

1 **Little evidence that Eurasian jays (*Garrulus glandarius*) protect their caches by**  
2 **responding to cues about a conspecific's desire and visual perspective**

3  
4  
5  
6  
7

8 Piero Amodio<sup>1,2,\*</sup>, Benjamin G. Farrar<sup>1,3</sup>, Christopher Krupenye<sup>4,5</sup>, Ljerka Ostojić<sup>1,6</sup> and Nicola  
9 S. Clayton<sup>1</sup>

10  
11

12 <sup>1</sup> Department of Psychology, University of Cambridge, Cambridge, UK

13 <sup>2</sup> Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn,  
14 Napoli, Italy

15 <sup>3</sup> Institute for Globally Distributed Open Research and Education (IGDORE)

16 <sup>4</sup> Department of Psychological & Brain Sciences, Johns Hopkins University, Baltimore, USA

17 <sup>5</sup> Department of Psychology, Durham University, Durham, UK

18 <sup>6</sup> Department of Psychology, Faculty of Humanities and Social Sciences, University of Rijeka,  
19 Rijeka, Croatia

20  
21

22 \* corresponding author: [piero.amodio@cantab.net](mailto:piero.amodio@cantab.net)

23  
24

25

26 **Short title**

27 Do jays protect their caches by responding to others' social cues?

28

29 **Key words**

30 Corvids, Eurasian jay, theory of mind, desires, perspective, caching, replication

31

32 **ABSTRACT**

33

34 Corvids appear to be capable of adjusting their behaviour according to another's perspective,  
35 knowledge and desire. For example, Eurasian jays have been found to employ a variety of  
36 cache protection strategies to minimise cache loss by responding to cues about the visual  
37 perspective or current desire of an observing conspecific. However, it is not known whether  
38 these jays (or any other corvid) can integrate multiple cues about different mental states and  
39 perform the optimal response accordingly. Across five experiments, we found little evidence  
40 that our Eurasian jays responded to either the visual perspective or current desire of another  
41 agent. In Experiments 1 and 2 we investigated whether Eurasian jays can limit the risk of  
42 cache loss by responding simultaneously to cues about the desire and perspective of a  
43 potential conspecific pilferer. Building on established paradigms, we used opaque and clear  
44 barriers to manipulate the observer's visual access to cache locations, and specific satiety to  
45 manipulate the observer's desire towards different types of food. Across both experiments the  
46 jays' caching pattern provided no evidence that they could integrate information about the  
47 observer's desire and perspective. Moreover, the results were also inconsistent with the  
48 previously reported effects that jays protect their caches by responding to either the visual  
49 access or specific satiety of the observer independently. To gain further insight into these  
50 unexpected results, we conducted three more experiments. In Experiments 3 and 4, we  
51 attempted to replicate the previous finding that Eurasian jays prefer to cache behind an  
52 opaque barrier over a clear barrier when observed by a conspecific than when caching in  
53 private. In Experiment 5, we further investigated the previous finding that jays preferentially  
54 cache a type of food that had been eaten to satiety by a conspecific over a food that had not  
55 been eaten by the conspecific. Experiments 3, 4 and 5 found no significant effects in the  
56 direction of the previously reported effects, questioning their robustness. We conclude by  
57 discussing the implications of our study for the field of corvid cognition and highlight several  
58 key issues that affect the reliability of comparative cognition research.

59

## 60 INTRODUCTION

61

62 Theory of mind is thought to provide a causal and flexible cognitive framework that allows  
63 humans to navigate complex social interactions (FeldmanHall and Shenhav, 2019; Tamir and  
64 Thornton, 2018). Through this framework, different observable social cues can be used to  
65 infer otherwise imperceptible mental states (e.g., perspectives, desires, knowledge, beliefs),  
66 such that the behaviour of other individuals can be interpreted, predicted, and manipulated  
67 based on an interplay of different mental states (Baker et al., 2017; Bartsch and Wellman,  
68 1989; Bennett and Galpert, 1992). In humans, theory of mind is thought to emerge as a  
69 stepwise process, with a meta-representational framework in place by at least 5 years of age  
70 (Wellman, 2018; Wellman and Liu, 2004). However, even before they develop this meta-  
71 representational theory of mind, infants can already respond to multiple pieces of social  
72 information in an integrated manner (e.g., Moses et al., 2001; Repacholi et al., 2014). Multiple  
73 accounts exist, arguing that they may do so with or without representing others' mental states  
74 as such (Apperly and Butterfill, 2009; Carruthers, 2020; Scott and Baillargeon, 2017;  
75 Southgate, 2019). At a minimum, though, infants seem to implicitly register social cues and  
76 exhibit adaptive responses accordingly (Apperly and Butterfill, 2009; Butterfill and Apperly,  
77 2013). Consequently, different mechanisms may allow individuals to integrate information  
78 from multiple social cues.

79

80 Inspired by Premack and Woodruff (1978), the past few decades have seen growing efforts to  
81 investigate whether non-human animals also possess something akin to theory of mind (e.g.,  
82 primates: Buttelmann et al., 2017; Drayton and Santos, 2014; Flombaum and Santos, 2005;  
83 Hare et al., 2000, 2001; Kano et al., 2019; Krupenye and Call, 2019; dogs: Horowitz, 2011;  
84 Maginny and Grace, 2014; corvids: Bugnyar et al., 2016; Dally et al., 2006; Emery and  
85 Clayton, 2001; Ostojić et al., 2013; Shaw and Clayton, 2012). This issue is significant, given  
86 foundational debates in cognitive science about whether minds can represent mental state  
87 concepts in the absence of language as well as a long-standing interest in the evolution of the  
88 key cognitive traits that make us human (e.g., Penn et al., 2008). However, within this line of  
89 research, studies most often focus on testing whether a given species/group has the ability to  
90 attribute one specific type of mental state, i.e., exclusively focusing on belief, or exclusively  
91 focusing on desire, within a single study. As a result, very little is known about whether other  
92 species can – like humans (e.g., Baker et al., 2017) – integrate multiple social cues that  
93 correlate with others' mental states and exhibit appropriate responses accordingly. This

94 question is far from trivial: in real life scenarios, an individual's behaviour is likely to result  
95 from the interplay of multiple factors (e.g., their perspective, desires and previous knowledge)  
96 that can be indirectly perceived simultaneously by another agent during social interactions.  
97 Therefore, by focusing on non-human animals' ability to respond to a single social cue at a  
98 time, comparative psychologists may overlook a crucial aspect of social cognitive complexity.  
99 Additionally, investigating whether and how other species may integrate different social cues  
100 can help shift the focus of social cognition studies in animals away from binary questions –  
101 e.g., does species X understand false beliefs? – towards more process-oriented and nuanced  
102 research questions, such as: what are the relative contributions of mechanisms A and B to  
103 how species X perform behaviour Y? (Buckner, 2013; Heyes, 2015). In doing so, this work may  
104 also identify mechanisms of varying complexity that operate in the absence of language and  
105 may feed into uniquely human social cognition.

106

107 Corvids are a group of large-brained birds that are hypothesised to have evolved  
108 sophisticated cognitive abilities independently from primates (Clayton and Emery, 2015;  
109 Emery and Clayton, 2004; Güntürkün and Bugnyar, 2016; Osvath et al., 2014; Seed et al.,  
110 2009). This group represents a good model for this line of research because they might be  
111 capable of responding – independently – to social cues correlating with different types of  
112 mental states (perspective: Bugnyar et al., 2016; Dally et al., 2004, 2005; Legg et al., 2016;  
113 Legg and Clayton, 2014; Shaw and Clayton, 2013; Stulp et al., 2009; desires: Ostojić et al.,  
114 2013, 2016, 2017; knowledge: Bugnyar and Heinrich, 2005; Emery and Clayton, 2001). In  
115 particular, previous research has reported that Eurasian jays (*Garrulus glandarius*) may be  
116 able to adjust their behaviour according to cues that correlate with the perspective and  
117 current desire of a conspecific. In the presence of a conspecific competitor, these jays have  
118 been reported to preferentially cache food in less visible locations (e.g., behind barriers, at  
119 distance), or in non-noisy substrates, which has been interpreted as a potential response to  
120 the visual (Legg et al., 2016; Legg and Clayton, 2014) or acoustic perspective (Shaw and  
121 Clayton, 2013) of a potential pilferer. In parallel, research investigating food-sharing  
122 behaviour found that male Eurasian jays change the type of food shared with their female  
123 partner, depending on which food the female has been sated on, and therefore which she  
124 desires (Ostojić et al., 2013, 2014, 2016). In a recent study, a similar response to another's  
125 satiety has also been reported in the context of caching, whereby Eurasian jays and Western  
126 scrub-jays preferentially cached food that an observer, and thus potential pilferer, was sated  
127 on (Ostojić et al., 2017). Notably, the effect reported in this study was unlikely to be based on

128 theory of mind, because the caching jays showed this effect also when they did not know what  
129 food the observer was pre-fed on and when the only cue available was the observer's  
130 behaviour during the caching event itself (Ostojić et al., 2017). Thus, taken together, these  
131 studies seem to indicate that Eurasian jays employ a variety of cache protection strategies to  
132 limit the risk of cache loss, by responding to cues correlating with the perspective *or* current  
133 desire of a potential pilferer. Although this evidence is not sufficient to demonstrate mental  
134 state attribution, nor to pinpoint the exact underlying cognitive mechanism, it does suggest  
135 that corvids are capable of behaving flexibly on the basis of different types of social cues.

136

137 The aim of the current study was to investigate whether Eurasian jays can integrate multiple  
138 cues that correlate with different types of mental states to solve social problems. Previous  
139 work has to some extent already, at least implicitly, tested whether animals can integrate  
140 perspective cues from different sensory modalities (corvids: Shaw and Clayton, 2013; Stulp et  
141 al., 2009; primates: Bray et al., 2014; Santos et al., 2006). While these studies investigate the  
142 ability to integrate different cues regarding the same type of mental state, namely another  
143 individual's perspective in assessing the significance of visual and acoustic perceptual cues,  
144 our study concerns the integration of cues correlating with different types of mental states,  
145 namely perspectives and desires. Building on previous studies in the caching context, we  
146 tested whether jays can integrate information about a conspecific's perspective and current  
147 desire to selectively protect those caches that are at most risk of being pilfered. To this end,  
148 we first conducted two experiments in which we manipulated the observer's visual access to  
149 caching locations *and* its current desire for different foods. We measured the caching pattern  
150 of the birds when viewed by this observer.

151

152 Our manipulation of the observer's perspective followed the procedure by Legg and Clayton  
153 (2014). The authors gave Eurasian jay cachers access to two locations, one that could and one  
154 that could not be seen from an adjacent compartment and thus by an observer bird (when  
155 present). One caching tray was positioned behind an opaque barrier (*out-of-view* tray), and  
156 the other behind a transparent barrier (*in-view* tray). Legg and Clayton's experiment  
157 encompassed three conditions: jays could cache when an observer bird was housed in the  
158 adjacent compartment – the observer was either a higher ranked individual (*Observed by*  
159 *dominant* condition) or a lower ranked individual (*Observed by subordinate* condition) – or  
160 when no conspecific was present (*Private* condition). The authors compared the jays' caching  
161 pattern in the *Private* condition with that in the *Observed* conditions (the two observed

162 conditions were merged together), and found that jays cached a higher proportion of food  
163 items in the *out-of-view* tray in the *Observed* than in the *Private* conditions.

164

165 Our manipulation of the observer's desire for different foods followed the procedure by  
166 (Ostojić et al. (2017)). The authors investigated whether Eurasian jays and California scrub-  
167 jays can protect their caches by preferentially caching the type of food that the observer was  
168 not currently motivated to pilfer. A cacher and an observer jay were housed in adjacent  
169 testing compartments. In the pre-feeding phase, the observer could feed to satiety on a  
170 specific type of food: maintenance diet in the baseline trial, and either food A or B in the two  
171 test trials. This procedure subsequently reduces the individual's motivation for eating and  
172 caching that specific food (but not different kinds of food), a phenomenon known as 'specific  
173 satiety' (Balleine and Dickinson, 1998; Clayton and Dickinson, 1999; Dickinson and Balleine,  
174 1994). Ostojić et al. found that the jays' preference for caching food A over food B was larger  
175 after the observer was sated on food A than after the observer was sated on food B.  
176 Interestingly, this pattern was exhibited not only in the *Seen* Condition – when the cacher had  
177 witnessed which particular food had been provided to the observer – but also in the *Unseen*  
178 Condition – when the cacher had not seen the pre-feeding of the observer – thus indicating  
179 that the observer's behaviour at the time of caching may have played a key role in the  
180 decision-making process of the cacher. Note that, when data were re-analysed for Eurasian  
181 jays only, the effect was still statistically significant in the *Seen* condition – although it was not  
182 statistically significant in the *Unseen* condition (Crosby, 2019).

183

184 In the current study we combined these two protocols. In Experiment 1, cacher jays were  
185 provided with only one type of food – which on one trial was the same food on which the  
186 observer sated, and on another trial was a different food from that on which the observer was  
187 sated – and two caching trays, one that the observer could see and one that the observer could  
188 not see. Thus, here, the jays could choose between two caching locations, one that was in-view  
189 and one that was out-of-view of the conspecific, allowing them to selectively cache food out-  
190 of-view when it was desired by the observer. In Experiment 2, cacher jays were provided with  
191 a single caching tray – which on one trial could be seen by the observer and on another trial  
192 could not be seen by the observer – and two types of food, one of which had previously been  
193 pre-fed to the observer. Thus, here, the jays could choose between two food types, one on  
194 which the observer was sated and one on which the observer was not sated, allowing them to  
195 selectively cache the less desired food when the observer could see them. Consequently, the

196 designs of the two experiments were complementary, such that jays could most effectively  
197 protect their caches by deciding *where* to cache in Experiment 1, and *what* to cache in  
198 Experiment 2.

## 199 200 **RESULTS**

### 201 202 **Integration of cues correlating with others' desire and perspective**

203  
204 In Experiment 1 we tested whether jays can integrate multiple cues to decide *where* to hide  
205 food to protect it from being pilfered. To do so, we manipulated the observer' visual access to  
206 two caching trays through a 'T'-shaped Perspex barrier (henceforth T-barrier). This T-barrier  
207 – which was the same barrier originally designed and used in Legg and Clayton (2014) –  
208 consisted of three plastic panels: one transparent panel forming one arm of the 'T' and two  
209 opaque panels forming the second arm and stem of the 'T' (see Material and Methods). The T-  
210 barrier could be placed around two caching trays (Figure 1) such that an observer could see  
211 the tray behind the transparent arm (*in-view* tray) but could not see the tray behind the  
212 opaque arm (*out-of-view* tray).

213  
214 To ensure that the birds (n=9) were comfortable caching in trays when these were placed in  
215 proximity of each of the two arms of the T-barrier, they initially received two familiarisation  
216 trials in private, in which only a single tray was present. The tray was placed once behind the  
217 opaque and once behind the transparent arm of the barrier. All birds except two reached the  
218 criterion in the familiarisation, i.e., they cached at least one item on each trial, and therefore  
219 proceeded to the test (leading to n=7 for the test).

220  
221 Following the basic design of Ostojić et al. (2017), the test trials comprised a pre-feeding  
222 phase and a caching phase (Figure 1). In the pre-feeding phase, the cacher jay could see a  
223 conspecific (the observer) eat a specific type of food (macadamia nuts, M, or peanuts, P) in an  
224 adjacent indoor compartment. In the subsequent caching phase, the cacher jay was presented  
225 with two caching trays, each placed behind one of the two arms of the T-barrier, and was  
226 allowed to cache while the observer jay was still present in the adjacent compartment. Birds  
227 were tested in two conditions (one trial per condition; Figure 1). In the *Different Food*  
228 Condition, the type of food received by the observer in the pre-feeding phase differed from the  
229 type of food received by the cacher in the caching phase (e.g., M for the observer, and P for the

230 cacher). In the *Same Food* Condition, the observer and the cacher received the same type of  
231 food (e.g., P for the observer, and P for the cacher).

232

233 If the jays can integrate information from the different cues correlating with the observer's  
234 desire and perspective, their caching pattern should meet two predictions. First, their  
235 preference for caching in the *out-of-view* tray should be greater in the *Different Food* condition  
236 than in the *Same Food* condition. This is because it is in the *Different Food* condition that the  
237 observer has a stronger desire toward the cacher's food, such that the caches would be more  
238 at risk from being stolen in this condition. Second, in the *Different Food* condition, the cacher  
239 should exhibit a clear preference for caching in the *out-of-view* tray, therefore in this condition  
240 the amount of caches in the *out-of-view* tray should be higher than expected by chance,  
241 namely if the cacher distributed its caches randomly across the *out-of-view* and the *in-view*  
242 trays.

243

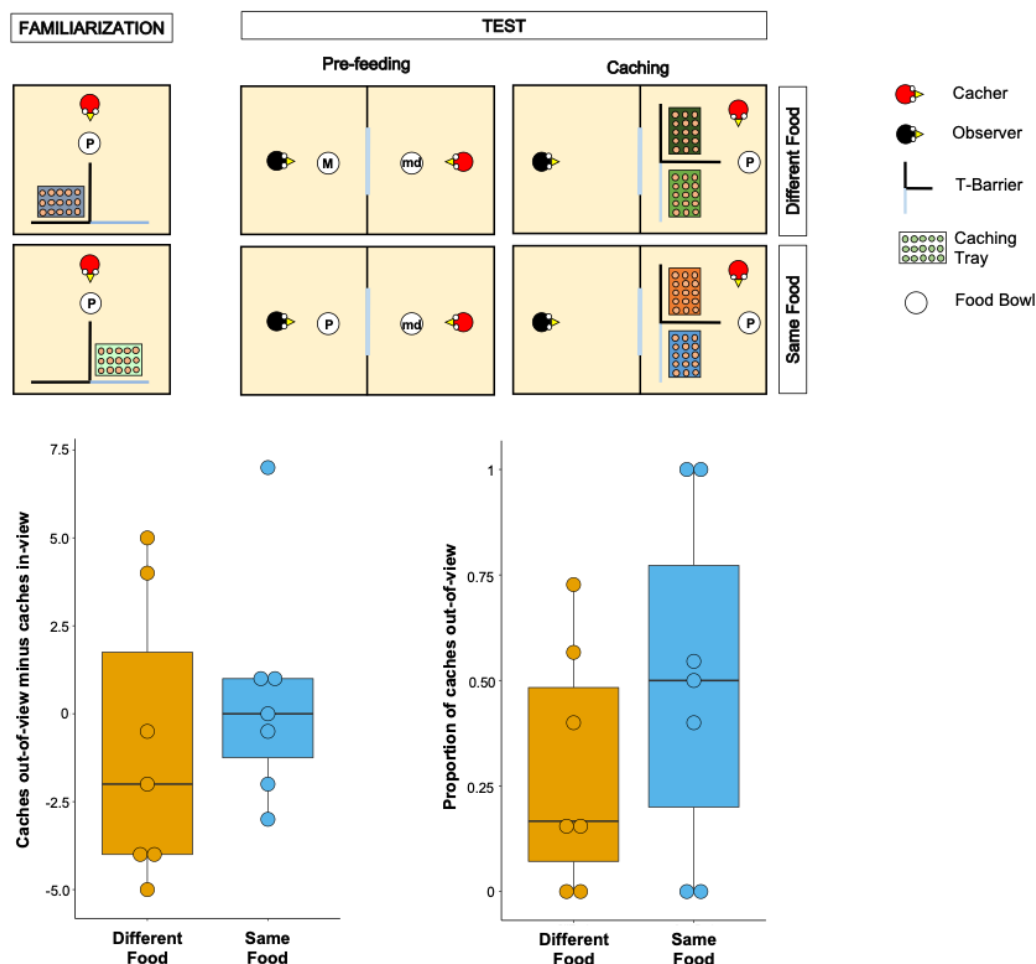
244 The jays' preference to cache in a specific location can be indexed either as a proportion of the  
245 items cached in the *out-of-view* tray (over the items cached in both trays) or as a difference  
246 score of the number of items cached in the *out-of-view* tray minus the number of items cached  
247 in the *in-view* tray. Because robust findings would necessitate that the two different indices do  
248 not lead to drastically different results, we conducted the analyses with both indices (for  
249 details see Analysis in the Methods section).

250

251 The results from the analyses with both indices were consistent. In contrast to the first  
252 prediction, a comparison of the proportion of items cached in the *out-of-view* tray between the  
253 *Different Food* and the *Same Food* conditions did not detect a statistically significant difference  
254 (Wilcoxon signed rank test:  $n=7$ ,  $W=13$ ,  $p=0.21$ ; Figure 1). In contrast to the second prediction,  
255 the proportion of items cached in the *out-of-view* tray was not significantly different from that  
256 expected by chance in the *Different Food* condition (one-sample Wilcoxon signed rank test:  
257  $n=7$ ,  $W=25$ ,  $p=0.11$ ). As an additional analysis, we conducted the same comparison in *Same*  
258 *Food* condition and also detected no statistically significant difference from chance (one-  
259 sample Wilcoxon signed rank test:  $n=7$ ,  $W=25$ ,  $p=1$ ). The same pattern was found when the  
260 difference score – number of items cached in the *out-of-view* tray minus the number of items  
261 cached in the *in-view* tray,  $[\text{Caches}_{\text{out-of-view}} - \text{Caches}_{\text{in-view}}]$  – was used. No statistically  
262 significant difference in the difference score was detected between the *Different Food* and the  
263 *Same Food* conditions (Wilcoxon signed rank test:  $n=7$ ,  $W=9$ ,  $p=0.40$ ; Figure 1) and a



264 comparison of the difference score to chance (i.e., 0) detected no statistically significant  
 265 difference in the *Different Food* condition (one-sample Wilcoxon signed rank test:  $n=7$ ,  $W=-7$ ,  
 266  $p=0.61$ ). The same comparison in *Same Food* condition also detected no statistically  
 267 significant difference from chance (one-sample Wilcoxon signed rank test:  $n=7$ ,  $W=1$ ,  $p=1$ ).  
 268 Notably, the numerical pattern of the jays' caching was in the opposite direction to the  
 269 prediction: jays' caching was more biased towards the *out-of-view* tray in the *Same food* than  
 270 the *Different food* condition (Proportion of items cached *out-of-view*: Median<sub>Same food</sub> = 0.5,  
 271 Median<sub>Different food</sub> = 0.17; Caches out-of-view minus Caches in-view: Median<sub>Same food</sub> = 0,  
 272 Median<sub>Different food</sub> = -2). Consequently, the observed data in Experiment 1 cannot be  
 273 interpreted as supporting the conclusion that the jays were integrating the information from  
 274 multiple cues to protect the caches that are most at risk of being pilfered by a conspecific.  
 275



276

277

278 **Figure 1: Top Panel:** Top-view schematic representation of the set-up and procedure used in Experiment 1. In  
 279 the familiarisation (left), the cacher bird received two trials, one in which the caching tray was placed near the  
 280 opaque arm of the barrier, and one in which the tray was placed near the clear arm of the barrier. In the test,  
 281 trials were composed by a pre-feeding phase (middle) and a caching phase (right). The cacher bird received two

282 trials that differed in the type of food that was provided to the observer bird in the pre-feeding phase. In one trial  
283 (*Different Food Condition*), the food provided to the observer in the pre-feeding phase differed from the food the  
284 cacher bird could subsequently cache. In the other trial (*Same Food Condition*), the food provided to the observer  
285 in the pre-feeding phase was the same as the food the cacher bird could subsequently cache. In the pre-feeding  
286 phase of both trials, the cacher bird was provided with a handful of maintenance diet (md). Note that, for  
287 explanatory purposes, the scheme shows the cacher as being provided with peanuts (P) in the familiarisation  
288 and in the caching phase of the two test trials. However, in the experiment, cacher birds were randomly assigned  
289 to one type of food (either peanuts or macadamia nuts), which was used consistently in the familiarisation and in  
290 the caching phase of all trials. **Bottom Panel:** Box and whisker plots of data in Experiment 1. The plot on the left  
291 shows the difference in the number of items cached in the *out-of-view* tray minus the number of items cached in  
292 the *in-view* tray in the two experimental conditions. The plot on the right shows the proportion of items cached  
293 in the *out-of-view* tray (out of total caches) in the two experimental conditions.

294

295

296 In Experiment 2, we used a complementary design to test whether jays can integrate multiple  
297 cues to decide *which* type of food to hide to protect their caches from being pilfered. To do so,  
298 jays had access to one caching location at a time (either in-view or out-of-view to the  
299 observing conspecific) and two types of food (one on which the observer was sated and one  
300 on which the observer was not sated). Following the general structure of Experiment 1, in the  
301 pre-feeding phase the cacher bird was first able to see an observer eat one particular type of  
302 food (macadamia nuts or peanuts) to satiety. In the subsequent caching phase, the cacher bird  
303 was presented with a single caching tray and the two types of food (macadamia nuts and  
304 peanuts). To manipulate the observer's visual access to the caching location, the tray was  
305 placed behind an 'U-shaped' Perspex barrier (henceforth U-barrier) that consisted of two  
306 lateral panels and one central panel forming two angles of approximately 45° (see Materials  
307 and Methods; Figure 2). The U-barrier was either transparent, thereby allowing the observer  
308 to see the caching location, or opaque, thereby preventing the observer from seeing the  
309 caching location.

310

311 Birds (n=8; for details see Methods) first received two familiarisation trials in private to  
312 ascertain that they were comfortable caching both types of food in a tray placed in proximity  
313 of each of the barriers. All birds except one reached the familiarisation criteria, namely to  
314 cache: i) at least one item of both types of food across the two trials, and; ii) at least one item  
315 (of any type of food) in a tray placed in proximity of both the transparent and the opaque U-  
316 barrier. These birds (n=7) were subsequently tested with the transparent barrier (*In-view*  
317 condition) and with the opaque barrier (*Out-of-view* condition). In each condition, the birds

318 received two trials, one in which the observer was pre-fed one type of food and one in which  
319 it was pre-fed the other type (Figure 2).

320

321 If the jays can integrate information from the different cues available and which should  
322 correlate with the observer's desire and perspective, their caching pattern might be expected  
323 to meet two predictions. First, the jays' preference to cache P when the observer was sated on  
324 P relative to when the observer was sated on M, should be higher in the *In-view* than in the  
325 *Out-of-view* condition. This is because it is in the *In-view* condition that the observer can see  
326 the caching locations such that here the caching bird could protect its caches by caching  
327 preferentially more of the food that the observer is sated on. Second, in the *In-view* condition,  
328 the preference to cache P should be higher when the observer was sated on P than when the  
329 observer was sated on M. As in Experiment 1, both proportion and difference scores were  
330 used as indexes to analyse the birds' preference (for details see Analysis in the Methods  
331 section).

332

333 Again, the results from the analyses using both indices were consistent. A comparison of the  
334 proportion of P cached between the *In-view* and the *Out-of-view* conditions –  
335  $[P_{\text{cached}}/(P_{\text{cached}}+M_{\text{cached}})]_{\text{pre-fed P}} - [P_{\text{cached}}/(P_{\text{cached}}+M_{\text{cached}})]_{\text{pre-fed M}}$  – did not detect a statistically  
336 significant difference (Median<sub>*In-view*</sub> = 0, Median<sub>*Out-of-view*</sub> = 0; Wilcoxon Signed Rank Test: n=7,  
337 W=3, p=0.83; Figure 2). In the *In-view* condition, no statistically significant difference in the  
338 proportion of P cached could be detected between the trials in which observer was sated on  
339 peanuts and the trials in which the observer was sated on macadamia nuts (Median<sub>*Pre-fed P*</sub> =  
340 0.33, Median<sub>*Pre-fed M*</sub> = 0.5; Wilcoxon signed rank test: n=7, W=3, p=0.83). Thus, neither  
341 prediction could be supported. An additional analysis of the same comparison for the *Out-of-*  
342 *view* condition also did not detect a statistically significant difference in the proportion of P  
343 cached between the trials (Median<sub>*Pre-fed P*</sub> = 0, Median<sub>*Pre-fed M*</sub> = 0.06; Wilcoxon signed rank test:  
344 n=7, W=1, p=1).

345

346 The same pattern of results was found when the jays' preference was analysed using the other  
347 index, namely difference scores. No statistically significant difference in the preference to  
348 cache P over M when the observer is sated on P relatively to when the observer is sated on M  
349 – i.e., the difference of difference score:  $[P_{\text{cached}} - M_{\text{cached}}]_{\text{pre-fed P}} - [P_{\text{cached}} - M_{\text{cached}}]_{\text{pre-fed M}}$  – was  
350 detected between the *In-view* and the *Out-of-view* conditions (Median<sub>*In-view*</sub> = -1, Median<sub>*Out-of-*</sub>  
351 <sub>*view*</sub> = 0; Wilcoxon Signed Rank Test: n=7, W=-4, p=0.80; Figure 2). Further, in the *In-view*

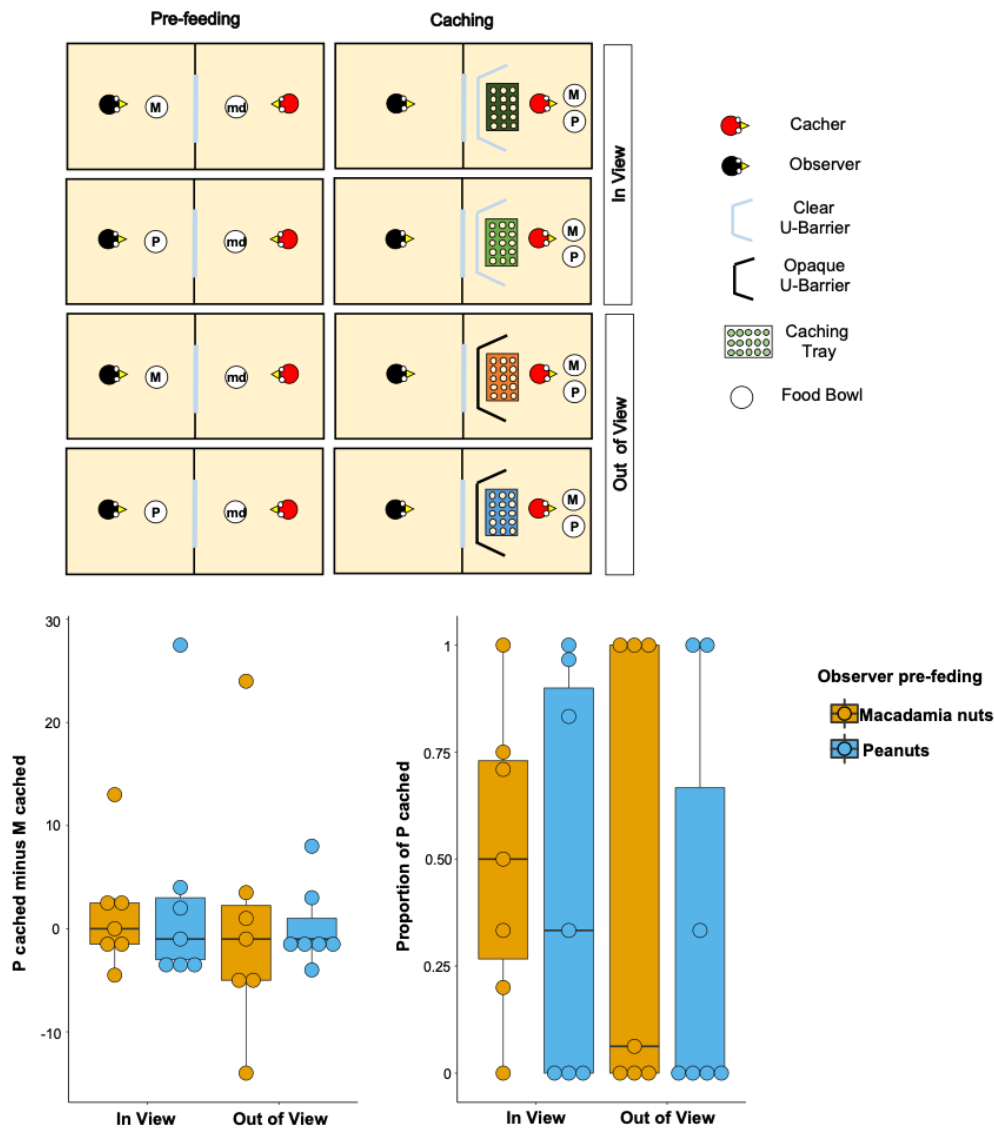
352 condition no statistically significant difference was detected between the trials in which  
353 observer was sated on peanuts and the trials in which the observer was sated on macadamia  
354 nuts ( $\text{Median}_{\text{Pre-fed } P} = -1$ ,  $\text{Median}_{\text{Pre-fed } M} = 0$ ; Wilcoxon signed rank test:  $n=7$ ,  $W=-3$ ,  $p=0.86$ ).  
355 The additional analysis of the same comparison for the *Out-of-view* condition also did not  
356 detect a statistically significant difference between trials ( $\text{Median}_{\text{Pre-fed } P} = -1$ ,  $\text{Median}_{\text{Pre-fed } M} = -$   
357  $1$ ; Wilcoxon signed rank test:  $n=7$ ,  $W=-1$ ,  $p=1$ ).

358  
359 Consequently, just like in Experiment 1, the observed data cannot be interpreted as support  
360 that jays could integrate the information from multiple cues to protect their caches when  
361 these were most at risk of being pilfered by a conspecific. The two experiments thus yielded  
362 consistent results. However, a clear interpretation of the results is impeded by the low power  
363 of the analyses. This is likely due to the small sample size and limited number of trials per  
364 condition of our experiments, two features that – despite being relatively representative of  
365 the research in this area, including the previously published studies on this topic – may have  
366 produced imprecise estimates (Farrar et al., 2020; Farrar and Ostojic, 2019). Therefore, to  
367 strengthen our confidence that Eurasian jays may not be able to integrate multiple cues to  
368 protect their caches, it will be essential for future research to conduct additional studies,  
369 ideally by employing larger sample sizes and procedures that can increase the precision of the  
370 analyses' estimates.

371  
372 Although birds could have used multiple cues to guide their caching decisions in Experiments  
373 1 and 2, they could also have adjusted their caching preference according to just one single  
374 type of cue, i.e., either the cues correlating with the observer's desire or the cues correlating  
375 with the observer's perspective. Both experiments used an experimental manipulation that,  
376 when applied separately, has already been reported – in previous studies – to have elicited a  
377 behavioural response that has been interpreted as a cache-protection strategy. Specifically,  
378 the caching phase of Experiment 1 involved the same procedure and set-up used by Legg and  
379 Clayton (2014), except for the specific types and quantities of food provided to the jays.  
380 Similarly, the *In-view* condition of Experiment 2 and the *Seen* condition of Ostojic et al.  
381 (2017)'s experiment employed the same procedure, with the exception that in the former the  
382 observers could see the caching location through a transparent barrier, whereas in the latter  
383 no barrier was present. However, in contrast to these previous studies, the results obtained in  
384 our experiments did not show a directional caching pattern in the predicted direction in these  
385 situations. Again, this may be a result of the low power of our statistical analyses, or possibly

386 from the greater demands associated with tracking and integrating multiple cues to inform  
387 decision-making. However, the inconsistencies with previous research could also be due to  
388 previously reported effects not being robust enough to form the basis of follow-up studies.  
389 Therefore, we conducted three further experiments to explore the robustness and reliability  
390 of the effects reported by Legg and Clayton (2014) and Ostojić et al. (2017). In Experiments 3  
391 and 4, we attempted a replication of Legg and Clayton (2014)'s findings. In Experiment 3,  
392 birds were tested in one trial per condition, mirroring Experiment 1, while in Experiment 4,  
393 we conducted a complete direct replication of the original study that encompassed two trials  
394 in the *Private* condition and two trials in each of the two *Observed* conditions. Finally, in  
395 Experiment 5, we tested whether the presence or absence of a transparent barrier – i.e., the  
396 minor difference in the set-up between Experiment 2 and Ostojić et al. (2017)'s experiment –  
397 may have affected the Eurasian jays' response in this caching situation.  
398

399



400

401

402 **Figure 2: Top Panel:** Top-view schematic representation of the set-up and procedure used in the test of  
 403 Experiment 2. Trials were composed by a pre-feeding phase (left panels) and a caching phase (right panels). The  
 404 cacher bird received two trials with the transparent U-barrier (*In-view* condition, top panels), and two trials with  
 405 the opaque U-barrier (*Out-of-view* condition, bottom panels). Within each condition, trials differed in the type of  
 406 food (either peanuts, P, or Macadamia nuts, M) that was provided to the observer in the pre-feeding phase. The  
 407 cacher bird was always presented with a handful of maintenance diet (md) in the pre-feeding phase of all trials.

408 **Bottom Panel:** Box and whisker plots of data in Experiment 2. The plot on the left shows the difference in the  
 409 number of peanuts cached minus the number of macadamia nuts cached for each condition; whereas the plot on  
 410 the right shows the proportion of P cached (out of total items cached) in each condition. In the *In-view* condition,  
 411 the observer had visual access to the caching tray, whereas in the *Out-of-view* condition, the observer did not  
 412 have visual access to the caching tray. The colour of the boxes in the plot differs on the basis of the type of food  
 413 that was provided to the observer in the pre-feeding phase: blue denotes that the observer had been pre-fed P  
 414 and orange denotes that the observer had been pre-fed M.

415 **Exploring the robustness of caching strategies based on either the perspective or the**  
416 **current desire of a competitor**

417

418 In Experiment 3, we investigated whether jays use information about an observer's visual  
419 perspective to protect their caches in a simplified version of Legg and Clayton (2014)'s  
420 experiment, i.e., the jays received only one trial in each of the two testing conditions. This  
421 mirrors the procedure in Experiment 1, where the same set-up was used and only one trial  
422 per testing condition was conducted. Following the original study, we presented cacher jays  
423 with two caching trays and manipulated the observer's visual access to cache locations by  
424 using the T-barrier. However, while Legg and Clayton (2014) tested jays in three conditions –  
425 *Observed by Dominant*, *Observed by Subordinate* and *Private* – and gave them two trials in each  
426 condition, in Experiment 3, jays (n=8) received only two trials: one with a conspecific present  
427 in the adjacent compartment (*Observed* Condition) and one with no conspecific present  
428 (*Private* Condition). Seven birds met the inclusion criterion (see Material and Methods for  
429 details).

430

431 The two analyses with the different indices yielded consistent results. The proportion of the  
432 items cached in the *out-of-view* tray was not significantly higher in the *Observed* condition  
433 than in the *Private* condition (Median<sub>Observed</sub> = 0.71, Median<sub>Private</sub> = 0.54; Wilcoxon signed-rank  
434 test, n=7, W=2, p<sub>one-tailed</sub>=0.59). The same pattern was found when the difference score, i.e., the  
435 number of items cached in the *out-of-view* tray minus the number of items cached in the *in-*  
436 *view* tray [Caches<sub>out-of-view</sub> – Caches<sub>in-view</sub>], was analysed. The difference score was not  
437 significantly higher in the *Observed* than in the *Private* condition (Median<sub>Observed</sub> = 1,  
438 Median<sub>Private</sub> = 0.5; Wilcoxon signed-rank test, n=7, W=11, p<sub>one-tailed</sub>=0.84).

439

440 In a subsequent experiment, Experiment 4, we conducted a direct replication of Legg and  
441 Clayton (2014)'s experiment. Here, the design and procedure were identical to those of the  
442 original study (Figure 3; Material and Methods). For this experiment only, we also tested the  
443 same colony of jays that originally participated in Legg and Clayton (2014)'s experiments.  
444 Because these birds had not recently participated in testing using the experimental set-up  
445 employed here and the T-barrier, we first conducted a familiarisation that followed the same  
446 procedure as that used in Experiment 1. Nine birds passed the familiarisation and proceeded  
447 to the test. In this experiment only, we conducted the same analyses as for all other  
448 experiments (i.e., Wilcoxon signed rank tests) but also an additional one, namely the same

449 analysis (permutation tests for paired data) that was also used in Legg and Clayton (2014).  
450 Again, a strong claim of an effect would require consistent results regardless of the analyses  
451 used.

452

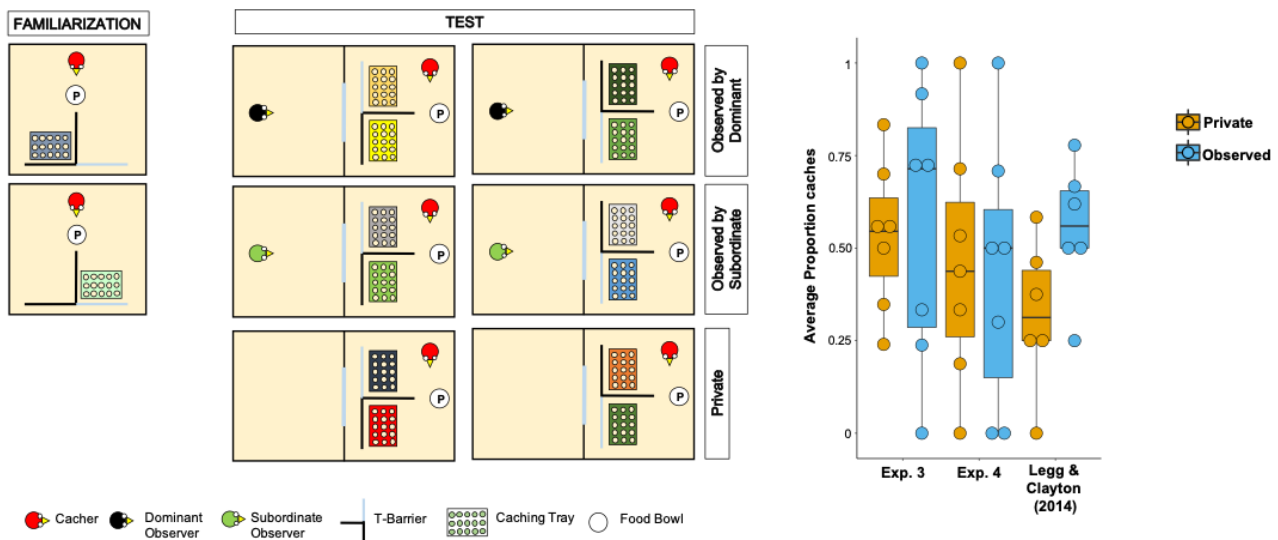
453 In line with the original study, we found that the average number of total items cached across  
454 both trays was not significantly higher when the jays were observed by a conspecific than  
455 when they were in private (Permutation test,  $n=9$ ,  $Z=0.79$ ,  $p=0.43$ ). Two birds cached no items  
456 in any of the *Private* and *Observed* trials, thereby they were excluded from further analyses of  
457 proportion scores because, given their performance, it was not possible to compare the  
458 proportion of items cached in the *out-of-view* tray between conditions. In the same analysis  
459 that was used by Legg and Clayton (2014), the average proportion of items cached in the *out-*  
460 *of-view* tray was not significantly higher in the *Observed* condition than in the *Private*  
461 condition (Median<sub>Observed</sub> = 0.5, Median<sub>Private</sub> = 0.44; Permutation test,  $n=7$ ,  $Z=0.15$ ,  $p_{\text{one-tailed}}=0.56$ ). The same results were found in the two analyses that used the same statistical test  
462 as in the other experiment reported in this study: average proportion of items cached in the  
463 *out-of-view* tray (Wilcoxon signed rank test,  $n=7$ ,  $W=2$ ,  $p_{\text{one-tailed}}=0.59$ ); average difference of  
464 the number of items cached in the *out-of-view* tray minus the number of items cached in the  
465 *in-view* tray (Median<sub>Observed</sub> = 0, Median<sub>Private</sub> = -0.5; Wilcoxon signed rank test,  $n=9$ ,  $W= 13$ ,  
466  $p_{\text{one-tailed}}=0.20$ ).

468

469 Taken together, Experiments 3 and 4 consistently did not detect the effect originally reported  
470 by Legg and Clayton (2014), whereby Eurasian jays adjusted their caching pattern to the  
471 transparency and opaqueness of the barrier around the caching tray specifically when an  
472 observer was present during the caching event (Figure 3). In addition, the results from  
473 Experiments 3 and 4 appear consistent with the negative results from the *Different Food*  
474 condition in Experiment 1.

475





476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

**Figure 3: Left:** Top-view schematic representation of the set-up and procedure used in Experiment 4. In the familiarisation (left panels), the cacher bird received two trials, one in which the caching tray was placed near the opaque arm of the barrier (top panel), and one in which the tray was placed near the clear arm of the barrier (bottom panel). In the test (central and right panel), birds were tested in three conditions: Observed by dominant, Observed by subordinate, Private. In each condition, the cacher received two trials that differed in the orientation of the T-barrier. **Right:** Box and whisker plot of data in Experiments 3, 4, and in Legg and Clayton (2014). The plot shows the average proportion of items cached in the out-of-view tray out of the total number of items cached, in the Private and Observed conditions. Note that data in the Observed by dominant and Observed by subordinate conditions were averaged for Experiment 4 and for Legg and Clayton (2014)'s study.

In Experiment 5, we investigated whether a minor difference in the set-up, i.e., the presence of a transparent barrier, may have caused the inconsistency in the results between Experiment 2 and the results reported in Ostojić et al. (2017)'s study. To this end, we employed the same experimental set-up and procedures used in Experiment 2, except that here, in one condition, jays were presented with the transparent U-barrier (*Barrier* condition) and in another condition, with no barrier (*No-barrier* condition). All birds (n=8) passed the familiarisation. In the test, one bird consistently cached no items, such that data of seven birds were analysed (see Material and Methods for details).

The two analyses using the two different indices yielded consistent results (Figure 4). No statistically significant difference could be detected in difference of the proportion of P cached when the observer was sated on P minus the proportion of P cached when observer was sated on M -  $[P_{\text{cached}} / (P_{\text{cached}} + M_{\text{cached}})_{\text{pre-fed P}}] - [P_{\text{cached}} / (P_{\text{cached}} + M_{\text{cached}})_{\text{pre-fed M}}]$  - between the *Barrier* and *No barrier* conditions (Median<sub>Barrier</sub> = 0, Median<sub>No Barrier</sub> = -0.04; Wilcoxon signed rank test: n=7, W=11, p=0.18). In addition, in both conditions, no statistically significant

502 difference could be detected in the proportion of P cached between the two pre-feeding trials  
503 (*Barrier* condition: Median<sub>Pre-fed P</sub> = 0.17, Median<sub>Pre-fed M</sub> = 0.1, Wilcoxon signed rank test, n=7,  
504 W=-2, p<sub>one-tailed</sub>=0.43; *No-barrier* condition: Median<sub>Pre-fed P</sub> = 0.12, Median<sub>Pre-fed M</sub> = 0.25, n=7,  
505 W=-9, p<sub>one-tailed</sub>=0.91).

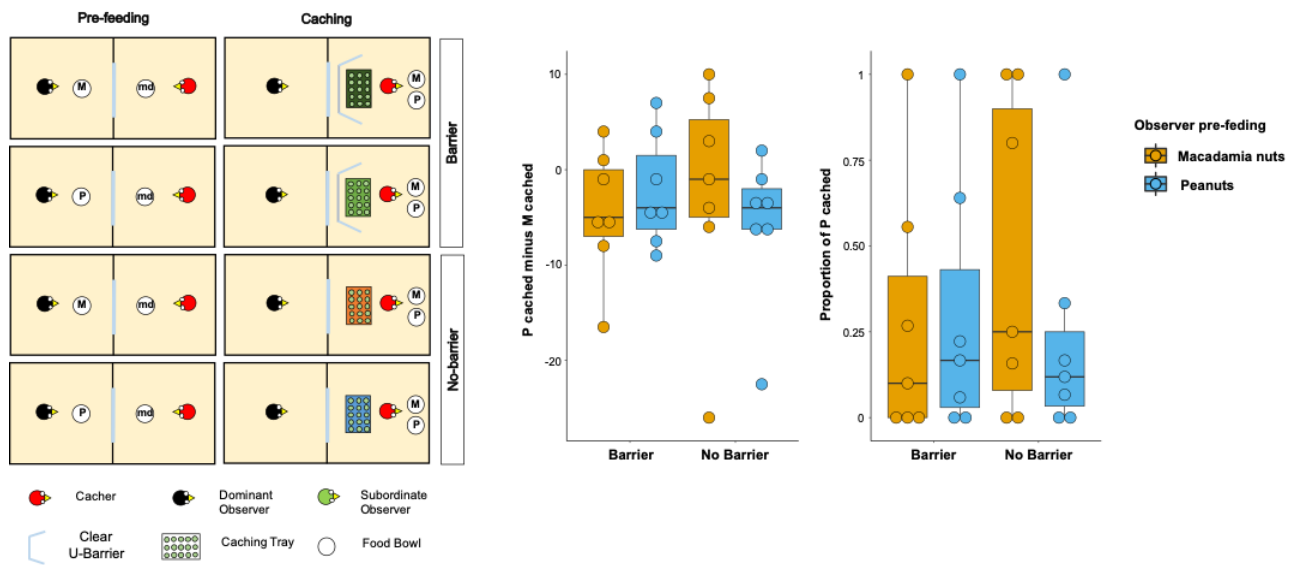
506

507 The same patterns of results were observed when the difference score of the number of P  
508 cached minus the number of M cached was analysed. No statistically significant difference  
509 could be detected in the differences of difference score - [P<sub>cached</sub> - M<sub>cached</sub>]<sub>pre-fed P</sub> - [P<sub>cached</sub> -  
510 M<sub>cached</sub>]<sub>pre-fed M</sub> - between the *Barrier* and *No barrier* conditions (Median<sub>Barrier</sub> = 0, Median<sub>No</sub>  
511 *Barrier* = 0; Wilcoxon signed rank test: n=7, W=15, p=0.14). In addition, in both conditions, no  
512 statistically significant difference could be detected in the difference score between the two  
513 pre-feeding trials (*Barrier* condition: Median<sub>Pre-fed P</sub> = -4, Median<sub>Pre-fed M</sub> = -5, Wilcoxon signed  
514 rank test, n=7, W=1, p<sub>one-tail</sub>=0.50; *No-barrier* condition: Median<sub>Pre-fed P</sub> = -4, Median<sub>Pre-fed M</sub> = -1,  
515 Wilcoxon signed rank test n=7, W=-9, p<sub>one-tail</sub>=0.91). Thus, the results from Experiment 5  
516 cannot be interpreted as providing support for the idea that the presence of the barrier may  
517 be the reason why the results in Experiment 2 did not detect the same pattern as the one  
518 reported in Ostojić et al. (2017). Crucially, like the *Transparent Barrier* condition in  
519 Experiment 1, both conditions in Experiment 5 also consistently could not detect the effect  
520 reported in Ostojić et al. (2017).

521

522

523



524

525

526 **Figure 4:** Left: Top-view schematic representation of the set-up and procedure used in the test of Experiment 5.  
 527 Trials were composed by a pre-feeding phase (left panels) and a caching phase (right panels). The cacher bird  
 528 received two trials with the transparent U-barrier (*Barrier* condition, top panels), and two trials with no barrier  
 529 (*No-barrier* condition, bottom panels). Within each condition, trials differed in the type of food (either peanuts, P,  
 530 or Macadamia nuts, M) that was provided to the observer in the pre-feeding phase. The cacher bird was always  
 531 presented with a handful of maintenance diet (md) in the pre-feeding phase of all trials. **Right:** Box and whisker  
 532 plots of data in Experiment 5. The central plot shows the difference in the number of peanuts cached minus the  
 533 number of Macadamia nuts cached in the *Barrier* condition (left) and *No-barrier* condition (right). The plot on  
 534 the right shows the proportion of P cached (over the total number of items cached) in the two conditions. The  
 535 colour of the boxes in the plot differs on the basis of the type of food that was provided to the observer in the  
 536 pre-feeding phase: blue denotes that the observer had been pre-fed P and orange denotes that the observer had  
 537 been pre-fed M.

538

539

## 540 DISCUSSION

541

542 In Experiments 1 and 2, we investigated whether Eurasian jays can take into account two  
 543 types of social cues simultaneously and perform the most advantageous behavioural output  
 544 accordingly. Specifically, we tested whether caching birds can integrate information from cues  
 545 correlating with a conspecific observer's desire and perspective to most effectively protect  
 546 their caches. Consistently across these two experiments, we did not detect effects that would  
 547 support such integration of information from different cues. In Experiment 1, jays did not  
 548 preferentially cache in the *out-of-view* tray when they were provided with a food that was

549 highly desired by the observer, and not more than when the food was not desired by the  
550 observer. Furthermore, in Experiment 2, jays did not preferentially cache the food for which  
551 observers had a decreased desire, and not more when the observer could see them than when  
552 they could not see them.

553

554 The negative results we obtained in both experiments appear inconsistent with previous  
555 effects in the literature, despite the use of set-ups that were very similar to those used in the  
556 original studies. Specifically, the negative results in the *Different Food* condition in Experiment  
557 1 appear incompatible with the effect reported in Legg and Clayton (2014), where jays were  
558 found to preferentially cache in an *out-of-view* tray specifically when they were observed by a  
559 conspecific. Similarly, the negative results in the *Clear barrier* condition in Experiment 2  
560 appear incompatible with the effect reported in Ostojić et al. (2017), where jays were found to  
561 preferentially cache a specific food when the observer was pre-fed on that food relative to  
562 when the observer was pre-fed on a different food.

563

564 Thus, we conducted three follow-up experiments to explore the robustness and reliability of  
565 the two previous findings (Legg and Clayton, 2014; Ostojić et al., 2017) that our first two  
566 experiments were built on. In Experiments 3 and 4, we attempted to replicate the effect  
567 reported by Legg and Clayton (2014), but – in contrast to the original study – no statistically  
568 significant difference between the experimental conditions was detected. Similarly, in  
569 Experiment 5 no statistically significant difference was detected between the experimental  
570 conditions, a result that contrasts with the effect reported by Ostojić et al. (2017). Thus,  
571 Experiments 3 to 5 also yielded negative results. However, evaluating the ‘success’ of a  
572 replication study from the statistical significance of a finding alone is overly simplistic,  
573 particularly for comparative cognition research, where – like in our experiment – the sample  
574 size of replication studies are often as equally small as the original studies (Farrar et al.,  
575 2020). However, the finding that we could not detect any significant effect in line with the  
576 original experiments of Legg and Clayton (2014) and Ostojić et al. (2017) across all five of our  
577 experiments was surprising, especially given that they were conducted in the same lab, with  
578 many if the same birds and experimenters. Specifically, in 4 out of the 7 tests of the hypothesis  
579 that the jays could use social cues to protect their caches, the results were not in the direction  
580 of the prediction: Experiment 1, prediction 1; Experiment 2, predictions 1 and 2; Experiment  
581 5, prediction 2, *No Barrier* condition. In the remaining 3 tests in which we had a directional  
582 prediction – Experiments 3 and 4, and Experiment 5 prediction 2, *Barrier* condition – the

583 effects were in the correct direction but were non-significant and much smaller than similar  
584 effects previously reported.

585

586 We propose two explanations for why our studies were unable to detect effects consistent  
587 with the previous literature, namely low power and the re-use of a unique bird sample. First,  
588 the sample sizes (often ~10) and trial numbers (often 1 to 3) used in cache protection studies  
589 are so low that the designs are only powered to consistently detect very large effects. Coupled  
590 with a likely publication bias, this may be taken to suggest that published effects are  
591 overestimated, and that the probability of a single replication study finding a similarly sized  
592 effect is low (Farrar et al., 2020; Fiedler and Prager, 2018; Hedges, 1984). Viewed in this light,  
593 it is perhaps not surprising that any one of our studies returned non-significant results.  
594 However, that we found no significant results across all five of the studies was surprising.  
595 Second, our five studies used the same populations of birds as tested in the previous studies  
596 (Legg and Clayton, 2014; Ostojić et al., 2017; Table 1), but they were around five years older.  
597 It is possible that the behaviour of these birds has changed over time, either due to learning  
598 effects, ageing, developing abnormal or stereotyped behaviour or change in motivation as a  
599 result of being kept in captivity for a long duration (Garner, 2005).

600

601 Our results contradict a relatively large body of literature on cache-protection strategies in  
602 corvids in general (Bugnyar et al., 2016; Dally et al., 2004, 2005; Emery and Clayton, 2001;  
603 Heinrich and Pepper, 1998) and Eurasian jays in particular (Legg et al., 2016; Shaw and  
604 Clayton, 2013). We were unable to elicit the cache protection strategies that this literature  
605 implies are consistently observable across corvid species, including in our Eurasian jays.  
606 While it is possible our findings were local failures to find these effects, it is also possible that  
607 the general research practices and methods that have produced the corvid social cognition  
608 literature are liable to producing unreliable findings or overestimated effect sizes. We  
609 currently do not know how many other studies have produced negative cache-protection  
610 results but have not been published, and understanding the magnitude of the publication bias  
611 (Fanelli, 2012; Scheel et al., 2020) in this literature is therefore a necessary step to evaluating  
612 the evidential strength within the field. Concerning our failed replication of Ostojić et al.  
613 (2017), a slightly different reasoning applies as, to our knowledge, no similar study has been  
614 conducted in another laboratory. As such, we believe we have access to all the data on this  
615 topic. These are the current study, the Ostojić et al. (2017) study, and a further, unpublished,  
616 replication attempt that also did not detect the originally reported effects (Crosby, 2019).

617 Overall, the data on these effects seem too uncertain to draw any firm conclusions about  
618 Eurasian jay cognition.

619

620 Our difficulty with replicating previous research, even in the same laboratory as the original  
621 findings and with many of the same birds and experimenters, highlights two ways in which  
622 research on corvid social cognition could make progress. First, understanding the extent of  
623 publication bias in our literatures is key to understanding their evidential value.  
624 Retrospectively, this may be achieved through meta-analysis techniques, and prospectively  
625 through effective pre-registering of hypothesis-testing research. Second, before building on  
626 findings, researchers can, where appropriate, build in reliability tests into their research  
627 programmes, which may be especially important for previous findings where the effects of  
628 publication bias are unknown. Such reliability tests may be useful when the same animals are  
629 tested in follow-up tests (as is the case in this study) to probe the reliability of the behavioural  
630 patterns over time.

631

632 In conclusion, the current study presents five experiments that are inconsistent with the  
633 previous literature on caching in Eurasian jays. Across all experiments, the effects were non-  
634 significant, and often in the opposite direction to predictions derived from the published  
635 literature. This suggests that previous effect sizes are likely overestimated, or at the very  
636 least, that the effects cannot be consistently elicited in the same or similar samples of birds. In  
637 Experiments 1 and 2, we investigated a follow-up question that assumed the reliability of  
638 previously reported statistical effects, which we later could not replicate in Experiments 3, 4  
639 or 5. The current series of experiments demonstrate the necessity to investigate the  
640 uncertainty of such effects and to adjust the claims – including those in previously published  
641 literature – accordingly. In regard to the behavioural effects investigated in this study, the  
642 caching patterns interpreted as cache-protection strategies in Legg and Clayton (2014) and  
643 Ostojić et al. (2017) do not seem to be reliable enough to form the basis for follow-up studies  
644 such as the ones reported in Experiments 1 and 2, at least in our sample of jays. It would be  
645 informative, but unfortunately not currently possible, to replicate these studies at other  
646 laboratories across the world.

647

648 **MATERIALS AND METHODS**

649

650 **Subjects**

651 Fourteen adult Eurasian jays from two separate colonies were tested in this study (Table. 1).  
 652 Most of the jays took part in multiple experiments and had previously been tested in  
 653 experiments that involved caching in a similar set-up as that used in the current study (details  
 654 about which jay participated in which experiment(s) are given in Table 1).

655

656

657 **Table 1:** Individual data of the birds that participated in this study (Experiments 1-5). The table reports also the  
 658 individual data of the birds that participated in Legg and Clayton (2014)'s study and to the caching experiment  
 659 by Ostojić et al. (2017).

660

Colony	Bird	Sex	Born	Experiment
1	Caracas	M	May 2006	1 <sup>S, Ob</sup> , 2 <sup>S, Ob</sup> , 3 <sup>S, Ob</sup> , 4 <sup>S, Ob</sup> , 5 <sup>S</sup> , 0 <sup>S, Ob</sup>
1	Dublin	M	May 2006	(1), 3 <sup>Ob</sup> , 4 <sup>S</sup> , 5 <sup>S, Ob</sup> , 0 <sup>S</sup>
1	Jerusalem	F	May 2006	1 <sup>S, Ob</sup> , 2 <sup>S, Ob</sup> , 3 <sup>S</sup>
1	Lima	M	May 2006	1 <sup>S</sup> , 2 <sup>S</sup> , 3 <sup>S, Ob</sup> , 4 <sup>S</sup> , 5 <sup>Ob, S</sup>
1	Lisbon	M	May 2006	(1 <sup>Ob</sup> ), (2), 3 <sup>Ob, S!</sup> , 4 <sup>S, Ob</sup> , 5 <sup>S!, Ob</sup>
1	Quito	F	May 2006	1 <sup>S</sup> , 2 <sup>S</sup> , 3 <sup>S</sup> , 4 <sup>S, Ob</sup> , 5 <sup>S</sup>
1	Rome	F	May 2006	1 <sup>S, Ob</sup> , 2 <sup>Ob, S</sup> , 3 <sup>S</sup> , 4 <sup>S, Ob</sup> , 5 <sup>Ob, S</sup> , 0 <sup>S</sup>
1	Washington	F	May 2006	1 <sup>S</sup> , 2 <sup>S, Ob</sup> , 3 <sup>S, Ob</sup> , 4 <sup>Ob, S</sup> , 5 <sup>S</sup> , 0 <sup>S</sup>
1	Wellington	F	May 2006	1 <sup>S, Ob</sup> , 2 <sup>S, Ob</sup> , 3 <sup>S</sup> , 4 <sup>S, Ob</sup> , 5 <sup>S, Ob</sup> ,
2	Hunter	F	May 2008	4 <sup>S</sup> , L <sup>S</sup>
2	Adlington	F	May 2008	(4 <sup>Ob</sup> ), L <sup>Ob, S</sup>
2	Webb	F	May 2008	(4)
2	Hoy	M	May 2008	(4), L <sup>S, Ob</sup> , 0 <sup>S, Ob</sup>

2	Romero	M	May 2008	(4), O <sup>S</sup>
2	Wilson	M	May 2008	L <sup>S</sup> , Ob
2	Ohurougu	F	May 2008	L <sup>S</sup> , Ob
2	Pendleton	M	May 2008	L <sup>S</sup> , O <sup>S</sup> , Ob
2	Ainslie	M	May 2008	L <sup>Ob</sup>
2	Purchase	F	May 2008	L <sup>Ob</sup>

661

662 'L' refers to Legg and Clayton (2014); 'O' refers to Ostojić et al. (2017). '( )' means that the bird participated only  
663 in a preliminary phase of the experiment (i.e., familiarisation). This is relevant only to the experiments reported  
664 in this study but not to Legg and Clayton (2014), Ostojić et al. (2017). 'Ob' means that the bird participated in the  
665 experiment as an observer. 'S' means that the bird participated in the experiment as a subject. '!' means that the  
666 bird did not complete the testing, such that its data was not included in the analyses. Note that the order in which  
667 'Ob' and 'S' are reported describes the order in which the bird was used as observer and as subject (e.g., '1<sup>Ob, S</sup>'  
668 means that the bird was used as observer *before* being used as subject in Experiment 1, whereas '1<sup>S, Ob</sup>' means  
669 that the bird was used as observer *after* being used as subject in Experiment 1).

670

671

672 All of the jays were hand-raised, having been taken as chicks from wild nests or from the  
673 natural nests of birds in a breeding programme. The birds from each colony were housed as a  
674 group in large outdoor aviaries each measuring 20m long x 10m wide x 3m high in Clayton's  
675 Comparative Cognition Lab at the Sub-Department of Animal Behaviour, University of  
676 Cambridge, Madingley, UK. At one end, the aviaries were divided such that birds had access to  
677 multiple smaller aviaries (approximately 6x2x3 m) and from these smaller aviaries birds  
678 could access indoor (colony 1) or fully sheltered (colony 2) testing compartments (2x1x3 m).  
679 Birds of colony 2 were housed in pairs in indoor cages until 2009 or 2010. Outside of testing  
680 the birds had *ad libitum* access to their maintenance diet of vegetables, eggs, seed and fruits.  
681 Water was available at all times. All procedures were approved by the University of  
682 Cambridge Animal Ethics Committee.

683

#### 684 **Experimental set-up**

685 Birds were tested in the testing compartments measuring 2m long x 1m wide x 3 m high,  
686 which were accessible from the smaller aviaries through flap windows. In trials requiring the



687 presence of an observer, two birds – a cacher bird and an observer bird – were located in  
688 adjacent compartments. These compartments were separated by wire mesh and additional  
689 opaque sheeting. A little mesh window (30x55 cm) was not covered by the opaque sheeting  
690 and through it the birds had visual access to the adjacent compartment. Testing  
691 compartments contained a suspended platform (1x1 m) approximately 1 m from the ground,  
692 onto which food bowls, caching trays and Perspex barriers could be placed. Each type of food  
693 used in the experiments was presented in a bowl of a specific colour, and these colours were  
694 kept consistent for all birds to minimise the likelihood of experimenter errors. Rectangular  
695 seedling trays (5 x 3 pots filled with sand) were used as caching trays. Trays were painted  
696 different colours and were trial-specific to minimise the probability that birds' caching  
697 behaviour in one trial would be influenced by its memory from previous trials.

698

699 In Experiments 1, 3, and 4, a 'T-barrier' was used to manipulate the observer's visual access to  
700 the caching trays. It was the same T-barrier that Legg and Clayton (2014) used. This barrier  
701 consisted of three plastic panels (25x40 cm) forming two arms and one stem. One arm of the  
702 'T' was constructed out of transparent Perspex, while the other arm and the stem were  
703 constructed out of white opaque Perspex. The T-barrier could be placed around two caching  
704 trays in the cacher's compartment, such that the observer could see the tray behind the  
705 transparent arm (*in-view* tray) but could not see the tray behind the opaque arm (*out-of-view*  
706 tray). Due to the height of the barrier, the observer could always see the cacher when the  
707 latter was standing upright in proximity of the trays. However, the observer could not see the  
708 exact location where the cacher hid the food when it was caching in the *out-of-view* tray.

709

710 In Experiments 2 and 5, a U-barrier was used to manipulate the observer's visual access to the  
711 caching trays. The barrier consisted of two lateral Perspex panels (26x25 cm) and one central  
712 Perspex panel (53x25 cm) forming two angles of approximately 45°. In Experiment 2, we used  
713 two U-barriers, one made of transparent Perspex and another made of white opaque Perspex.  
714 In Experiment 5, only the transparent barrier was used. The U-barrier was placed around a  
715 single tray in the cacher's compartment, and if opaque, it impaired the observer's visual  
716 access to the caching tray.

717

## 718 **General procedures**

719 In all experiments, the birds' maintenance diet was removed from the aviary approximately  
720 1.5h prior to the start of each trial to ensure that the birds were mildly hungry and thus likely  
721 to interact with food provided during testing.

722

723 **Familiarisation.** In all experiments in which the birds had not experienced the set-up and  
724 apparatuses just prior to testing (i.e., in Experiments 1, 2, 4, and 5), a familiarisation  
725 procedure was conducted to ascertain that birds were comfortable caching in trays placed in  
726 proximity of the respective barriers (see Specific Procedures for further details). During the  
727 familiarisation, each bird was tested in isolation, i.e., with no other birds present in the test  
728 area. Compartments used during the familiarisation were not used in the test phase to  
729 minimise the probability of carry-over effects.

730

731 **Test.** In Experiments 1, 2, and 5, test trials involved a pre-feeding phase followed by a caching  
732 phase. Before the start of a test trial, two birds (a cacher and an observer) were given access  
733 to two adjacent compartments. Subsequently, the experimenter placed a bowl containing the  
734 pre-feeding food (macadamia nuts or peanuts) on the suspended platform in the observer's  
735 compartment and a bowl containing a handful of maintenance diet on the platform in the  
736 cacher's compartment (Figures 1, 2, and 4). Both bowls were placed in front of the mesh  
737 window to ensure that the birds could see each other whilst eating and to maximise the  
738 likelihood that the cacher could see on which food the observer was pre-fed. The  
739 experimenter then left the test room and the birds could eat the pre-feeding food for 15  
740 minutes. Next, the experimenter entered the test room again and removed the bowls as well  
741 as any food remains on the platforms. In the subsequent caching phase, the caching trays, as  
742 well as the barrier and the food bowl, were positioned in front of the mesh window in the  
743 cacher's compartment (Figures 1, 2, and 4). The experimenter then left the test room and the  
744 birds were given 15 minutes during which the cacher could eat and cache the food in the  
745 trays. In Experiments 3 and 4, the test trials involved only a caching phase (Figure 3). Before  
746 the start of a trial, the cacher bird was given access to the testing compartment where the  
747 caching trays and the T-barrier had already been positioned. In the test trials of the *Observed*  
748 condition, a second bird (i.e., the observer) was also induced to enter the adjacent  
749 compartment. Subsequently, a food bowl was placed on the suspended platform in the  
750 cacher's compartment. The experimenter then left the test room and the cacher was given the  
751 opportunity to eat and cache food for 15 minutes.

752

753 At the end of each familiarisation and test trial, the experimenter opened the flap windows to  
754 allow the bird(s) to re-join the rest of the group in the aviary and recorded the amount of food  
755 eaten and the number and location of caches by manually checking the food bowls and trays.  
756 Approximately three hours after each trial, the cacher was allowed to re-enter the caching  
757 compartment. No other birds were present in the test area and the flap window was kept  
758 open so that the bird had access not only to the test compartment but also to the adjacent  
759 smaller aviary. Note that the door connecting the small aviary to the main aviary was kept  
760 closed such that no other bird could enter the cacher's aviary or compartment. The cacher  
761 could retrieve the hidden items and re-cache them in the compartment and in the adjacent  
762 small aviary. This retrieval phase was conducted only to reduce the probability that birds  
763 would stop caching in the trays, and thus these data were not analysed. Birds received a single  
764 test trial per day.

765

766

## 767 **Specific procedures**

### 768 **Experiment 1.**

769 **Familiarisation.** Birds ( $n = 9$ ; Table 1) received two familiarisation trials on two separate  
770 days to ensure that they were comfortable caching in trays when these were placed in  
771 proximity of each of the two arms of the T-barrier. On each trial, the bird was presented with  
772 the T-barrier, a single caching tray and a food bowl containing either 50 macadamia nut  
773 halves (M) or 50 whole peanuts with skin (P). The type of food (macadamia nuts or peanuts)  
774 was randomly assigned to birds but each bird was provided with the same type of food in  
775 both trials. The bird was given the opportunity to eat and cache for 15 minutes. On one trial,  
776 the tray was placed behind the opaque arm of the T-barrier and on the other trial it was  
777 placed behind the transparent arm. The order in which birds experienced the tray in the two  
778 locations was counterbalanced across birds. The orientation of the barrier within the  
779 compartment was different from that later used during testing and was kept consistent for  
780 each bird across the two familiarisation trials (Figure 1). This procedure was chosen to  
781 ensure that the birds were not more familiar with one of the two orientations of the barrier in  
782 a specific spatial set-up (e.g., opaque arm facing the outdoor aviary) in the subsequent test. To  
783 proceed to the test, birds had to cache at least one food item in the tray on each trial. If a bird  
784 did not meet this criterion, it was excluded from further testing. All birds except two (i.e.,  
785 Dublin and Lisbon; Table 1) passed the familiarisation and proceeded to the test.

786

787 **Test.** During the pre-feeding phase, cachers ( $n = 7$ ; Table 1) could see a conspecific eat a  
788 specific type of food: either the same type of food they were going to receive in the  
789 subsequent caching phase (*Same Food* condition) or a different one (*Different Food* condition;  
790 Figure 1). The order in which the birds experienced the *Different Food* and *Same Food*  
791 conditions was counterbalanced across birds. In the subsequent caching phase, cachers were  
792 provided with the same food used in the familiarisation and with two caching trays, each one  
793 placed behind one of the two arms of the T-barrier (Figure 1). The food given to the observer  
794 during the pre-feeding phase and to the cacher during the caching phase was either 50  
795 macadamia nut halves or 50 whole peanuts with skin. All birds received one trial per  
796 condition, i.e. two test trials in total. If a bird cached no items on both trials, it was paired with  
797 a different observer and the two trials were repeated. If it again did not cache on both trials,  
798 these data were not included in the analysis. In contrast, if the bird cached with the second  
799 observer, then these data were included in the analysis. This procedure was decided during  
800 data collection, after one bird (Lima) did not cache any food across both trials, but before the  
801 analysis was conducted. For all other birds, test trials were not repeated. The analysis  
802 included the data of all seven birds. Experiment 1 was conducted from October to November  
803 2017 by LO, BF and PA.

804

## 805 **Experiment 2.**

806 **Familiarisation.** Birds ( $n = 8$ ; Table 1) received two familiarisation trials on two separate  
807 days to ascertain that they were comfortable caching both types of food (macadamia nuts and  
808 peanuts) in a tray placed in proximity of each of the U-barriers (transparent and opaque). On  
809 each trial, the bird was presented with a U-barrier, a single caching tray and two food bowls,  
810 which were presented sequentially. The food bowls contained either 50 macadamia nut  
811 halves or 50 whole peanuts with skin. The bird was given the opportunity to eat and cache for  
812 20 minutes: during the first 10 minutes it was provided with one type of food and during the  
813 next 10 minutes with the other. The order in which the birds experienced the two types of  
814 foods was counterbalanced across birds and across trials, such that each bird experienced one  
815 order on their first trial and the opposite order on their second trial. On each trial, the barrier  
816 was either transparent or opaque. The order in which the birds experienced the two types of  
817 the U-barrier was counterbalanced across birds. To proceed to testing, birds had to i) cache at  
818 least one item on each trial, and ii) cache at least one item of each kind of food across the two

819 trials. If a bird did not meet these criteria, it was excluded from further testing. All birds  
820 except one (Lisbon; Table 1) passed the familiarisation and proceeded to the test.

821

822 **Test.** The pre-feeding phase was conducted in the same way as in Experiment 1. In the  
823 subsequent caching phase, the cacher was provided with a single tray placed within the U-  
824 barrier and two food bowls (one containing 50 macadamia nut halves and the other  
825 containing 50 whole peanuts with skin). Birds ( $n = 7$ ; Table 1) received four trials in total:  
826 there were two conditions, namely the *In-view* condition (clear U-barrier) and the *Out-of-view*  
827 condition (opaque U-barrier), and within each condition, there were two trials, one in which  
828 the observer was pre-fed on one type of food (e.g., macadamia nuts), and one in which the  
829 observer was pre-fed on the other type food (e.g., peanuts). Birds first received both trials of  
830 one condition, and then the two trials of the other condition. The order of conditions was  
831 counterbalanced across birds. The order in which the observer was pre-fed on the two types  
832 of food was counterbalanced across birds within condition, but kept consistent across  
833 conditions (i.e., the order of the two trials was the same in both conditions). If a bird did not  
834 cache any items in one or two trials, those trials were repeated at the end of the experiment.  
835 However, if a bird did not cache any items in more than two trials, that bird was not tested  
836 any further and was excluded from the analysis. A specific trial was repeated for a maximum  
837 of two times, such that a bird was excluded from the analysis if it cached no item in both  
838 repeated trials. It was necessary to repeat trials for three birds (i.e., Jerusalem: two repeated  
839 trials; Lima: two repeated trials; Rome: one repeated trial). The analysis included the data of  
840 all seven birds. Experiment 2 was conducted from January to February 2018 by PA.

841

### 842 **Experiment 3.**

843 **Familiarisation.** Here, we conducted no familiarisation because this experiment took place  
844 shortly after Experiment 1 (which also involved the T-barrier) and involved the same birds.  
845 Thus, participation in Experiment 1 already insured that birds were comfortable caching in  
846 trays next to the T-barrier.

847

848 **Test.** The procedure of the test was simplified from the procedure in Legg and Clayton  
849 (2014). The main difference was that there was only one trial per condition. In addition, we  
850 used either 50 whole peanuts with skin or macadamia nut halves (counterbalanced across  
851 birds) as food for the cachers, whereas the original study used 30 peanut halves. Birds ( $n = 8$ ;  
852 Table 1) were given two trials in total: one with an observer present in the adjacent

853 compartment (*Observed* condition) and one in which no observer was present (*Private*  
854 condition). The order of the conditions was counterbalanced across birds. On each trial, two  
855 trays were positioned behind the T-barrier, one behind the opaque and one behind the  
856 transparent arm. The orientation of the T-barrier was counterbalanced across birds but kept  
857 constant across trials for each bird. The cacher was given 15 minutes during which it could  
858 each or cache the food. If a bird did not cache on a trial, the trial was repeated. If the bird did  
859 not cache again, the data were not included in the analysis. A trial was repeated for one bird  
860 (Caracas). Due to timing constraints, another bird (Lisbon) was not given the possibility to  
861 repeat the trials in which no item was cached. Thus, although this was not pre-specified as an  
862 exclusion criterion, Lisbon's data were excluded from the analysis before it was conducted.  
863 The analysis included the data of seven birds, i.e., all birds except one (Lisbon). Experiment 3  
864 was conducted in December 2017 by LO.

865

#### 866 **Experiment 4.**

867 **Dominance Hierarchy.** In Legg and Clayton (2014)'s experiment, cacher birds received four  
868 trials in the *Observed* condition: two trials in which they were observed by a higher ranked  
869 individual (*Observed by dominant* condition), and two trials in which they were observed by a  
870 lower ranked individual (*Observed by subordinate* condition). Thus, to replicate the original  
871 design it was necessary to determine the dominance hierarchy within each colony. To this  
872 end, *ad libitum* observations were conducted for each colony. Birds were observed as a group  
873 in the main outdoor aviaries after their maintenance diet had been removed for  
874 approximately 2h. For each observation session, maintenance diet was presented on a single  
875 food platform in the aviary. This procedure was chosen to increase activity within the colony.  
876 To solicit competitive interactions among birds, higher value food items (e.g., wax worm  
877 larvae, *Galleria mellonella*) were also presented in a bowl or scattered around on the floor of  
878 the aviary. The identity of both actor and recipient involved in any displacement (i.e., Bird X  
879 approaches Bird Y causing Bird Y to leave) was recorded. If necessary, to obtain data for all  
880 birds, higher ranked birds were locked into separate compartments to favour interactions  
881 among lower ranked birds. Observation sessions were conducted on multiple days, until data  
882 were collected to establish a clear social hierarchy within each colony. Sessions lasted  
883 approximately 40 min each. Dominance hierarchy data were not collected for colony 2  
884 because only one bird of this group (Hunter) passed the familiarisation.

885

886 **Familiarisation.** Unlike birds from colony 1, birds of colony 2 had not recently had any  
887 experience with the T-barrier. Thus, all birds (n=13; Table 1) received two familiarisation  
888 trials on separate days to ascertain that the birds were comfortable caching in proximity of  
889 both the transparent and the opaque arm of the T-barrier (see also pre-registration addition  
890 from 8 Dec 2018). This familiarisation followed the same procedure as the familiarisation in  
891 Experiment 1, except that here, each bird was provided with 30 peanut halves as in the  
892 original study (Legg and Clayton, 2014). To proceed to the test, birds were required to cache  
893 at least one food item in each of the two familiarisation trials. If no item was cached in a trial,  
894 then that trial was repeated for a maximum of two times. Thus, a bird could receive a  
895 maximum of 6 trials in total. The repeated trials were conducted at the end (for example, if a  
896 bird's first trial had to be repeated, then the bird received the second, pre-planned trial on day  
897 2, and subsequently it received the first trial again on day 3). It was necessary to repeat trials  
898 for two birds (i.e., Lisbon: one repeated trial; Lima: one repeated trial). Nine birds (Table 1)  
899 passed the familiarisation and proceeded to the test. Due to an experimenter's error, the raw  
900 data for one bird (Hunter) in the familiarisation was not archived.

901  
902 **Test.** Following the procedure in Legg and Clayton (2014), birds (n=9) received 6 trials in  
903 total. There were three conditions (*Private* condition, *Observed by dominant* condition, and  
904 *Observed by subordinate* condition) and in each condition, the cacher experienced the T-  
905 barrier in two different orientations (opaque arm of the barrier was facing the outdoor aviary,  
906 transparent arm facing outdoor aviary) on two separate trials. Each bird was first tested in all  
907 conditions with the barrier being kept consistent in one specific orientation, then  
908 subsequently received the remaining trials with the barrier being kept consistent in the  
909 alternative orientation. The order in which the two orientations of the barrier were  
910 experienced, was counterbalanced among birds.

911  
912 On each trial, the cacher was given access to the testing compartment and presented with the  
913 T-barrier, the *in-view* tray (i.e. the tray placed behind the transparent arm of the barrier), the  
914 *out-of-view* tray (i.e. the tray placed behind the opaque arm of the barrier) and a bowl  
915 containing 30 peanuts halves. The bowl was placed close to the stem of the "T" such that it was  
916 equidistant from the two caching trays (Figure 4). The cacher could eat and cache for 15  
917 minutes and was subsequently released back into the aviary.

918

919 All birds were tested in all three conditions, except the highest and lowest ranked bird in each  
920 colony. The former could only be tested in the *Private* and *Observed by subordinate* conditions  
921 and the latter could only be tested in the *Private* and *Observed by dominant* conditions. Thus,  
922 most birds received six trials in total (two trials per condition), whereas the highest and  
923 lowest ranked individuals in each colony received four trials in total because they could only  
924 be tested in two of the three conditions. In line with the procedure of the original study, test  
925 trials were not repeated if the bird cached no item. However, we decided to repeat a specific  
926 trial (Dublin's first trial in the *Observed by dominant* condition) because the bird that served  
927 as observed (Rome) appeared to experience issues with flying. This trial was repeated with a  
928 different observer after the remaining pre-planned trials were completed. The analysis  
929 included the data of all nine birds. This experiment was pre-registered on the Open Science  
930 Framework (<https://osf.io/8p4tx/>). The pre-registration was conducted after the  
931 familiarisation was completed but before the start of the test. Experiment 4 was conducted  
932 from October to December 2018 by PA (colony 1) and Rachel Crosby (colony 2).

933

#### 934 **Experiment 5.**

935 **Familiarisation.** Birds (n=8; Table 1) received two familiarisation trials on separate days to  
936 ensure that they were motivated to cache both types of food and were comfortable caching in  
937 a tray both when it was positioned close to the U-barrier and when no barrier was present.  
938 Thus, the familiarisation followed the procedure of the familiarisation in Experiment 2, except  
939 that here, one trial involved the transparent U-barrier and the other one no barrier. To  
940 proceed to testing, birds had to i) cache at least one item on each trial (i.e. both with barrier  
941 present and with no barrier present), and ii) cache at least one item of each type of food  
942 across the two trials. If no item was cached in a trial, then that trial was repeated for a  
943 maximum of two times. Thus, a bird could receive a maximum of six trials in total. The  
944 repeated trials were conducted at the end: for example, if a bird's first trial had to be repeated,  
945 then this bird received the second, pre-planned trial on day 2, and subsequently it received  
946 the first trial again on day 3. It was necessary to repeat trials for two birds (i.e., Lisbon: one  
947 repeated trial; Wellington: one repeated trial). All eight birds (Table 1) passed the  
948 familiarisation and proceeded to the test.

949

950 **Test.** The procedure in the test phase was the same as in Experiment 2, except that instead of  
951 two different U-barriers being used (clear and opaque), here there was either a clear U-  
952 barrier (*Barrier* condition) or no barrier at all (*No Barrier* condition; Figure 4). Birds first



953 received both trials (observer pre-fed macadamia nuts and observer pre-fed peanuts) of one  
954 condition (e.g. *Barrier* condition), and then the two trials of the other condition. The order in  
955 which the two conditions were conducted was counterbalanced across birds. The order in  
956 which observers were pre-fed the two kinds of food within a condition was counterbalanced  
957 across birds, but kept consistent across conditions such that the order of the two trials for  
958 each bird was the same in both conditions. If a bird did not cache any items in one or two  
959 trials, those trials were repeated at the end of the experiment. Each trial could be repeated no  
960 more than two times (i.e., three attempts in total). If a bird cached no item in more than two  
961 trials, that bird was not tested any further and was excluded from the analysis. It was  
962 necessary to repeat trials for two birds (i.e., Quito: one trial; Wellington: two trials). The  
963 analysis included the data of seven birds (Table 1), i.e., all birds except one (Lisbon), that  
964 cached no items in three trials. This experiment was pre-registered on the Open Science  
965 Framework (<https://osf.io/8p4tx/>). Experiment 5 was conducted in November 2018 by PA.

966

#### 967 **Data Collection**

968 In all experiments, we recorded the number and type of food items cached on each trial by  
969 manually checking the trays. The experimenters were not blind to the conditions while  
970 counting the food items. These data were used to test whether the birds had a preference for  
971 caching a specific type of food or for caching in a specific tray. In all experiments, we also  
972 recorded i) the number of items taken from the bowl by observers (during pre-feeding) and  
973 by cachers, and ii) the number of items recovered by cachers during retrieval sessions. These  
974 data were collected such that all data available for each trial are archived and available, but  
975 these data were not relevant to the experimental question so that they were not analysed.

976

#### 977 **Statistical Analysis**

978 The birds' preference for a specific type of food or tray was analysed according to two indices:  
979 proportion scores (e.g., the proportion of items cached in one location out of total number of  
980 items cached in both locations) and difference scores (e.g., number of items cached in one  
981 location minus the number of items cached in the other location). As stated in the pre-  
982 registrations of Experiments 4 and 5 (<https://osf.io/8p4tx/>), we originally planned to analyse  
983 the data of all five experiments only through proportion scores. However, when a bird caches  
984 no item in a trial, then the individual performance in that specific trial cannot be analysed  
985 through the proportion scores, yet it can still be analysed through the difference scores. This  
986 issue is relevant only to Experiment 4, where – in line with the procedure of the original study

987 by Legg and Clayton (2014), and in contrast with the procedure of Experiments 1, 2, 3 and 5 –  
988 the trials in which no item was cached were not repeated. Nevertheless, after the study was  
989 conducted, we decided to analyse the data of all experiments – not only the data of  
990 Experiment 4 – also by using the difference scores. We reasoned that, if there are large  
991 discrepancies between the results obtained with both types of indices, then this may be  
992 important information regarding the robustness of any effects because such discrepancies  
993 would show that results from small sample sizes are easily susceptible to change based on the  
994 type of analysis used.

995

996 **Experiment 1.** For each trial we calculated the proportion of items cached in the *out-of-view*  
997 tray out of the total number of items cached in the *out-of-view* and *in-view* trays [ $Caches_{out-of-view} / (Caches_{out-of-view} + Caches_{in-view})$ ]. In parallel, for each trial we calculated the difference  
998 score, i.e., the number of items cached in the *out-of-view* tray minus the number of items  
999 cached in the *in-view* tray [ $Caches_{out-of-view} - Caches_{in-view}$ ]. Both indices indicate a preference  
1000 for caching in the *out-view* tray over the *in-view* tray. Wilcoxon signed rank tests were used to  
1001 test whether the two indices of preference for caching in the *out-of-view* tray differed between  
1002 the *Same Food* condition and the *Different Food* condition. Further, in the *Different Food*  
1003 condition, one-sample Wilcoxon signed rank tests were used to test whether the preference  
1004 for caching in the *out-of-view* tray was different from that expected by chance, i.e., 0.5 for the  
1005 proportion score, and 0 for the difference score. As an additional, exploratory analysis, one-  
1006 sample Wilcoxon signed rank tests were used to investigate whether the preference for  
1007 caching in the *out-of-view* tray differed from chance (again, 0.5 for the proportion score, and 0  
1008 for the difference score) in *Same Food* condition.

1010

1011 **Experiment 2.** For each trial, we calculated i) the proportion of peanuts (P) cached out of the  
1012 total number of peanuts and macadamia nuts (M) cached [ $P_{cached} / (P_{cached} + M_{cached})$ ] and; ii) the  
1013 difference score, i.e., the number of P cached minus the number of M cached [ $P_{cached} - M_{cached}$ ].  
1014 These scores indicate a potential preference for caching P over M. For each condition (*In-view*  
1015 and *Out-of-view* condition), we further calculated i) the difference of proportions score,  
1016 namely the proportion of P cached when the observer was pre-fed on P minus the proportion  
1017 of P cached when the observer was pre-fed on M: [ $P_{cached} / (P_{cached} + M_{cached})$ ]<sub>pre-fed P</sub> -  
1018 [ $P_{cached} / (P_{cached} + M_{cached})$ ]<sub>pre-fed M</sub> and; ii) the difference of difference score, namely the difference  
1019 score when the observer was pre-fed on P minus the difference score when the observer was  
1020 pre-fed on M: [ $P_{cached} - M_{cached}$ ]<sub>pre-fed P</sub> - [ $P_{cached} - M_{cached}$ ]<sub>pre-fed M</sub>. These scores indicate a

1021 potential preference to cache P over M when the observer was sated on P relative to when the  
1022 observer was sated on M. Consequently, Wilcoxon signed rank tests were used to test whether  
1023 the difference of proportion scores and the difference of differences scores were statistically  
1024 different between the *In-view* and in the *Out-of-view* conditions. In addition, Wilcoxon signed  
1025 rank tests were also used to test whether – in the *In-view* condition – the proportion score and  
1026 the difference score differed when the observer was pre-fed on P relative to when the  
1027 observer was pre-fed on M. The same tests were also conducted in the *Out-of-view* condition  
1028 as an exploratory analysis.

1029

1030 **Experiment 3.** As in Experiment 1, for each trial we calculated the proportion of items cached  
1031 in the *out-of-view* tray –  $[\text{Caches}_{\text{out-of-view}} / (\text{Caches}_{\text{out-of-view}} + \text{Caches}_{\text{in-view}})]$  – and the difference  
1032 between the number of items cached in the *out-of-view* tray and the number of items cached  
1033 in the *in-view* tray –  $[\text{Caches}_{\text{out-of-view}} - \text{Caches}_{\text{in-view}}]$ . Subsequently, we used one tailed  
1034 Wilcoxon signed rank tests to investigate whether each of the two indices was significantly  
1035 higher in the *Observed* condition than in the *Private* condition.

1036

1037 **Experiment 4.** As in Experiment 1, for each trial we calculated the proportion of items cached  
1038 in the *out-of-view* tray –  $[\text{Caches}_{\text{out-of-view}} / (\text{Caches}_{\text{out-of-view}} + \text{Caches}_{\text{in-view}})]$  – and the difference  
1039 between the number of items cached in the *out-of-view* tray and the number of items cached  
1040 in the *in-view* tray –  $[\text{Caches}_{\text{out-of-view}} - \text{Caches}_{\text{in-view}}]$ . Following Legg and Clayton (2014), for  
1041 both indexes we calculated individual mean values in the *Private* condition and *Observed*  
1042 condition. In the latter case, all trials in which an observer was present (i.e., *Observed by*  
1043 *dominant* condition and *Observed by subordinate*) were taken into account to calculate  
1044 individual mean values. It should be noted that for the proportion score, the actual number of  
1045 trials that could be used to calculate the individual means in each condition was not  
1046 consistent for all birds. This is because the trials in which no item was cached across both  
1047 trays had to be excluded. We also calculated the mean number of caches made in both trays by  
1048 each individual, in each condition.

1049

1050 The data were analysed using the same statistical analysis used by Legg and Clayton (2014).  
1051 In particular, we used a one-tailed permutation test to investigate whether the individual  
1052 mean of the proportion of the caches in the *out-of-view* tray was higher in the *Observed*  
1053 condition than in the *Private* condition. Following Legg and Clayton (2014) we also tested

1054 whether the average number of items cached across both trays differed between the *Observed*  
1055 and *Private* conditions by using a two-tailed permutation test.

1056

1057 In parallel, we also analysed the data using Wilcoxon signed rank tests like in all other  
1058 experiments reported here. In particular, one-tailed tests were used to investigate whether  
1059 individual mean values (for both the proportion score and the difference score) were  
1060 significantly higher in the *Observed* condition than in the *Private* condition.

1061

1062 **Experiment 5.** As in Experiment 2, for each trial we calculated the proportion of P cached –  
1063  $[P_{\text{cached}} / (P_{\text{cached}} + M_{\text{cached}})]$  – and the difference of P cached –  $[P_{\text{cached}} - M_{\text{cached}}]$ . Further, as in  
1064 Experiment 2, we also calculated for each condition (*Barrier* and *No-barrier* conditions) the  
1065 difference of proportion scores –  $[P_{\text{cached}} / (P_{\text{cached}} + M_{\text{cached}})]_{\text{pre-fed P}} - [P_{\text{cached}} / (P_{\text{cached}} + M_{\text{cached}})]_{\text{pre-}}$   
1066  $\text{fed M}$  – and the difference of differences scores –  $[P_{\text{cached}} - M_{\text{cached}}]_{\text{pre-fed P}} - [P_{\text{cached}} - M_{\text{cached}}]_{\text{pre-fed}}$   
1067  $\text{M}$ . Consequently, we used Wilcoxon signed rank tests to investigate whether the difference of  
1068 proportions score and the difference of differences score were statistically different between  
1069 the *Barrier* and *No-barrier* conditions. In addition, one tailed Wilcoxon signed rank tests were  
1070 also used to test whether – in each condition – the proportion score and the difference score  
1071 were higher when the observer was pre-fed P relative to when the observer was pre-fed M.

1072

1073 All statistical analyses were performed in R (R.3.5) using the RStudio 1.1.447 wrapper  
1074 (RStudio Team, 2018). Permutation tests were conducted with the package *coin* (Hothorn et  
1075 al., 2006). All tests were two tailed, unless stated otherwise. Alpha was set to 0.05.

1076

1077 **DATA AVAILABILITY**

1078 Data and analyses of all experiments are available at <http://doi.org/10.5281/zenodo.4636561>

1079

1080

1081

1082 **ACKNOWLEDGEMENTS**

1083 We are grateful to Rachel Crosby for collecting data for Experiment 4 in colony 2, and for her  
1084 feedback on the manuscript. During the preparation of the manuscript, PA received support  
1085 from the Leverhulme Trust (Grant reference: SAS-2020-004\10). BGF was supported by the  
1086 University of Cambridge BBSRC Doctoral Training Programme (BB/M011194/1). CK was  
1087 supported by European Commission Marie Skłodowska-Curie Fellowship MENTALIZING  
1088 ORIGINS (Grant reference: 752373). NSC was funded by the European Research Council under  
1089 the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant  
1090 Agreement No. 3399933, awarded to NSC, and provided financial support conducting this  
1091 research.

1092

1093 **REFERENCES**

1094

- 1095 Apperly, I. A., and Butterfill, S. A. (2009). Do humans have two systems to track beliefs and  
1096 belief-like states? *Psychol. Rev.* 116, 953–970. doi:10.1037/a0016923.
- 1097 Baker, C. L., Jara-Ettinger, J., Saxe, R., and Tenenbaum, J. B. (2017). Rational quantitative  
1098 attribution of beliefs, desires and percepts in human mentalizing. *Nat. Hum. Behav.* 1,  
1099 0064. doi:10.1038/s41562-017-0064.
- 1100 Balleine, B. W., and Dickinson, A. (1998). The role of incentive learning in instrumental  
1101 outcome reevaluation by sensory-specific satiety. *Anim. Learn. Behav.* 26, 46–59.  
1102 doi:10.3758/BF03199161.
- 1103 Bartsch, K., and Wellman, H. (1989). Young Children’s Attribution of Action to Beliefs and  
1104 Desires. *Child Dev.* 60, 964. doi:10.2307/1131035.
- 1105 Bennett, M., and Galpert, L. (1992). Complex belief-desire reasoning in children. *Soc. Dev.* 1,  
1106 201–210. doi:10.1111/j.1467-9507.1992.tb00124.x.
- 1107 Bray, J., Krupenye, C., and Hare, B. (2014). Ring-tailed lemurs (*Lemur catta*) exploit  
1108 information about what others can see but not what they can hear. *Anim. Cogn.* 17, 735–  
1109 744. doi:10.1007/s10071-013-0705-0.
- 1110 Buckner, C. (2013). Morgan’s Canon, meet Hume’s Dictum: Avoiding anthropofabulation in  
1111 cross-species comparisons. *Biol. Philos.* 28, 853–871. doi:10.1007/s10539-013-9376-0.
- 1112 Bugnyar, T., and Heinrich, B. (2005). Ravens, *Corvus corax*, differentiate between  
1113 knowledgeable and ignorant competitors. *Proceedings. Biol. Sci.* 272, 1641–6.  
1114 doi:10.1098/rspb.2005.3144.
- 1115 Bugnyar, T., Reber, S. A., and Buckner, C. (2016). Ravens attribute visual access to unseen  
1116 competitors. *Nat. Commun.* 7, 10506. doi:10.1038/ncomms10506.
- 1117 Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., and Tomasello, M. (2017). Great apes  
1118 distinguish true from false beliefs in an interactive helping task. *PLoS One* 12, e0173793.  
1119 doi:10.1371/JOURNAL.PONE.0173793.
- 1120 Butterfill, S. A., and Apperly, I. A. (2013). How to Construct a Minimal Theory of Mind. *Mind*  
1121 *Lang.* 28, 606–637. doi:10.1111/mila.12036.
- 1122 Carruthers, P. (2020). Representing the Mind as Such in Infancy. *Rev. Philos. Psychol.* 11, 765–  
1123 781. doi:10.1007/s13164-020-00491-9.
- 1124 Clayton, N. S., and Dickinson, A. (1999). Motivational control of caching behaviour in the scrub  
1125 jay, *Aphelocoma coerulescens*. *Anim. Behav.* 57, 435–444. doi:10.1006/anbe.1998.0989.
- 1126 Clayton, N. S., and Emery, N. J. (2015). Avian Models for Human Cognitive Neuroscience: A  
1127 Proposal. *Neuron* 86, 1330–1342. doi:10.1016/J.NEURON.2015.04.024.
- 1128 Crosby, R. M. (2019). A comparative investigation of the attribution of desires and  
1129 preferences. PhD Thesis, University of Cambridge, UK.
- 1130 Dally, J. M., Emery, N. J., and Clayton, N. S. (2004). Cache protection strategies by western  
1131 scrub-jays (*Aphelocoma californica*): hiding food in the shade. *Proceedings. Biol. Sci.* 271  
1132 Suppl 6, S387-90. doi:10.1098/rsbl.2004.0190.
- 1133 Dally, J. M., Emery, N. J., and Clayton, N. S. (2005). Cache protection strategies by western  
1134 scrub-jays, *Aphelocoma californica*: implications for social cognition. *Anim. Behav.* 70,  
1135 1251–1263. doi:10.1016/J.ANBEHAV.2005.02.009.
- 1136 Dally, J. M., Emery, N. J., and Clayton, N. S. (2006). Food-caching western scrub-jays keep track  
1137 of who was watching when. *Science (80-. )*. 312, 1662–5. doi:10.1126/science.1126539.
- 1138 Dickinson, A., and Balleine, B. (1994). Motivational control of goal-directed action. *Anim.*  
1139 *Learn. Behav.* 22, 1–18. doi:10.3758/BF03199951.
- 1140 Drayton, L. A., and Santos, L. R. (2014). Capuchins’ (*Cebus apella*) sensitivity to others’ goal-  
1141 directed actions in a helping context. *Anim. Cogn.* 17, 689–700. doi:10.1007/s10071-013-  
1142 0700-5.

- 1143 Emery, N. J., and Clayton, N. S. (2001). Effects of experience and social context on prospective  
1144 caching strategies by scrub jays. *Nature* 414, 443–446. doi:10.1038/35106560.
- 1145 Emery, N. J., and Clayton, N. S. (2004). The mentality of crows: convergent evolution of  
1146 intelligence in corvids and apes. *Science* 306, 1903–1907. doi:10.1126/science.1098410.
- 1147 Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries.  
1148 *Scientometrics* 90, 891–904. doi:10.1007/s11192-011-0494-7.
- 1149 Farrar, B., Boeckle, M., and Clayton, N. (2020). Replications in Comparative Cognition: What  
1150 Should We Expect and How Can We Improve? *Anim. Behav. Cogn.* 7, 1–22.  
1151 doi:10.26451/abc.07.01.02.2020.
- 1152 Farrar, B., and Ostojic, L. (2019). The illusion of science in comparative cognition. *PsyArXiv*.  
1153 doi:10.31234/OSF.IO/HDUYX.
- 1154 FeldmanHall, O., and Shenhav, A. (2019). Resolving uncertainty in a social world. *Nat. Hum.*  
1155 *Behav.* 3, 426–435. doi:10.1038/s41562-019-0590-x.
- 1156 Fiedler, K., and Prager, J. (2018). The Regression Trap and Other Pitfalls of Replication  
1157 Science—Illustrated by the Report of the Open Science Collaboration. *Basic Appl. Soc.*  
1158 *Psych.* 40, 115–124. doi:10.1080/01973533.2017.1421953.
- 1159 Flombaum, J. I., and Santos, L. R. (2005). Rhesus Monkeys Attribute Perceptions to Others.  
1160 *Curr. Biol.* 15, 447–452. doi:10.1016/J.CUB.2004.12.076.
- 1161 Garner, J. P. (2005). Stereotypies and Other Abnormal Repetitive Behaviors: Potential Impact  
1162 on Validity, Reliability, and Replicability of Scientific Outcomes. *ILAR J.* 46, 106–117.  
1163 doi:10.1093/ilar.46.2.106.
- 1164 Güntürkün, O., and Bugnyar, T. (2016). Cognition without Cortex. *Trends Cogn. Sci.* 20, 291–  
1165 303. doi:10.1016/J.TICS.2016.02.001.
- 1166 Hare, B., Call, J., Agnetta, B., and Tomasello, M. (2000). Chimpanzees know what conspecifics  
1167 do and do not see. *Anim. Behav.* 59, 771–785. doi:10.1006/ANBE.1999.1377.
- 1168 Hare, B., Call, J., and Tomasello, M. (2001). Do chimpanzees know what conspecifics know?  
1169 *Anim. Behav.* 61, 139–151. doi:10.1006/ANBE.2000.1518.
- 1170 Hedges, L. V. (1984). Estimation of Effect Size under Nonrandom Sampling: The Effects of  
1171 Censoring Studies Yielding Statistically Insignificant Mean Differences. *J. Educ. Stat.* 9, 61–  
1172 85. doi:10.3102/10769986009001061.
- 1173 Heinrich, B., and Pepper, J. W. (1998). Influence of competitors on caching behaviour in the  
1174 common raven, *Corvus corax*. *Anim. Behav.* 56, 1083–1090.
- 1175 Heyes, C. (2015). Animal mindreading: what’s the problem? *Psychon. Bull. Rev.* 22, 313–327.  
1176 doi:10.3758/s13423-014-0704-4.
- 1177 Horowitz, A. (2011). Theory of mind in dogs? Examining method and concept. *Learn. Behav.*  
1178 39, 314–317. doi:10.3758/s13420-011-0041-7.
- 1179 Hothorn, T., Hornik, K., van de Wiel, M. A., and Zeileis, A. (2006). A Lego System for  
1180 Conditional Inference. *Am. Stat.* 60, 257–263. doi:10.1198/000313006X118430.
- 1181 Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., and Call, J. (2019). Great apes use self-  
1182 experience to anticipate an agent’s action in a false-belief test. *Proc. Natl. Acad. Sci. U. S. A.*  
1183 116, 20904–20909. doi:10.1073/pnas.1910095116.
- 1184 Krupenye, C., and Call, J. (2019). Theory of mind in animals: Current and future directions.  
1185 *Wiley Interdiscip. Rev. Cogn. Sci.*, e1503. doi:10.1002/wcs.1503.
- 1186 Legg, E. W., and Clayton, N. S. (2014). Eurasian jays (*Garrulus glandarius*) conceal caches from  
1187 onlookers. *Anim. Cogn.* 17, 1223–1226. doi:10.1007/s10071-014-0743-2.
- 1188 Legg, E. W., Ostojic, L., and Clayton, N. S. (2016). Caching at a distance: a cache protection  
1189 strategy in Eurasian jays. *Anim. Cogn.* 19, 753–758. doi:10.1007/s10071-016-0972-7.
- 1190 Maginny, M. E., and Grace, R. C. (2014). Visual perspective taking by dogs (*Canis familiaris*)  
1191 in a Guesser–Knower task: evidence for a canine theory of mind? *Anim. Cogn.* 17, 1375–  
1192 1392. doi:10.1007/s10071-014-0773-9.

- 1193 Moses, L. J., Baldwin, D. A., Rosicky, J. G., and Tidball, G. (2001). Evidence for Referential  
1194 Understanding in the Emotions Domain at Twelve and Eighteen Months. *Child Dev.* 72,  
1195 718–735. doi:10.1111/1467-8624.00311.
- 1196 Ostojić, L., Legg, E. W., Brecht, K. F., Lange, F., Deininger, C., Mendl, M., et al. (2017). Current  
1197 desires of conspecific observers affect cache-protection strategies in California scrub-jays  
1198 and Eurasian jays. *Curr. Biol.* 27, R51–R53. doi:10.1016/J.CUB.2016.11.020.
- 1199 Ostojić, L., Legg, E. W., Dits, A., Williams, N., Brecht, K. F., Mendl, M., et al. (2016). Experimenter  
1200 expectancy bias does not explain Eurasian jays' (*Garrulus glandarius*) performance in a  
1201 desire-state attribution task. *J. Comp. Psychol.* 130, 407–410. doi:10.1037/com0000043.
- 1202 Ostojić, L., Legg, E. W., Shaw, R. C., Cheke, L. G., Mendl, M., and Clayton, N. S. (2014). Can male  
1203 Eurasian jays disengage from their own current desire to feed the female what she  
1204 wants? *Biol. Lett.* 10, 20140042–20140042. doi:10.1098/rsbl.2014.0042.
- 1205 Ostojić, L., Shaw, R. C., Cheke, L. G., and Clayton, N. S. (2013). Evidence suggesting that desire-  
1206 state attribution may govern food sharing in Eurasian jays. *Proc. Natl. Acad. Sci. U. S. A.*  
1207 110, 4123–8. doi:10.1073/pnas.1209926110.
- 1208 Osvath, M., Kabadayi, C., and Jacobs, I. (2014). Independent evolution of similar complex  
1209 cognitive skills: the importance of embodied degrees of freedom. *Anim. Behav. Cogn.* 1,  
1210 249–264.
- 1211 Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008). Darwin's mistake: Explaining the  
1212 discontinuity between human and nonhuman minds. *Behav. Brain Sci.* 31, 109–130.  
1213 doi:10.1017/S0140525X08003543.
- 1214 Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav.*  
1215 *Brain Sci.* 1, 515. doi:10.1017/S0140525X00076512.
- 1216 Repacholi, B. M., Meltzoff, A. N., Rowe, H., and Toub, T. S. (2014). Infant, control thyself:  
1217 Infants' integration of multiple social cues to regulate their imitative behavior. *Cogn. Dev.*  
1218 32, 46–57. doi:10.1016/J.COGDEV.2014.04.004.
- 1219 RStudio Team (2018). RStudio: Integrated development environment for R.  
1220 <http://www.rstudio.org>.
- 1221 Santos, L. R., Nissen, A. G., and Ferrugia, J. A. (2006). Rhesus monkeys, *Macaca mulatta*, know  
1222 what others can and cannot hear. *Anim. Behav.* 71, 1175–1181.  
1223 doi:10.1016/J.ANBEHAV.2005.10.007.
- 1224 Scheel, A. M., Schijen, M. R. M. J., and Lakens, D. (2020). An excess of positive results:  
1225 Comparing the standard Psychology literature with Registered Reports.  
1226 doi:<https://doi.org/10.31234/osf.io/p6e9c>.
- 1227 Scott, R. M., and Baillargeon, R. (2017). Early False-Belief Understanding. *Trends Cogn. Sci.* 21,  
1228 237–249. doi:10.1016/J.TICS.2017.01.012.
- 1229 Seed, A., Emery, N., and Clayton, N. (2009). Intelligence in corvids and apes: A case of  
1230 convergent evolution? *Ethology* 115, 401–420. doi:10.1111/j.1439-0310.2009.01644.x.
- 1231 Shaw, R. C., and Clayton, N. S. (2012). Eurasian jays, *Garrulus glandarius*, flexibly switch  
1232 caching and pilfering tactics in response to social context. *Anim. Behav.* 84, 191–1200.  
1233 doi:10.1016/j.anbehav.2012.08.023.
- 1234 Shaw, R. C., and Clayton, N. S. (2013). Careful cachers and prying pilferers: Eurasian jays  
1235 (*Garrulus glandarius*) limit auditory information available to competitors. *Proceedings.*  
1236 *Biol. Sci.* 280, 20122238. doi:10.1098/rspb.2012.2238.
- 1237 Southgate, V. (2019). Are Infants Altercentric? The Other and the Self in Early Social  
1238 Cognition. *Psychol. Rev.* 127, 505–523. doi:10.1037/rev0000182.
- 1239 Stulp, G., Emery, N. J., Verhulst, S., and Clayton, N. S. (2009). Western scrub-jays conceal  
1240 auditory information when competitors can hear but cannot see. *Biol. Lett.* 5, 583–5.  
1241 doi:10.1098/rsbl.2009.0330.
- 1242 Tamir, D. I., and Thornton, M. A. (2018). Modeling the Predictive Social Mind. *Trends Cogn. Sci.*



- 1243 22, 201–212. doi:10.1016/j.tics.2017.12.005.  
1244 Wellman, H. M. (2018). Theory of mind: The state of the art. *Eur. J. Dev. Psychol.* 15, 728–755.  
1245 doi:10.1080/17405629.2018.1435413.  
1246 Wellman, H. M., and Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Dev.* 75, 523–541.  
1247 doi:10.1111/j.1467-8624.2004.00691.x.  
1248