<sup>45</sup>. This task was run using PsychoPy v1.84.2<sup>44</sup>, on either a Microsoft Surface tablet, or a Lenovo Yoga 520 touchscreen laptop in tablet mode, both using Windows 10. We had originally intended to use this measure to assess whether children’s socio-cognitive reasoning was related to their success on the task. However, children’s performance was close to ceiling in all age groups. Furthermore, the group and condition differences that were identified were not consistent with expectations based on previous literature on children’s first- and second-order mental state reasoning (e.g., we found no clear age trend, and children in all age categories performed better on test items purportedly measuring second-order reasoning, compared with those classed as first-order). There were therefore reasons to doubt the validity of this test, so we did not use it as a measure of mental state understanding. Further details of the test, including the results of our analysis, can be found in Supplementary Methods.

Participants were provided with thorough instructions on screen before beginning the main task. These instructions were adapted from the adult task instructions to be easier for children to understand. These instructions included picture examples of an unselected grid square, a selected hit, a selected miss, and a full  $3 \times 3$  target. Participants were also given instructions about the constraints on their search activity for any given grid, i.e., that the program would advance to a new grid when either all nine search attempts had been used up, or when all 27 scoring tiles (i.e., three complete targets) had been selected (see previous description of experimental task). Following the instructions, participants were presented with an understanding check quiz (see Fig. S4) that ensured they could identify a  $3 \times 3$  target, recall how many full targets were hidden, and, with the exception of the first participant in the chain, correctly identify hit and miss squares as they would be depicted in a transmission from a previous participant. Children were given the quiz as soon as the instruction screens ended and were encouraged to discuss the instructions until they had correctly answered all of the questions. The quiz was repeated at the end of the task to ensure they were still aware of the instructions. Children could refer to the quiz for reminders of the shape of the targets throughout the task. In the event that participants stalled their search due to distraction, the experimenter gave them a prompt (such as, “Find the alien spaceship.”). After all five grids had been searched, participants were shown their final score (total hits selected across all five grids). Participants were given the opportunity to choose a reward sticker for participating.

## Methods of analysis

All linear models were performed using R<sup>46</sup> and lme4<sup>47</sup>. Graphs were created using ggplot2 performed on R Studio<sup>48</sup>.

Problems with model fit (e.g. singular fit error) were addressed by removing random variance, however models with the maximum random effects structure<sup>49</sup> were considered first. Post-hoc comparisons were done using the emmeans package<sup>50</sup>. Tests were two-tailed, and  $p$ -values of  $< 0.05$  were taken as statistically significant. Emmeans using the Tukey adjustment was used for multiple comparisons, and emtrends was used to assess interactions between conditions in effects of chain position.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-62734-6>.

**Author contributions** C.A.C., G.M. and M.A. conceived and designed the study. M.A. wrote the code for the experimental task. G.M. collected the data. C.A.C., M.A. and E.R. advised on statistical analysis and presentation of results. G.M. and C.A.C. conducted analyses and produced figures. G.M. and C.A.C. wrote the manuscript with contributions from M.A and E.R.

**Funding** This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 648841 RATCHETCOG ERC-2014-CoG.

**Data availability** Source data are available at: <https://osf.io/d8jpb/files/osfstorageCode> for reproducing the analyses can be found at: <https://osf.io/d8jpb/files/osfstorage>.

**Code availability** Code for reproducing the analyses can be found at: <https://osf.io/d8jpb/files/osfstorage>.

## Declarations

**Competing interests** The authors declare no competing interests.

**Ethical approval** Experiment 1 (Adults) was approved by the University of Stirling General University Ethics Panel with approval reference number GUEP 295, and all participants gave full informed, written consent before the experiment took place. Experiment 2 (Children) was approved by the University of Stirling General University Ethics Panel with approval reference number GUEP 557. Formal consent to conduct Experiment 2 was obtained from the Highland Council education manager, and all head teachers gave consent for the experimenter to enter the school. Full informed, written consent was obtained from each child's parent or legal guardian prior to testing, confirming that the child was permitted to take part in the experiment, and that this could be done during class time. Participants and their teachers also gave verbal consent before the participant left the classroom to participate in the study. All experiments were performed in accordance with relevant guidelines and regulations and in accordance with the Declaration of Helsinki.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Hill, K., Barton, M. & Hurtado, A. M. The emergence of human uniqueness: Characters underlying behavioral modernity. *Evol. Anthropol.* **18**, 187–200. <https://doi.org/10.1002/evan.20224> (2009).
2. Boyd, R., Richerson, P. J. & Henrich, J. The cultural niche: Why social learning is essential for human adaptation. *Proc. Natl. Acad. Sci. USA*. **108** (Suppl 2), 10918–10925. <https://doi.org/10.1073/pnas.1100290108> (2011).
3. Vaesen, K. & Houkes, W. Is human culture cumulative? *Curr. Anthropol.* **62**, 218–238. <https://doi.org/10.1086/714032> (2021).
4. Heyes, C. *Cognitive Gadgets* (Harvard University Press, 2018).
5. Henrich, J. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton University Press, 2015).
6. Tomasello, M. *Becoming Human a Theory of Ontogeny* (Belknap, 2019).
7. Laland, K. N. *Darwin's Unfinished Symphony: How Culture Made the Human Mind* (Princeton University Press, 2017).
8. Heyes, C. & Moore, R. in *The Oxford Handbook of Cultural Evolution*. 230–240 (eds Kendal, R., Tehrani, J. & Kendal, J.) (Oxford University Press, 2025).
9. Nichols, R., Moll, H. & Mackey, J. L. How to create a cultural species: Evaluating three proposals. *Philosophical Psychol.* <https://doi.org/10.1080/09515089.2021.1929915> (2021).
10. Tomasello, M. *The cultural origins of human cognition* (Harvard University Press, 1999).
11. Laland, K. N., Odling-Smee, J. & Myles, S. How culture shaped the human genome: Bringing genetics and the human sciences together. *Nat. Rev. Genet.* **11**, 137–148. <https://doi.org/10.1038/nrg2734> (2010).
12. Sasaki, T. & Biro, D. Cumulative culture can emerge from collective intelligence in animal groups. *Nat. Commun.* **8**, 15049. <https://doi.org/10.1038/ncomms15049> (2017).
13. Jesmer, B. R. et al. Is ungulate migration culturally transmitted? Evidence of social learning from translocated animals. *Science* **361**, 1023–1025 (2018).
14. Mesoudi, A. & Thornton, A. What is cumulative cultural evolution? *Proc. R. Soc. London B* <https://doi.org/10.1098/rspb.2018.0712> (2018).

15. Pinker, S. The cognitive niche: Coevolution of intelligence, sociality, and language. *Proc. Natl. Acad. Sci. USA*. **107** (Suppl 2), 8993–8999. <https://doi.org/10.1073/pnas.0914630107> (2010).
16. Osiurak, F. & Reynaud, E. The elephant in the room: What matters cognitively in cumulative technological culture. *Behav. Brain Sci.* **43**, e156. <https://doi.org/10.1017/S0140525X19003236> (2019).
17. Derex, M., Bonnefon, J. F., Boyd, R. & Mesoudi, A. Causal understanding is not necessary for the improvement of culturally evolving technology. *Nat. Hum. Behav.* **3**, 446–452 (2019).
18. Harris, J. A., Boyd, R. & Wood, B. M. The role of causal knowledge in the evolution of traditional technology. *Curr. Biol.* **31**, 1798–1803. <https://doi.org/10.1016/j.cub.2021.01.096> (2021). e1793.
19. Danchin, E., Giraldeau, L. A., Valone, T. J. & Wagner, R. H. Public information: from nosy neighbors to cultural evolution. *Science* **305**, 487–491. <https://doi.org/10.1126/science.1098254> (2004).
20. Smith, J. W., Benkman, C. W. & Coffey, K. The use and misuse of public information by foraging red crossbills. *Behav. Ecol.* **10**, 54–62 (1999).
21. Templeton, J. J. & Giraldeau, L. Vicarious sampling: The use of personal and public information by starlings foraging in a simple patchy environment. *Behav. Ecol. Sociobiol.* **38**, 105–114 (1996).
22. Whiten, A. et al. Cultures in chimpanzees. *Nature* **399**, 682–685 (1999).
23. Thornton, A. & Raihani, N. J. Identifying teaching in wild animals. *Learn. Behav.* **38**, 297–309. <https://doi.org/10.3758/LB.38.3.297> (2010).
24. Whiten, A., Horner, V. & Marshall-Pescini, S. Cultural panthropology. *Evol. Anthropol.: Issues News Rev.* **12**, 92–105. <https://doi.org/10.1002/evan.10107> (2003).
25. Thornton, A. & McAuliffe, K. Teaching in wild meerkats. *Science* **313**, 227–229 (2006).
26. Franks, N. R. & Richardson, T. Teaching in tandem-running ants. *Nature* **439**, 153. <https://doi.org/10.1038/439153a> (2006).
27. Kline, M. A. How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behav. Brain Sci.* **38**, e31. <https://doi.org/10.1017/S0140525X14000090> (2015).
28. Caldwell, C. A., Renner, E. & Atkinson, M. Human teaching and cumulative cultural evolution. *Rev. Philos. Psychol.* <https://doi.org/10.1007/s13164-017-0346-3> (2017).
29. Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E. & Kendal, R. L. Human cumulative culture: A comparative perspective. *Biol. Rev.* **89**, 284–301. <https://doi.org/10.1111/brv.12053> (2014).
30. Caldwell, C. A. & Millen, A. E. Social learning mechanisms and cumulative cultural evolution: Is imitation necessary? *Psychol. Sci.* **20**, 1478–1483 (2009).
31. Zwirner, E. & Thornton, A. Cognitive requirements of cumulative culture: Teaching is useful but not essential. *Sci. Rep.* **5**, 16781. <https://doi.org/10.1038/srep16781> (2015).
32. Morgan, T. J. et al. Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat. Commun.* **6**, 6029. <https://doi.org/10.1038/ncomms7029> (2015).
33. Lucas, A. et al. (ed, J.) The value of teaching increases with tool complexity in cumulative cultural evolution. *Proc. R. Soc. Lond. B* **287** 20201885 <https://doi.org/10.1098/rspb.2020.1885> (2020).
34. Lomas, M., Psychogiou, L., Caldwell, C. A., Happé, F. & Thornton, A. The cognitive foundations of teaching. *Evol. Anthropol.* **35**, e70032. <https://doi.org/10.1002/evan.70032> (2026).
35. Wellman, H. M., Cross, D. & Watson, J. Meta-analysis of theory-of-mind development: The truth about false belief. *Child Dev.* **72**, 655–684. <https://doi.org/10.1111/1467-8624.00304> (2001).
36. Perner, J. & Wimmer, H. John thinks that Mary thinks that. : Attribution of second-order beliefs by 5- to 10-year-old children. *J. Exp. Child Psychol.* **39**, 437–471 (1985).
37. Ruffman, T. Do children understand the mind by means of a simulation or a theory? Evidence from their understanding of inference. *Mind Lang.* **11**, 388–414 (1996).
38. Blakey, K. H. et al. Development of strategic social information seeking: Implications for cumulative culture. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0256605> (2021).
39. Blakey, K. H., Renner, E., Atkinson, M., Rafetseder, E. & Caldwell, C. A. Children transition from simple associations to explicitly reasoned social learning strategies between age four and eight. *Sci. Rep.* <https://doi.org/10.1038/s41598-022-09092-1> (2022).
40. Morin, O. *ie (Foundations of Human Interaction)* (OUP, 2015).
41. Scott-Phillips, T. Biological adaptations for cultural transmission? *Biol. Lett.* <https://doi.org/10.1098/rsbl.2022.0439> (2022).
42. Brune, M. & Brune-Cohrs, U. Theory of mind - evolution, ontogeny, brain mechanisms and psychopathology. *Neurosci. Biobehav. Rev.* **30**, 437–455. <https://doi.org/10.1016/j.neubiorev.2005.08.001> (2006).
43. Warneken, F. & Tomasello, M. Altruistic helping in human infants and young chimpanzees. *Science* **311**, 1301–1303. <https://doi.org/10.1126/science.1121448> (2006).
44. Peirce, J. et al. PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods.* **51**, 195–203. <https://doi.org/10.3758/s13428-018-01193-y> (2019).

45. Shamay-Tsoory, S. G. & Aharon-Peretz, J. Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia* **45**, 3054–3067. <https://doi.org/10.1016/j.neuropsychologia.2007.05.021> (2007).
46. R Core Team. R: A language and environment for statistical computing. *R Foundation Stat. Computing* <https://www.R-project.org/> (2022).
47. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v067.i01> (2015).
48. RStudio Team. RStudio: Integrated development environment for R. *RStudio, PBC* <http://www.rstudio.com/> (2022).
49. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* <https://doi.org/10.1016/j.jml.2012.11.001> (2013).
50. Lenth, R. V. emmeans: Estimated marginal means, aka least-squares means. *R Package Version 1.7.5* <https://CRAN.R-project.org/package=emmeans> (2022).

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

## Authors and Affiliations

Gemma Mackintosh<sup>1</sup> · Mark Atkinson<sup>2</sup> · Elizabeth Renner<sup>3</sup> · Christine A. Caldwell<sup>1</sup> 

✉ Gemma Mackintosh  
gemma.r.mackintosh@gmail.com

✉ Christine A. Caldwell  
c.a.caldwell@stir.ac.uk

<sup>1</sup> Division of Psychology, University of Stirling, Stirling, UK

<sup>2</sup> Present address: Psychology, School of Social Sciences, Heriot-Watt University, Edinburgh, Scotland

<sup>3</sup> Present address: School of Psychology, Northumbria University, Newcastle upon Tyne, England