

1 Recognition of rotated objects and cognitive offloading in
2 dogs

3
4 Lucrezia Lonardo^{1*}, Elisabetta Versace² and Ludwig Huber¹

5
6 ¹Comparative Cognition, Messerli Research Institute, University of Veterinary Medicine Vienna, Medical
7 University of Vienna and University of Vienna, 1210 Austria

8 ²School of Biological and Behavioural Sciences, Department of Biological and Experimental Psychology,
9 Queen Mary University of London, E1 4NS United Kingdom.

* Corresponding author and lead contact: lucrezia.lonardo@vetmeduni.ac.at

10 Summary

11 Recognition of rotated images can challenge visual systems. Humans often diminish the load
12 of cognitive tasks employing bodily actions (cognitive offloading). To investigate these
13 phenomena from a comparative perspective, we trained eight dogs (*Canis familiaris*) to
14 discriminate between bi-dimensional shapes. We then tested the dogs with rotated versions
15 of the same shapes, while measuring their accuracy and head tilts. Although generalisation to
16 rotated stimuli challenged dogs (overall accuracy: 55%), three dogs performed differently
17 from chance level with rotated stimuli. The amplitude of stimulus rotation did not influence
18 dogs' performance. Interestingly, dogs tilted their head following the direction and amplitude
19 of rotated stimuli. These small head movements did not influence their performance. Hence,
20 we show that dogs might be capable of recognising rotated 2D objects, but they do not use a
21 cognitive offloading strategy in this task. This work paves the way to further investigation of
22 cognitive offloading in non-human species.

23

24 Introduction

25 When engaging in mental tasks, we spontaneously try to overcome the limitations of
26 our cognitive systems by transferring part of the intellectual demands onto our bodies or onto
27 the external world. This strategy is known as cognitive offloading (for a review of the human
28 literature see Risko & Gilbert 2016). Little is known on cognitive offloading in non-human
29 animals. Are physical actions that reduce mental load a uniquely human strategy or a
30 behaviour shared across species? To answer this question, we assessed whether pet dogs
31 (*Canis familiaris*) spontaneously try to decrease the cognitive load of a mental rotation task
32 by tilting their heads in the same direction as the stimuli, like humans do.

33 Mental rotation has been hypothesised as a cognitive process used by humans to
34 recognise an object seen from an orientation different than a reference stimulus, by mentally
35 rotating and aligning the internal (mental) representation of the object with the reference
36 (e.g., Shepard & Metzler, 1971; Cooper & Shepard, 1973). However, the mental rotation
37 account has been challenged and may be outdated. In particular, it seems that this account is
38 best suited to explain results from tasks requiring to judge whether two objects are identical
39 or mirror images of each other (Gauthier et al., 2002). Several alternative models have been
40 proposed to explain how the human visual system recognises whether two objects seen from
41 different points of view are the same or not. For example, some researchers proposed that
42 object recognition might be based on more than just one processing mechanism (the "multiple
43 routes" hypothesis, Vanrie et al. (2001)) or that it might be based on interpolation between
44 the limited views of an object stored in memory (Edelman and Poggio, 1991; Riesenhuber and
45 Poggio, 2000). The strategies used by non-human species to recognize rotated objects are
46 debated too, as discussed below.

47 The cognitive load of recognizing the rotated stimuli in the generalization task can be
48 reduced by physical actions that reduce the degree of mismatch between the test and the
49 training stimulus, such as tilting the head (Risko & Gilbert, 2016). In humans, cognitive
50 offloading is common in different domains: people use reminders for future events, rotate
51 maps to match one's own point of view, count with the help of fingers, write things down, and
52 tilt their head during the perception of ambiguous images (Carlson et al., 2007; Chu and Kita,
53 2011; Gilbert, 2015; Goldin-Meadow et al., 2001; Risko et al., 2014). Indeed, cognitive

54 offloading strategies improve performance also for recognition of rotated objects, for instance
55 by increasing the speed of reading rotated text (Risko et al., 2014).

56 Two lines of evidence support the idea of an analogue representation and mental
57 rotation model in humans. First, a larger number of mistakes when the object and the
58 reference have a greater angular disparity (Bauer and Jolicoeur, 1996; Delius and Hollard,
59 1995; Foster, 1978; Hall and Friedman, 1994; Hollard and Delius, 1982; Parsons, 1987;
60 Wohlschläger, 2001; Wohlschläger and Wohlschläger, 1998). Second, a near-linear increase in
61 latency with rotation difference. In humans, the increment in reaction times has been
62 interpreted as the effect of an analogue mode of visual information processing in which an
63 object's mental representation is transformed in a serial, time-consuming process (Shepard
64 and Metzler, 1971; Stich et al., 2003). In a classical study, Shepard & Metzler (1971)
65 investigated how humans decide whether two differently rotated objects are the same or not.
66 Eight participants were presented with pairs of differently rotated drawings of solids, a
67 reference and a target. They judged whether the two objects were the same but rotated or
68 whether they were different objects (mirror images, that could not be rotated into
69 congruence). Rotations consisted of 20° steps, from 0° to 180°. On average, only 3.2% of the
70 responses were incorrect (ranging from 0.6 to 5.7% for individual subjects). The time used to
71 make the judgement was a linear function of the angular disparity between the two figures.
72 Based on this result, the authors described mental rotations as an analogue transformation
73 process of image-like representations of visual information.

74 Neuroimaging data have supported both the analogue representation view and the
75 hypothesis that mental rotation depends on motor simulation, i.e. the planning of motor
76 processes (Zacks, 2008). Mental rotation could be considered an imagined (covert) action or
77 at least partly produced in conjunction with the motor system (Lamm et al., 2007; Wexler et
78 al., 1998; Wohlschläger, 2001).

79 With analogue mental representations, the cognitive load imposed by the effort of
80 mental rotation can be reduced either by rotating the objects in the external world or by tilting
81 the head of the observer. The conditions under which humans engage in cognitive offloading
82 while processing rotated stimuli have been investigated by Risko et al. (2014). In these
83 experiments, participants were asked to read rotated letters and text. This task demands
84 "normalization" of the viewpoint, i.e., alignment of the rotated stimuli to their canonical
85 orientation. When presented with sets of 1, 5 or 15 rotated letters, participants spontaneously
86 exhibited head tilts on approximately 16% of the trials but this tilting did not improve nor
87 hinder their reading accuracy. Head tilts were exhibited more frequently with increasing set
88 size (and hence increasing mental effort required by the task). Indeed, participants tilted their
89 head on average on 3% of the trials presenting only 1 letter; on 18% of the trials presenting 5
90 letters and on 37% of the trials presenting 15 letters. Risko and colleagues therefore showed
91 that human head tilts can be systematically investigated as an instance of cognitive offloading
92 in a controlled environment.

93 While cognitive offloading has been extensively investigated in humans, much less is
94 known on non-human species. Pigeons (*Columba livia*) can discriminate mirror-image shapes
95 equally fast and well regardless of orientation disparities, a skill known as orientation
96 invariance, presumably based on a parallel mode of information processing (Hollard and
97 Delius, 1982). Even highly intelligent humans could not match the birds' performance. Several
98 types of stimuli did not lead to a rotation effect in pigeons: novel mirror-image stimuli,
99 rotation of sample shapes, a delayed display of comparison shapes, and a mixed use of original
100 and reflected sample shapes (Delius and Hollard, 1995). Interestingly, with misaligned
101 arbitrary shapes, humans failed to show a mental rotation effect, similarly to pigeons. This

102 finding led Delius and Hollard (1995) to conclude that the complete absence of a rotation
103 effect in pigeons is due to an advantage in discriminating mirror-image shapes compared with
104 arbitrary shapes. It is possible that humans perceive the orientation differences of arbitrary
105 shapes but are not obstructed by them in the same way as when discriminating mirror-image
106 shapes.

107 Evidence coming from the arboreal living lion-tailed macaque seems to hint at a hybrid
108 status of monkeys' information processing mode. Indeed, these macaques recognized non-
109 rotated stimuli faster than rotated ones, but showing no clear relationship between reaction
110 times and angle of rotation (Burmam et al., 2005). The testing of a more terrestrial living
111 Rhesus monkey have also yielded inconsistent results, with some evidence for both processes,
112 mental rotation and rotational invariance (Köhler et al., 2005), supporting the view of two
113 separately evolved information processing systems that may be coexisting to a certain extent
114 in species with correspondingly overlapping ecological demands.

115 Delius and Hollard (1995) have speculated on why humans do not benefit from the
116 potential rotational invariance capability of the primate visual system by suggesting bio-
117 evolutionary adaptations due to special demands of the lifestyle. While pigeons operate
118 visually on the horizontal ground plane both in flight and walking, humans, who have
119 abandoned the arboreal lifestyle of our primate ancestors, mainly operate visually on the
120 vertical plane. Due to an upright gait, humans mostly see the environment in the vertical plane
121 and therefore are used to a rather restricted number of environmental perspectives. If
122 orientation invariance is neurally elaborate and costly, humans might have secondarily lost it.
123 The fact that pigeons – a species lacking hands or similar effectors allowing continuous object
124 rotation – are not affected in their discrimination performance by the rotation of stimuli has
125 also been interpreted as consistent with the view that the motor system might play a role in
126 the mental rotation process (Wohlschläger and Wohlschläger, 1998).

127 In non-human animals, no study has investigated the link between mental rotation and
128 cognitive offloading. The only study with at least an implicit answer to this question was by
129 Hollard and Delius (1982) who reported that, while recognizing rotated images, pigeons rarely
130 inclined their heads by more than 30° and that the head position did not relate in any obvious
131 way with the orientation of the forms. However, pigeons might benefit from a kind of parallel
132 processing, with which they achieve rotational invariance, thus they might have no need for
133 facilitating the internal transformation to bring the mental representation into alignment with
134 the object.

135 Here we focus on pet dogs, to clarify whether a species that has lived in the same
136 household as humans for more than 14 thousand years (Janssens et al., 2018) has evolved
137 similar capacities for recognizing rotated objects and cognitive offloading. Based on the shared
138 evolutionary history of ecological constraints on perception between dogs and humans, we
139 expected dogs to recognise rotated stimuli similarly to humans. Therefore, we predicted that
140 their accuracy should have decreased with increasing angular disparity between probe and
141 reference. In our setting, it was not possible to measure dogs' reaction times because our
142 main focus was observing their possible head tilts prior to choice. Hence, we let dogs wait a
143 fixed time interval before allowing them to give a response.

144 To date, apart from humans, no purely terrestrial mammal has been tested for their
145 ability to generalize to rotated stimuli. An interesting mammalian model is the domestic dog,
146 as these animals do not only see a very similar environment as their human caregivers, but
147 they also perceive many objects in the vertical plane due to gravity.

148 Domestic dogs have been tested in many visual tasks, ranging from simple (e.g.,
149 Milgram et al., 1994; for a comprehensive review see Bensky et al., 2013) to more complex

150 discriminations. At a larger scale, they have proven the ability to discriminate between visual
151 classes, such as dog and landscape images, according to a perceptual response rule (Range et
152 al., 2008). In sum, dogs can clearly learn to discriminate between various arbitrary stimuli
153 based on differential reward contingencies, as reviewed in Byosiere et al. (2018). In the
154 present study, we trained our dogs on a visual discrimination task, in which they had to
155 distinguish between two abstract geometrical shapes having the same area and colour.

156 Here we modified for dogs an object recognition task to test the ability of pet dogs to
157 first acquire a discrimination of two geometrical shapes and then transfer this ability to
158 rotated versions of the same shapes. By looking at the number of errors in relation to different
159 degrees of stimulus rotation, we aimed at understanding whether dogs exhibit a human-
160 typical performance (lower performance at larger rotation distance), or a pigeon-like rotation
161 invariance performance. If dogs are able to discriminate between the two shapes when these
162 are presented at rotation angles different from the training one, evidence that the task is
163 solved using a human-like strategy would be that (1) they are more accurate for angles of
164 stimuli rotation closer to the training orientation and (2) they exhibit wider head tilts for
165 greater stimuli rotation angles, in an attempt to relieve the increasingly demanding cognitive
166 process.

167 The second aim of the study was to examine the use of cognitive offloading with
168 rotated stimuli. Only recently, Sommese et al. (2021) found a relationship between head-
169 tilting and the processing of auditory stimuli, but only if those had been relevant and
170 meaningful. From their data, the authors concluded that head tilts are a sign of increased
171 attention. Dogs might tilt the head also as a kind of external way of alignment instead of
172 rotating an internal (mental) representation of the presented stimulus before making the
173 judgment. If the mechanism that transforms an input shape into the orientation of the
174 presented shape is cognitively demanding, dogs might use the cognitive offloading strategy to
175 simplify the task.

176 A third aim of the experiment was to explore sex and individual differences. A study
177 investigating how dogs respond to a violation of size constancy (Müller et al., 2011) found
178 surprisingly large sex differences. Female dogs looked significantly longer when the size of a
179 rolling ball seemed to “magically” change after rolling temporarily behind a barrier while
180 males did not. The authors suggested the existence of cognitive differences between sexes in
181 their task is a by-product of other sex differences in spatial cognition. Building on this finding,
182 we were interested in testing whether sex differences would have emerged in this mental
183 rotation task as well and, if so, if they would have emerged in the same direction, with females
184 outperforming males. Moreover, not only individual differences in performance but also in
185 the adoption of different strategies in solving the mental rotation task have been found with
186 Rhesus monkeys (Köhler et al., 2005). Similarly, we expected to find large inter-individual
187 variation in dogs too and hence have analysed each subject’s performance separately.
188

189 **Results**

190 **Learning curves**

191 The tested dogs needed between 15 and 127 sessions to learn the visual discrimination
192 task (see Table S12 for individual results). The huge inter-individual variability in learning
193 speed is portrayed in Figure S1 (Supplementary Materials).

194 **Accuracy during training**

195 The sample for this model comprised 23 dogs, trained over 1159 sessions for a total of
196 22919 trials. The number of correct choices was 13621, while the number of incorrect choices
197 was 9298 (see Table S1 for each of the tested dogs' accuracy during training). Each dog was
198 trained for at least 30 sessions. The vast majority of sessions consisted of 20 trials. However,
199 a minority of sessions were terminated before the dog completed all 20 trials. This happened
200 mainly due to system malfunctioning or if the dog stopped spontaneously approaching the
201 touchscreen.

202 Overall, there was no effect of sex nor of its interaction with session number on the
203 proportion of correct choices (likelihood ratio test comparing full and null model: $\chi^2(2) = 1.526$,
204 $P = 0.466$). None of the interactions was significant, as shown in Table S2.

205 There was no main effect of sex on accuracy and the interaction between sex and
206 session number was not significant either. There was no main effect of age on accuracy and
207 the interactions between age and age squared and session number were not significant either
208 (Table S3). This suggests that, during training, there was no difference in the performance of
209 dogs (both females and males) of all ages (from 5 months to 14 years), who reached similar
210 levels of accuracy. The only significant effect was that of session number (Table S4). As the
211 number of training sessions increased, dogs became significantly more accurate in
212 discriminating between the two upright shapes.

213 **Accuracy during test**

214 Dogs were tested on a total of 3264 upright trials (rotation: "none") and on 576 rotated
215 trials (half of which with clockwise rotations). See Table S5 for each of the 8 tested dogs'
216 accuracy with upright and rotated stimuli and Table S6 for their performance with each angle
217 of rotation.

218 Because the full-null model comparison was significant ($\chi^2 = 6.993$, $Df = 2$, $p\text{-value} = 0.030$), we
219 further proceeded in testing the significance of the individual predictors on accuracy.

220 First, direction of rotation had a significant influence on accuracy ($\chi^2(2) = 17.127$, $p\text{-value} < 0.001$). This effect was due to a significantly worse performance with rotated stimuli
221 relative to upright stimuli (Table S6). Indeed, all dogs were less accurate with rotated stimuli
222 than with upright stimuli. The proportion of correct responses averaged among the 8 subjects
223 was approximately 88% for upright stimuli and 55% for rotated stimuli.

224 Second, there was no effect of the angle of stimulus rotation on accuracy (Table S6,
225 predictor "Rotation"). Hence, within the rotated stimuli, the wideness and direction of
226 rotation ($\pm 45^\circ$, $\pm 90^\circ$, $\pm 135^\circ$) did not have an influence on performance (Table S6 and Figure 1).
227 This means that performance was similarly inaccurate for smaller and wider stimulus
228 rotations. We additionally compared the accuracy with clockwise and counter-clockwise
229 rotated stimuli using a Wald test. This confirmed that performance was similarly inaccurate
230 for clockwise and counter-clockwise rotations (pairwise comparison ccw-cw rotations:
231 $z = 0.342$, $p\text{-value} = 0.732$).

232 Third, we found an effect of sex on performance, with females being significantly more
233 accurate than males. As in the training phase, age did not have an influence on performance.
234 Finally, there was an effect of session number on accuracy: with increasing session number,
235 accuracy increased as well.

236 A post-hoc analysis of solely the trials that presented rotated stimuli revealed no
237 significant effect of session number on performance (estimate \pm SE = 0.085 ± 0.087 , $\chi^2(1) =$
238

239 0.948, p-value= 0.330). Therefore, over the course of the 24 test sessions, dogs improved their
240 performance in rewarded trials (upright stimuli) but not in unrewarded trials (rotated stimuli).
241 Interestingly, analysing the performance with rotated stimuli only, the sex effect disappeared
242 (estimate \pm SE= χ^2 (1) = 0.522, p-value= 0.470) too, meaning that females were more accurate
243 than males only with upright reinforced stimuli but not with the novel, rotated and
244 unrewarded stimuli.

245 The estimate of stimulus rotation was approximately 0.02, while the estimated
246 standard deviations among reinforced stimulus varied (depending on the random slope within
247 reinforced stimulus) to a maximum of 0.61. This indicates that the estimated variation among
248 different reinforced stimuli was large compared to that among different angles of stimuli
249 rotation.

250

251 **Head tilts and their influence on performance**

252 The average head tilt after the stimulus onset was 4.66° wide (SD= \pm 3.94) and we
253 observed heads being inclined by more than 10°, after the stimulus onset, in only 43 of 508
254 trials.

255 At a group level, we found a significant effect of the degrees of stimulus rotation on head
256 rotations measured after the stimulus onset (χ^2 (1)= 5.48, P=0.019). This means that, as
257 expected, head rotation was influenced by stimulus rotation in a linear fashion, with dogs
258 rotating their heads in the same direction as the stimulus rotation (clockwise or counter-
259 clockwise) and with wider head tilts in response to wider stimulus rotations. We did not find
260 any effect of sex on the wideness of head tilts (χ^2 (1)= 0.29, P=0.589), as shown in Figure 3 (see
261 also Table S7). We found that only one dog's (subject 29) wideness of head tilts was linearly
262 influenced by the angle of stimulus rotation (see Table S10 and Figure S2).

263 Overall, there was no effect of the test predictors on the difference between dogs'
264 head rotations after and before the stimulus onset (likelihood ratio test comparing full and
265 null model: χ^2 (2)= 1.51 , P= 0.469). Likewise, in contrast to the previous model, we did not find
266 an effect of the angle of stimulus rotation when we using the difference between head
267 rotation after and before the stimulus onset as dependent variable (χ^2 (1)= 1.512, P= 0.219).
268 Similarly to the previous model, sex did not affect (χ^2 (1)= 0.001, P= 0.973; see Table S8).

269

270 Overall, there was no effect of the test predictors on the proportion of correct choices
271 (likelihood ratio test comparing full and null model: χ^2 (2)= 0.64 , P= 0.727). Likewise, the
272 absolute difference between the degrees of head tilt and the degrees of rotation of the stimuli
273 and sex had no influence on performance (Table S9). These results suggest that the wideness
274 of head tilts did not have any influence on accuracy.

275

276 **Discussion**

277 Little is known on the link between the recognition of rotated objects and cognitive offloading
278 in non-human animals. We investigated how dogs process rotated stimuli by analysing their
279 generalization performance with rotated versions of the training stimuli and their head tilts as
280 a means to ease the cognitive load of the seemingly difficult task.

281 Overall, dogs' generalisation performance with rotated stimuli decreased compared to
282 their performance with upright stimuli. The overall performance with rotated stimuli was

283 significantly different from chance level for three (38%) of the eight tested dogs: significantly
284 above chance level for two dogs and significantly below chance level for one dog. These results
285 show that dogs are capable of solving a task similar to those that humans solve with mental
286 rotation.

287

288 We investigated whether dogs recur to external normalization (head tilting) while
289 observing rotated stimuli from a fixed position on a chinrest. We found that, while using a
290 chinrest, dogs tilted their heads very rarely. They made small adjustments in their head
291 positions after stimuli were presented. However, the angle of stimulus rotation had a linear
292 influence on dogs' head tilts. At the individual level, head rotations of only one dog (dog 29,
293 the one performing significantly below chance level in the recognition of rotated objects task)
294 were linearly associated to the angle of stimulus rotation. Overall, this evidence suggests that
295 dogs might use a cognitive offloading strategy. However, the effect of stimulus rotation on
296 head tilts was not significant when we considered as dependent variable the difference in
297 head tilt between after and before the stimulus onset. This might be due to the fact that dogs
298 straightened their head out for smaller rotations of the stimuli but left their head tilted for
299 wider stimulus rotations. In any case, head tilting did not impact accuracy with rotated stimuli.

300 It has been proposed that differences in the visual information processing systems
301 across species might have emerged in response to the ecological and evolutionary demands
302 of adapting to different ecological niches (Delius and Hollard, 1995; Köhler et al., 2005). Given
303 that dogs have massively shared their habitat with humans, we expected similarities in the
304 strategies used by the two species in this experiment. In different studies involving mental
305 rotation and recognition of rotated objects, human participants have shown a tendency to
306 commit more mistakes as angular disparity increased (Bauer and Jolicoeur, 1996; Delius and
307 Hollard, 1995; Foster, 1978; Hall and Friedman, 1994; Hollard and Delius, 1982; Parsons, 1987;
308 Wohlschläger, 2001; Wohlschläger and Wohlschläger, 1998). Based on these findings, our
309 initial prediction was that dogs would have been less accurate with wider rotations of stimuli,
310 which would be consistent with a mental rotation process, as described in humans and in a
311 sea lion (Mauck and Dehnhardt, 1997). However, our results did not support this hypothesis.
312 Indeed, differently from what observed in humans, the different angles of stimulus rotation
313 did not influence dogs' accuracy.

314 Although all dogs were more accurate in discriminating upright stimuli than rotated
315 stimuli, consistently with the performance observed in humans (Hollard & Delius, 1982), a sea
316 lion (Stich et al., 2003), rhesus monkeys (Köhler et al., 2005) and a lion-tailed macaque
317 (Burmam et al., 2005), we found no linear effect of the amplitude of stimulus rotation on
318 performance. In addition, while all subjects performed significantly above chance level with
319 upright stimuli, at the group level their performance with rotated stimuli did not differ
320 significantly from chance. When we looked at the effect of the degree of stimulus rotation,
321 we found no differences for clockwise and counter-clockwise rotations of different angles
322 (Figure 1). It is possible that a mental-rotation-like effect of stimulus rotation on generalization
323 performance would appear only when testing a group of subjects who perform above chance
324 level with rotated stimuli. However, also undergraduate students' accuracy was not
325 influenced by increasing angular disparity in a letter naming task (Risko et al., 2014), probably
326 due to the simplicity of the task. Indeed, the students' average mistake rate when reading
327 blocks of 15 letters was 10.0% for upright letters, 11.7% for 45° rotated letters and 8.8% for
328 90° rotated letters. Hence, it is possible that an effect of stimulus rotation on accuracy would
329 be best observed for intermediate levels of task difficulty. This possibility should be addressed
330 in further studies.

331 We also investigated sex and individual differences. We found no sex difference in the
332 accuracy with rotated stimuli, while females were significantly more accurate than males
333 during the test trials that presented the training (upright) shapes. We found large inter-
334 individual variability in the number of sessions that dogs needed to learn the visual
335 discrimination task. However, no sex nor age differences emerged during training, meaning
336 that females and males of all ages reached comparable levels of accuracy during the training
337 phase, at similar speed. Previous research based on owners' reports (e.g., Hsu and Serpell,
338 2003; Kubinyi et al., 2009; Serpell and Hsu, 2005), identified "trainability" as one of the factors
339 explaining dogs' inter-individual differences. The definition of trainability included both the
340 willingness to obey to already acquired commands and the speed, distractability and
341 resistance to correction when learning new tasks. According to this definition, none of these
342 studies found conclusive sex differences over large samples. Based on our results, it is possible
343 that the distinction between a training (when the behaviour to be performed is not
344 understood yet) and a test phase (when the behaviour has already been acquired) should not
345 be overlooked. Indeed, males and females might differ in their motivation to comply with
346 tasks but not in their ability to learn them. Only the eight dogs that reliably acquired the visual
347 discrimination task were tested on the rotated versions of the same stimuli.

348 The plots on individual performance (Figure 2) show that at least one female dog (dog
349 21) passed the generalisation test consistently recognising above chance level the reinforced
350 stimulus even when this was rotated of all different angles, as indicated by the confidence
351 intervals for the fitted values being above 0.5. Because the model fitted to the data is blind to
352 the performance being significantly above or below chance level, we also compared each
353 individuals' overall accuracy with rotated stimuli to chance level, using binomial tests. The
354 results showed that one additional male dog (dog 22) performed significantly above chance
355 level with rotated stimuli and one female (dog 29) performed significantly below chance level
356 with rotated stimuli. Based on the performance of dogs 21 and 22, who selected the
357 reinforced stimulus overall above chance level even when stimuli were rotated, we conclude
358 that dogs have the (neuro-cognitive) potential to recognise rotated bi-dimensional objects.
359 This result implies that dogs might have the capability to build a mental representation of the
360 training stimuli, to remember it even when the stimuli are no longer present (during the test),
361 and to compare it with the rotated point of view presented during test.

362 Interestingly, the two dogs performing above chance level were trained and tested on
363 the same pair of stimuli (pair 2 in Figure 5), whereas other dogs (7 and 20) did not pass the
364 generalisation test with the same pair. On the other hand, the dog who performed significantly
365 below chance level was trained and tested on the same stimuli (pair 3 in Figure 5) as the other
366 two dogs (10 and 11) who tended to perform below chance level with counter-clockwise 90°
367 rotations. The fact that dogs were misled by these specific stimuli being rotated of 90° can be
368 explained in terms of local attention. The upright pair of stimuli is reported in the upper part
369 of Figure S3. When this pair is rotated (lower part of Figure S3), a pattern present on the
370 upright negative stimulus is recreated on the rotated reinforced stimulus. It is possible that
371 dogs focused mainly on the highlighted part of the negative stimulus and consequently
372 discarded the 90° rotated positive stimulus.

373 Taken together, these results suggest that dogs might use individual strategies in
374 solving this recognition of rotated objects task and that their performance was influenced by
375 the specific stimuli, probably due to a local rather than a global focus of attention.

376 Individual differences in the strategies employed to solve a mental rotation task were
377 already reported for three Rhesus monkeys (Köhler et al., 2005). Based on these, the authors
378 speculated about the possible co-existence of two different mechanisms of visual processing
379 (mental rotation and rotational invariance) within the same species. We currently cannot
380 assess whether this variability characterises other species as well due to the prevalence of
381 single-case studies (e.g., Burmann et al., 2005; Stich et al., 2003) and group-level only analyses
382 (e.g., Hollard & Delius, 1982; Hopkins et al., 1993; Shepard & Metzler, 1971).

383 Pitteri et al. (2014) tested dogs on the Navon task with compound stimuli (Navon,
384 1977) and found great inter-individual variability in global/local precedence. The authors
385 concluded that the global/local strategy of processing in dogs is determined by the individuals'
386 life experiences in using visual information more than by a predisposition of the whole species.
387 Hence, our eight dogs' different life experiences until the beginning of the experiment might
388 explain the inter-individual variability observed in the strategies to solve our task.
389 Approximately two years later, Mongillo et al. (2017) re-tested the same dogs on the same
390 task as Pitteri et al. (2014) and found a clearer global bias than in the original study. Instead,
391 in a subsequent experiment, dogs were trained on a novel composite stimulus and, when
392 tested, did not show an overall global precedence. In fact, they tended to invert their strategy
393 relative to the original study. On average, despite being slightly higher than wide, our stimuli
394 subtended a comparable area as those in these two studies.

395 We currently cannot pinpoint the strategy that dogs used to solve this task. In contrast
396 to what is to be expected from the mental rotation and the object recognition accounts, dogs'
397 accuracy did not decrease systematically with increasing stimulus rotation. Moreover, dogs'
398 overall inaccuracy with rotated stimuli suggests that rotational invariance is unlikely to explain
399 their performance. For better comparison with the existing literature, in a future experiment,
400 it would be interesting to add mirror images discrimination to the task and to introduce three-
401 dimensional perspective drawings of stimuli. Furthermore, a fundamental variable to measure
402 in this kind of tasks is the reaction time. Indeed, previous studies with other species have
403 mainly distinguished between rotational invariance and mental rotation processes based on
404 reaction times. And even the well documented sex difference found in humans is mainly due
405 to males being faster – though not more accurate – than females at all ages through the
406 human lifespan (Linn and Petersen, 1985).

407 The amplitude and occurrence of head tilts in this setting were very low. Unlike
408 previous studies (Risko et al., 2014; Sommesse et al., 2021) we did not code head tilts
409 dichotomously as present/absent but rather we measured how inclined dogs' heads were
410 before and after the stimulus onset. Indeed, we were interested in any head rotation around
411 the Z-axis, even very small ones, because those might already facilitate the recognition of
412 rotated objects. Moreover, while Risko et al. (2014) coded as head tilts only rotations larger
413 than 10° and those that happened within one second from the stimulus onset, we coded the
414 degrees of dogs' head tilt within three seconds from stimulus onset, in each trial presenting
415 rotated stimuli. We observed spontaneous head rotations wider than 10° on approximately
416 8% of the trials, while Risko et al. (2014) report that human participants, with unrestricted
417 heads, exhibited this behaviour on average in 16% of the trials of a letter naming task
418 presenting stimuli rotated of 0°, ± 45° or ± 90°. The limited movements we observed might be
419 due to the chinrest inhibiting wider actions, apart from obvious anatomical differences
420 between dogs and humans. Also, the extreme difficulty of the task might explain the floor
421 effect in the number of attempts dogs made to externally normalise the pictures. Indeed, if

422 dogs did not recognise that the test stimuli consisted of rotations of the familiar (training)
423 ones, they might have not realised that tilting their heads would decrease the cost of the
424 mental task.

425 In our study, dogs did not exhibit head tilts in a preferred direction, but rather tended
426 to tilt their heads in the direction in which the stimuli were rotated. Moreover, most of the
427 observed head movements were characterised by small amplitude (average amplitude
428 observed after stimulus onset: ca. 5°) and in general, they occurred at a low rate. Taken
429 together, these elements let us speculate that the behaviours observed in this study in
430 response to visual (rotated) stimuli and the head tilts observed by Sommesse et al. (2021) in
431 response to auditory stimuli might subtend different mechanisms and serve different
432 functions. However, it is important to notice that in our controlled setting dogs could have not
433 exhibited such wide head tilts (maximum head tilt observed after stimulus onset: ca. 26°)
434 without lifting their head from the chinrest, a behaviour they rarely exhibited, probably due
435 to the previous training.

436 The limited and very brief head tilts we observed in the current study might be
437 indicative of an attempt to ease the cognitive demands of the task, but they are also consistent
438 with other interpretations. For example, the dogs might have been impatient to make a choice
439 during the observation period and tilting their head could have constituted the first step in
440 trying to overcome the chinrest that prevented them from moving forward. However, we
441 found that the degrees of stimulus rotation had a significant effect on the wideness of head
442 rotations measured after the stimulus onset. This means that, after the stimulus onset, dogs'
443 heads were more inclined for wider stimulus rotations and they were inclined in the same
444 direction (clockwise or counter-clockwise) as the stimuli.

445 The wideness of dogs' head tilts did not have an influence on accuracy in our task. The
446 lack of difference in the amplitude of head tilts exhibited by females and males is consistent
447 with the absence of sex differences in dogs' accuracy with rotated stimuli. In Risko et al.
448 (2014)'s experiments, no effect of head tilts was found on participants' accuracy in a letter
449 naming task, while both spontaneous and forced head tilts improved performance when
450 participants had to read whole paragraphs rather than single letters. However, in Risko et al.
451 (2014)'s third experiment, participants heads were unrestrained, unlike our dogs' heads.
452 Moreover, due to our relatively small sample size and due to the different level of task
453 difficulty, it is possible that our study is underpowered to show the effect of head tilts on
454 accuracy.

455 External normalisation (e.g., head tilting) needs not improve performance to be
456 considered an instance of cognitive offloading. However, given that tilting the head did not
457 improve dogs' accuracy, it remains unclear whether such a behaviour offloads on the body an
458 internal computation (Risko et al., 2014). Hence, we conclude that our results provide no
459 evidence for cognitive offloading in dogs.

460 In conclusion, although recognition of rotated two-dimensional shapes proved to be
461 challenging in our setting, we showed that dogs have the potential to solve the task. We
462 additionally showed that dogs' head tilts could be systematically studied in a controlled
463 setting, thus we provided a methodology for studying cognitive offloading in non-human
464 species. We did not find clear evidence that dogs tilted their heads as a means to offload a
465 cognitive process onto their bodies. Future research should investigate whether non-human

466 animals engage in cognitive offloading when facing moderately demanding tasks in this and in
467 other domains, such as memory.

468 **Limitations of the study**

469 In the present study, we chose to train dogs to lay their head on a chinrest before
470 presenting them with visual stimuli. We chose to stabilise the dogs' heads for different
471 reasons. First, we aimed at ensuring that the dogs would observe the stimuli consistently
472 across trials with regards to their body orientation and distance relative to the screen. Second,
473 we wanted to minimise dogs' impulsivity by forcing them to wait a pre-determined amount of
474 time before allowing them to walk forward and touch the screen. Finally, the dogs' head was
475 required to remain in a pre-determined and consistent position in space to ensure precise
476 coding of the video frames. The chinrest itself did not prevent dogs from tilting their head, as
477 proven by the (small) tilts we observed. However, we cannot know if the chinrest training
478 might have inhibited dogs' willingness to tilt their head more widely, whether this effect might
479 have been more pronounced for some individuals than others and if wider head tilts would
480 have influenced performance. Therefore, future studies assessing dogs' head tilts should leave
481 their heads unrestrained.

482 While previous studies have shown dogs are able to recognise 2D stimuli on the basis
483 of visual cues alone (e.g., Müller et al., 2015; Pitteri et al., 2014b), it is likely that, under more
484 ecological conditions, dogs do not need to rely exclusively on their visual modality to recognise
485 previously encountered entities. Therefore, dogs' scant performance with rotated stimuli
486 might be due to a true difficulty of this species with recognising familiar objects presented
487 from a rotated perspective, or to the lack of ecological validity of the task and abstract nature
488 of the stimuli. Future studies will need to assess which factors influence dogs' performance
489 (for example, biological relevance of the stimuli, ecological setting with cross-modal cues and
490 absence of postural constraints). While the visual discrimination between these particular
491 shapes and the rotated object recognition tasks might have been too difficult for most of the
492 dogs, in order to address the phenomenon of cognitive offloading, we had to confront subjects
493 with a challenging task. If the task had been too simple, the dogs would have had no need
494 to offload cognitive processing.

495 Finally, the limited sample size hinders the generalizability of our findings. In particular,
496 the post-hoc speculation about a possible female advantage in performing an already acquired
497 task but not in acquiring the task needs to be tested in future experiments across different
498 tasks and with larger sample sizes.

499 **Acknowledgements**

500 The authors wish to thank: Roger Mundry and Remco Folkertsma for their contribution
501 to data analysis; Karin Bayer, Peter Füreder, Michael Pichler and Wolfgang Berger for
502 administrative and technical assistance; Giulia Pedretti and Ana Grilc for their help with data
503 collection; Giorgio Vallortigara for fruitful discussions on the experimental design.

504 Funding for this study was provided by the Austrian Science Fund (FWF; W1262-B29),
505 by the Vienna Science and Technology Fund (WWTF), the City of Vienna and ithuba Capital AG
506 through project CS18-012, and the Messerli Foundation (Sörenberg, Switzerland). E.V.

507 received funding from the Royal Society Leverhulme Trust Senior Research
508 Fellowship SRF\R1\21000155.

509 **Author contributions**

510 Conceptualization: L.L., E.V., L.H.; Methodology: L.L., E.V., L.H.; Software: L.L.; Validation: L.L.,
511 Formal analysis: L.L.; Investigation: L.L.; Resources: L.H.; Data curation: L.L.; Writing: L.L., E.V.,
512 L.H.; Visualization: L.L.; Supervision: E.V., L.H.; Project administration: L.L.; Funding
513 acquisition: L.H.

514 **Declaration of interests**

515 The authors declare no competing interests.

516 **Main figure titles and legends**

517 Figure 1. Accuracy during test. Percentage of correct responses for each sex and direction of
518 stimulus rotation (“none” indicates performance with upright stimuli, “ccw” with counter-
519 clockwise rotated stimuli and “cw” with clockwise rotated ones). Females are represented on
520 the left, males on the right. The horizontal black lines indicate the fitted values and the error
521 bars refer to their confidence intervals. Each coloured bubble indicates the mean percentage
522 of correct responses of an individual dog with each rotation (each dog is shown with the same
523 colour across conditions). The area of the bubbles is proportional to the number of
524 observations per dog and condition. The grey dashed lines connect the observations of the
525 same individuals across conditions. The red dotted line shows the chance level. See also Table
526 S6.

527 Figure 2. Individual’s performance as a function of the angle of stimulus rotation. Females
528 are shown in the left column (from top to bottom: dog 11, 21, 29 and 37), males in the right
529 column (top to bottom: dog 7, 10, 20 and 22). The crosses indicate the mean observed
530 performance; the dashed line indicates the fitted values. Error bars represent the confidence
531 intervals of the fitted values. Negative numbers on the x-axis refer to counter-clockwise
532 rotations, while positive numbers refer to clockwise rotations. Dog 21 (second plot of the
533 left column, outlined in blue) achieved significantly above chance level performance with all
534 rotations. The chance level is indicated by the red dotted line. See also Table S5.

535 Figure 3. Degrees of head tilt as a function of stimulus rotation. On both the x- and the y-
536 axis, negative and positive numbers refer to counter clockwise and clockwise rotations. The
537 y-axis shows the degrees of head rotations measured after stimulus onset. Data are
538 represented as median \pm interquartile range. Outliers are represented as individual points
539 outside the whiskers of the boxplots. See also Figure S2 and Tables S7 and S8.

540 Figure 4. Experimental set-up.

- 541 A) The chinrest, opposite to the touchscreen, was operated through the metallic lever
542 on the right. Hence, the experimenter always stood on the dog’s right side.
- 543 B) One of the subjects shows the starting position, watching the screen with head on
544 the chinrest.
- 545 C) The automatic feeder behind the dog.

546 Figure 5. The three pairs of stimuli used in the whole experiment, portrayed in the upright
547 orientation. Each dog was randomly assigned to one of the pairs and, within the pair, to one
reinforced stimulus.

Figure 6. Examples of test stimuli. Pair 1 is shown in all 3 rotated versions, rotated by 45°
(top row), 90° (central row) and 135° (bottom row). Counter-clockwise rotations are shown
on the left, clockwise rotations on the right.

548 Figure 7. Examples of head tilt coding from video frames. Frame A) and B) come from the
549 same test trial. Frame A) is the one immediately preceding the stimulus onset, while frame
550 B) is the one containing the widest head tilt for that trial after the stimulus onset. The box in
551 the upper part of both frames shows the content of the screen as viewed from the dog's
552 perspective. The yellow lines on the dog's forehead show how the angles of head rotation
553 were measured.

554 **STAR Methods**

555 **Resource availability**

556 **Lead contact**

557 Inquiries should be addressed to the lead contact, Lucrezia Lonardo
558 (lucrezia.lonardo@vetmeduni.ac.at)

559 **Materials availability**

560 This study did not generate new unique materials

561 **Data and code availability**

- 562
- 563 • The raw datasets have been deposited and are currently available at:
564 <https://data.mendeley.com/datasets/khhkn6kcpm/draft?a=99aa5d91-cf7b-44cb-86b5-693eb60c2216>
565 The DOI is listed in the Key Resources Table.
 - 566 • The R code used to analyse and plot the data has been deposited and is
567 currently available at:
568 <https://data.mendeley.com/datasets/khhkn6kcpm/draft?a=99aa5d91-cf7b-44cb-86b5-693eb60c2216>
569 The folder "Training and learning curves" contains the scripts and data files for
570 Figure S1 and Tables from S1 to S4. The folder "Recognition of rotated
571 objects_accuracy during test" contains the script, workspace and data file for
572 Figures 1 and 2 and for Tables S5 and S6. The folder "Cognitive offloading_head
573 tilts" contains the script, workspace and data file for Figures 3 and S2 and Tables
574 from S7 to S10. The folder "R functions" contains the custom R functions used
575 for model stability, diagnostics and confidence intervals.
576 The DOI is listed in the Key Resources Table.
 - 577 • Any additional information required to reanalyse the data reported in this
578 paper is available from the lead contact upon request.
579

580 **Experimental model and subject details**

581 **Dogs (*Canis familiaris*)**

582 All dogs who took part in this experiment were pets, brought to the lab by their
583 volunteer owners. Prior to the beginning of the experiment, owners were informed about the

584 aim and procedures of the study, and gave an informed written consent for their dogs. Breed,
585 age, sex, previous touchscreen experience, number of training sessions and reinforced
586 stimulus of each of the 38 dogs are reported in the Supplementary Materials (Table S11).

587 Overall, 38 dogs (see Table S11) started in the experiment. Of these, five were excluded
588 from the study at an early stage due to behavioural issues that were not compatible with
589 learning or that posed a threat to the integrity of the experimental set-up (e.g., constant
590 barking, excessive impulsivity, fear of the wooden apparatus). Four of these five dogs were
591 excluded during the pre-training, the fifth after 3 training sessions. One additional dog passed
592 away. A further nine subjects dropped out of the study at different stages due to limitations
593 in owner availability. All other dogs completed at least 30 training sessions (600 trials) and
594 were included in the analysis of this phase. All other dogs completed at least 30 training
595 sessions (600 trials) and were included in the analysis of this phase. The resulting sample size
596 comprised therefore 23 dogs trained over a total of 1157 sessions. Of these 23 dogs, only 8
597 met the learning criteria (see Paragraph “Procedure” below) and were therefore tested with
598 rotated stimuli.

599 The 8 tested dogs (4 females) had a mean age of 7 years (age range 3-9 years). Seven
600 dogs had previously taken part in other experiments at the Clever Dog Lab. Table S12
601 summarises the main information regarding the 8 tested dogs. Dogs were randomly assigned
602 to one of the three pairs of stimuli shown in Figure 5.

603 The study was discussed and approved by the ethics and animal welfare committee of
604 the University of Veterinary Medicine of Vienna in accordance with GSP guidelines, national
605 legislation and EU regulations.

606 **Method details**

607 The experiment was conducted at the Clever Dog Lab, Messerli Research Institute,
608 University of Veterinary Medicine (Vienna). Dogs were trained and tested in the same 6 x 3 m
609 room, with the help of a semi-automated touchscreen and feeder. The automatic feeder used
610 was a Premier Treat & Train, filled with dry food pellets. It was positioned 1.20 meter behind
611 the dog (Figure 4C). The touchscreen was a Thin Film Transistor (TFT), with refresh rate of 60
612 Hz. It measured 46.5 x 27cm (height x wideness) and it was inserted in a white wooden
613 apparatus measuring 100 x 45.5 x 49 cm (height x depth x wideness). This minimised the
614 possibility of human cueing and distractions for the dogs (Figure 4A). To approach the
615 touchscreen, dogs had to walk on a black platform and lay their head on a chinrest (Figure 4B).
616 The chinrest ensured a standardised position of the dogs’ head at the beginning of each trial
617 (hence a rigorous video coding afterwards). The black platform measured 110 x 50 cm and the
618 chinrest 9.2 x 32.7 (l x w). This was a rubber foam pillow with v-shaped indentation in the
619 middle (depth: 2.2 cm). It was 5.5 cm deep on the sides. Dogs watched the stimuli on the
620 screen over a distance of approximately 50 cm. The experimental set-up is portrayed in Figures
621 4 and 7.

622 **Stimuli**

623 As stimuli, we adapted the two-dimensional Hollard and Delius (1982)’s shapes. We
624 changed the colour of the stimuli to blue because this is one of the two hues falling in dogs’
625 visible spectrum (violet and blue-violet range: 430 to 475 nm wavelengths). Each stimulus was
626 composed of nine adjacent squares forming an asymmetrical shape. The two shapes of each
627 pair differed in the positioning of four of the squares. Each square measured 4 cm² on the
628 screen. We used three different pairs of stimuli to assess whether different shapes would have

629 influenced dogs' accuracy. All stimuli were between 8 to 10 cm high and between 6 to 10 cm
630 wide. They subtended approximately between 9.2° to 11.4° (height) and between 6.9° to 11.4°
631 (width) of visual angle. Variations in wideness and height depend on the specific stimulus
632 considered. All the upright stimuli are reported in Figure 5 and a rotated pair is in Figure 6.

633 Mirror invariance, the tendency to identify two mirror images as the same picture
634 across left-right inversions, has been shown in humans and other primates (Biederman and
635 Cooper, 1991; Logothetis et al., 1995; Pegado et al., 2014). Hence, we did not include mirror
636 images in the task, as we wanted to increase the dogs' possibilities of seeing a difference
637 between the two alternatives.

638 **Procedure**

639 With one exception, the dog's owner was present in the room throughout the whole
640 experiment. To prevent interference, when present in the room, owners sat or stood more
641 than 2 meters behind their dog. During test sessions, owners sat with their backs towards their
642 dogs so that they could not see the screen, the experimenter nor the dog.

643 The experiment included three consecutive training phases (until dogs reached
644 predefined learning criteria, outlined in the following Paragraph) and a subsequent test phase.

645 **Training**

646 We trained pet dogs to discriminate between the elements of a pair of geometrical
647 shapes. Each dog was trained and tested on a single pair of stimuli. This means that,
648 throughout the whole experiment, only one stimulus (S+) was associated to a food reward
649 while the other (S-) never was. To avoid side biases, the left/right position of the S+ changed
650 pseudo-randomly, having no more than three consecutive presentations of the correct
651 stimulus on the same side.

652 Both training and test followed a simultaneous two-choice discrimination paradigm,
653 previously used in touchscreen experiments for dogs (e.g., Müller et al., 2015). Dogs were
654 trained to lay their heads on the chinrest and watch the screen prior to each trial. Once the
655 dog was in this initial position, the experimenter (standing on the right side of the apparatus)
656 presented the stimuli on the screen. After 3 seconds in which the dog watched the stimuli, the
657 experimenter manually moved the chinrest to the floor using the lever shown in Figure 4A.
658 Dogs were then free to step forward and touch the screen with their nose. After each touch,
659 the stimuli disappeared. If the choice was correct, and the trial was a rewarded one, a tone
660 and a food pellet were automatically emitted by the feeder behind the dog. Otherwise, the
661 experimenter lifted the chinrest back in place and started a new trial when the dog was ready
662 again. To shape this complex behaviour, training was sub-divided in the following stages:

663 1) **Approach.** Only for dogs with no or very little touch screen experience.

664 Dogs were trained to approach the touchscreen as soon as a stimulus (a large black dot)
665 appeared on the white screen. They were encouraged with food to touch the black dot with
666 their nose. Touch responses were rewarded with a dog food pellet that was automatically
667 dispensed from the feeder behind the dog. Each session consisted of 20 trials. When dogs had
668 reliably performed the approach-touch response in this phase, as judged by the experimenter,
669 they moved on to the pre-training.

670 2) **Pre-training: one shape - the reinforced stimulus.** For every dog, at least 5 sessions.
671 Dogs needed to touch the only stimulus appearing on the screen (one of the blue upright
672 shapes, figure 5) to get a reward from the automatic feeder. At this stage, there was no

673 possibility of doing wrong because if dogs touched the white screen, nothing happened while
674 if they touched the shape, they got a reward. For each dog, the shape appearing on the screen
675 in this phase was always the same and it was the reinforced stimulus for that dog. When dogs
676 reliably mastered this phase, as judged by the experimenter, they moved on to the next one.
677 Each session consisted of 20 trials.

678 3) **Simultaneous two choice discrimination (100% rewarded)**

679 Dogs needed to discriminate between two simultaneously presented stimuli, the correct
680 stimulus S+ (the same as seen in pre-training) and the S-, which is a different blue training
681 shape. The left/right position of the S+ changed pseudo-randomly with never more than 3
682 consecutive presentations of the same disposition. A reward was automatically dispensed for
683 every correct response.

684 When dogs chose the S-, the correction procedure started: the stimuli would
685 disappear and the screen would turn red until the experimenter pressed the forward button.
686 The chinrest was lifted back in place so that the dog could start the new trial with the head in
687 a standardised position. Each wrong choice was followed by a correction trial, presenting the
688 same configuration of stimuli again. Correction trials were excluded from the analysis of
689 performance.

690 Once a dog had reached the learning criterion (16 out of 20, i.e. 80% correct responses
691 within a session in each of 3 consecutive sessions), it moved on to the next phase. If dogs
692 found the discrimination too difficult, the experimenter could choose to insert a pre-training
693 (phase 2) session (only one image) to recover the dog's motivation. In addition, the
694 experimenter could decide to start a training day (phase 3) with a pre-training (phase 2)
695 session.

696 In this crucial training phase, dogs needed to learn to be persistent with their choice. If a dog
697 was stuck, the experimenter could cover (with her hand) the wrong alternative or she could
698 point to the right stimulus to make the task obvious for the dogs.

699 4) **Simultaneous two-choice discrimination training with partial reinforcement (85% 700 rewarded)**

701 In the last training phase, dogs needed to discriminate between two simultaneously presented
702 stimuli, the correct stimulus S+ (the same seen in pre-training) and the S-, which was a
703 different blue training shape. If dogs chose the S+, they had around 85% of chances of being
704 rewarded.

705 On average 3 out of 20 trials in each session were not rewarded. During these
706 unrewarded trials, nothing signalled the dog if they chose correctly or not. This partial
707 reinforcement phase familiarised dogs with the reward contingency of the test sessions, in
708 which 3 trials (those presenting the rotated stimuli) would have always been unrewarded.

709 Each dog was trained until they reached the learning criterion of 80% correct responses
710 in 3 consecutive sessions (48 correct/60 trials) or for at least 30 sessions without reaching this
711 criterion. Some of the dogs were trained for longer as the owners were available to continue
712 with the training. Only eight out of 23 dogs reached this learning criterion and moved on to
713 the test phase.

714 **Test**

715 Each test day started with a training session for rehearsal purposes. Then we
716 conducted the test sessions, consisting of 20 trials each. Within these 20 trials, 17 presented

717 the familiar training (upright) stimuli and 3 presented novel stimuli, which were the rotated
718 versions of both the S+ and the S-. With this schedule we prevented dogs from losing
719 motivation during the test sessions, as rotated stimuli were never followed by a reward.
720 The left-right position of the two stimuli on the screen was pseudo-randomly balanced.

721 During each test session, the 3 trials presenting rotated stimuli (test trials) were
722 pseudo-randomly interspersed among trials presenting upright stimuli (training trials), with a
723 test trial always followed at least by one training trial. Furthermore, the first trial in a test
724 session was always a training trial.

725 The transitional probabilities from one angle of rotation to another were balanced
726 across the 24 test sessions.

727 **Quantification and statistical analyses**

728 The touchscreen automatically scored whether the response was correct or not for
729 each trial.

730 For each dog, we looked at the individual learning curve by plotting the accuracy for every
731 session over the course of all training sessions (Figure S1).

732 A camera mounted on the top of the screen recorded the dogs' head while subjects
733 were watching the screen with the head positioned centrally and straight on the chinrest Fig.
734 4). Therefore, we were able to code precisely whether and how much dogs inclined their heads
735 while watching the rotated stimuli. We extracted 180 frames (60 per second) from the test
736 videos. The time window we considered went from the moment before the stimulus onset
737 until the experimenter's first movement to lower the chinrest (approximately 3 seconds,
738 during which dogs were observing the screen with their head on the chinrest). We visually
739 inspected the frames to determine the one with maximum head tilt (after stimulus onset). We
740 measured the angle of head tilt as the angle between the dogs' forehead-middle of the eyes
741 line and the vertical axis (angles drawn in yellow in Figure 7) by using the angle tool of the
742 software ImageJ (Schneider et al., 2012).

743 We calculated the difference between the head tilt visible in the frame with widest
744 rotation after the stimulus onset and the head tilt in the frame immediately preceding the
745 stimulus onset. To assess the agreement between two coders on this variable and on the
746 degrees of head tilts after the stimuli onset, we calculated Intra-class Correlation Coefficients
747 on a subset of 60 observations of different subjects. In detail, we used a two-way random-
748 effects model to assess absolute agreement between the two coders, one of which was blind
749 to the experimental hypothesis and conditions. Because we were interested in the reliability
750 of our dependent variables (as coded by a single rater) we calculated the ICC on single
751 measurements (ICC type: "single"). The resulting inter-raters reliability was acceptable (for
752 both variables, ICC: 78%; N=60, $p < 0.001$).

753 To validate the reliability of the automatic scoring made by the touchscreen, a second
754 coder also scored the dogs' accuracy from video recordings. We calculated an unweighted
755 Cohen's Kappa on 60 observations of different subjects. The agreement between touchscreen
756 and second rater was almost perfect (Kappa = 0.966, N=60, $p < 0.001$).

757
758 To analyse dogs' performance during training (phases 3 and 4), we fitted a generalized
759 linear mixed model (GLMM; Baayen, 2008) with binomial error structure and logit link function
760 (McCullagh and Nelder, 1989). We modelled the proportion of correct responses per session

761 as a function of sex and its interaction with session number (fixed effects). In addition to the
762 test predictors, we included the session number, age and age squared and also the
763 interactions between age and age squared, on the one hand, and session number, on the
764 other, as control predictors with fixed effect. We included the interaction between sex and
765 session number because a sex difference could manifest in one sex acquiring the capability of
766 performing successfully faster than the other. We included age squared because, given the
767 wide age range, we hypothesised that middle-aged dogs might have learned the task faster
768 (i.e., they might have reached higher levels of accuracy within fewer sessions). We included
769 dog identity, reinforced stimulus and session ID (nested within dog, as all subjects were
770 trained for months) as random intercept effects. Finally, in this and in all the following models,
771 all theoretically identifiable random slopes were included. This ensured that type-I error rate
772 was kept at the nominal level of 0.05 and avoided overconfidence in the precision of the fixed
773 effects estimates (Barr et al., 2013; Schielzeth and Forstmeier, 2009). Namely, in this model
774 we included: the random slope of session number within dog and that of sex, of age and of
775 the interaction between age squared and session number within reinforced stimulus.

776 We checked the distribution of the random effects and we verified that the model was
777 not overdispersed. The result showed no issue in this regard (dispersion parameter= 1.004).

778 To test the significance of the main effect of sex and its possible interaction with
779 session number, we used a likelihood ratio test (Dobson, 2002). This compared the fit of the
780 full model with that of a null model, lacking the fixed effects of sex and its interaction with
781 session number but retaining the same random effects structure and all other fixed effects
782 present in the full model (Schielzeth and Forstmeier, 2009). For this and the following models,
783 to draw inference about the individual predictors, we always used the function drop1
784 (Chambers and Hastie, 1992), which drops each fixed effect from the model (one at a time)
785 and uses a likelihood ratio test to compare the full with the respective reduced models (Barr
786 et al., 2013).

787 We assessed model stability with regards to the estimated coefficients and standard
788 deviations by excluding the levels of the random effects one at a time (Nieuwenhuis et al.,
789 2012). This revealed the model to be of good stability.

790 We fitted the model in R (version 3.6.3, R Core Team 2020) using the function glmer of
791 the package lme4 (version 1.1-21; Bates et al., 2015).

792
793 A second GLMM with the same error structure and link function was fitted to analyse
794 dogs' accuracy during test. We modelled the proportion of correct responses as a function of
795 the angle of rotation of the stimuli and sex (fixed effects). In addition to the test predictors,
796 we included direction of rotation (factor with levels: none, clockwise, and counter-clockwise),
797 session number and age as control predictors with fixed effect. We included dog identity,
798 reinforced stimulus, and session ID (nested within dog) as random intercept effects. The latter
799 allowed the possibility of variation among sessions within dogs as all dogs were tested across
800 several weeks. Finally, all theoretically identifiable random slopes were included in the model.
801 Namely, these were the random slope of direction of rotation, stimulus rotation and session
802 number within dog; direction of rotation, stimulus rotation, session number, sex and age
803 within reinforced stimulus; and, finally, direction of rotation and stimulus rotation within
804 session ID.

805 Initially we fitted a maximal model (Barr et al., 2013), including also the correlations
806 among random intercepts and slopes. However, all absolute correlation parameters were
807 close to 1 and hence not identifiable (Matuschek et al., 2017). Therefore, we fitted a second
808 model excluding the correlations. We assessed the capability of the two models to fit the data
809 comparing their log-likelihoods (-1545.569 (df=60) for the maximal model; -1559.429 (df=23)
810 for the model without the correlations). As removing the correlation parameters only led to a
811 moderate decrease in model fit, we used the model without correlations for further analysis.
812 Prior to fitting the model, we z-transformed rotation, session number and age to a mean of
813 zero and a standard deviation of one to increase the likelihood of the model to converge. We
814 manually dummy coded and centred the factors direction of rotation (levels: no rotation,
815 clockwise rotation and counterclockwise rotation; no rotation was set as reference category)
816 and sex before including them as random slopes.

817 We assessed model stability with regards to the estimated coefficients and standard
818 deviations by excluding the levels of the random effects one at a time (Nieuwenhuis et al.,
819 2012). This revealed the model to be of good stability. We bootstrapped model estimates
820 using the function bootMer of the package lme4 (Bates et al., 2015). We conducted an
821 additional bootstrap, conditioning in the particular levels of the random effects (setting the
822 argument use.u to “true”) which allowed to infer about the performance of the individual
823 dogs.

824 We fitted the model in R (version 3.6.3, R Core Team 2020) using the function glmer of
825 the package lme4 (version 1.1-21; Bates et al., 2015).

826 The sample for this model comprised 3840 observations, 576 of which were trials with rotated
827 stimuli. Each one of the 8 dogs contributed equally to the number of observations (480 trials,
828 obtained over 24 test sessions for each participant). The total number of incorrect choices was
829 657 while correct choices were 3183.

830 Because the results of this model suggested an effect of session number on accuracy,
831 we further analysed only the performance with rotated stimuli (that were never reinforced)
832 to test whether the effect of session number was driven by the reinforced trials with upright
833 stimuli. To run this post-hoc manipulation check, we fitted a second model, identical to the
834 first one but comprising only the 576 trials with rotated stimuli.

835 We additionally compared each individual’s overall performance with rotated stimuli
836 to chance level. To avoid multiple testing and the consequent risk of increasing the likelihood
837 of type I error, we did not run a significance test for each dog and angle of stimulus rotation.
838 Instead, to infer about individual performance, we used the confidence intervals for the fitted
839 values shown in Figure 2. Confidence intervals not comprising the value of 0.5 are indicative
840 of performance significantly above chance level.

841 Finally, to quantify the relative contribution of the reinforced stimulus on accuracy, we
842 compared the estimate of stimulus rotation (fixed effect) to the estimated standard deviations
843 within reinforced stimulus (random intercept).

844

845 We modelled the degrees of head tilt after the stimulus onset as a function of the
846 amplitude of stimulus rotation and sex using a linear mixed model (Baayen, 2008) with
847 Gaussian error distribution. . For both stimulus and head rotations, we transformed counter
848 clockwise rotations to negative numbers. Prior to fitting the model, we z-transformed stimulus

849 rotation (45°, 90° and 135°) to a mean of 0 and a standard deviation of 1. As we did not find
850 any effect of direction of stimulus rotation on accuracy (paragraph “results - accuracy during
851 test”), we did not include the direction of stimulus rotation (clockwise vs. counter clockwise)
852 in this model. We included dog identity and reinforced stimulus as random intercept effects.
853 Finally, all theoretically identifiable random slopes (stimuli rotation within subject and within
854 reinforced stimulus) were included in the model. We fitted the model in R (version 3.6.3, R
855 Core Team 2020) using the function lmer of the package lme4 (version 1.1-21; Bates et al.,
856 2015).

857 Additionally, we run a simple regression (general linear model with Gaussian error
858 structure and identity link) for each dog, to investigate also at individual level the effect of the
859 angle of stimulus rotation on wideness of the head tilts exhibited after the stimulus onset. As
860 described for the model above, we transformed counter clockwise rotations to negative
861 numbers and z-transformed the angles of stimulus rotation. To fit the models, we used the
862 function lm in R.

863 We checked the homogeneity and normality of the distribution of the residuals by
864 inspecting qqplots. Model stability was evaluated through standardised DFFit-values (which
865 compare the fitted values of a model using all data with those of a model with cases excluded
866 one at a time) and Cook’s distance, a measure of the influence of each data point on model
867 estimates (Queen and Keough, 2002)

868 To analyse the influence of head tilt on performance, we fitted a generalized linear
869 mixed model (GLMM; Baayen, 2008) with binomial error structure and logit link function
870 (McCullagh and Nelder, 1989). We modelled the proportion of correct responses as a function
871 of the absolute difference between the degrees of head tilt after the stimulus onset and the
872 degrees of head tilt before the stimulus onset. The larger this absolute angle, the more
873 accurate we expected dogs to be. We additionally included sex as test predictor and dog
874 identity and reinforced stimulus as random intercept effects. Finally, as in the previous
875 models, all theoretically identifiable random slopes (absolute angle of head rotation within
876 subject and within reinforced stimulus) were included.

877 To evaluate the main effect of head tilts and sex on accuracy we compared the fit of
878 the full model with that of a null model lacking these two effects in the fixed effects part using
879 a likelihood ratio test (Dobson, 2002).

880 For 79 trials it was not possible to measure an angle of head rotation before stimulus
881 onset due to experimenter’s mistake in the procedure. For one trial it was not possible to
882 measure the head rotation after the stimulus onset and whether the choice was correct or not
883 due to touchscreen malfunctioning. Hence, the sample size for both models described in this
884 section comprised 508 trials, 252 for counter clockwise-rotated stimuli and 256 for clockwise
885 rotations measured from 8 subjects (between 28 and 78 trials per subject) and 5 different
886 reinforced stimuli (between 28 and 78 trials per reinforced stimulus). In detail, for dogs 10 and
887 22, the model included 37 and 28 complete observations respectively. For dog 21, 68 complete
888 observations; for dog 07, 76 complete observations, for dog 37, 71 complete observations and
889 for dog 29, 72 complete observations. Dogs 07, 11 and 20 were tested on 2 additional sessions
890 (6 trials with rotated stimuli) relative to the other dogs. For dogs 11 and 20 all the 78
891 observations were available.

892 For both models, we assessed stability by excluding the levels of the random effects
893 one at a time (Nieuwenhuis et al., 2012) and calculated 95% confidence intervals conducting
894 a parametric bootstrap based on 1000 repetitions. Individual effects were tested using
895 likelihood ratio tests.

896

897 **References**

- 898 Baayen, R.H., 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*.
 899 Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511801686>
- 900 Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory
 901 hypothesis testing: Keep it maximal. *J. Mem. Lang.* 68, 255–278.
 902 <https://doi.org/10.1016/j.jml.2012.11.001>
- 903 Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B.,
 904 Grothendieck, G., Green, P., Bolker, M.B., 2015. Package ‘lme4.’ *Convergence* 12, 2.
- 905 Bauer, B., Jolicoeur, P., 1996. Stimulus dimensionality effects in mental rotation. *J. Exp. Psychol. Hum.*
 906 *Percept. Perform.* 22, 82–94. <https://doi.org/10.1037/0096-1523.22.1.82>
- 907 Bensky, M.K., Gosling, S.D., Sinn, D.L., 2013. Chapter Five - The World from a Dog’s Point of View: A
 908 Review and Synthesis of Dog Cognition Research, in: Brockmann, H.J., Roper, T.J., Naguib, M.,
 909 Mitani, J.C., Simmons, L.W., Barrett, L. (Eds.), *Advances in the Study of Behavior*. Academic
 910 Press, pp. 209–406. <https://doi.org/10.1016/B978-0-12-407186-5.00005-7>
- 911 Biederman, I., Cooper, E.E., 1991. Evidence for Complete Translational and Reflectional Invariance in
 912 Visual Object Priming. *Perception* 20, 585–593. <https://doi.org/10.1068/p200585>
- 913 Burmann, B., Dehnhardt, G., Mauck, B., 2005. Visual information processing in the lion-tailed
 914 macaque (*Macaca silenus*): mental rotation or rotational invariance? *Brain. Behav. Evol.* 65,
 915 168–176.
- 916 Byosiere, S.-E., Chouinard, P.A., Howell, T.J., Bennett, P.C., 2018. What do dogs (*Canis familiaris*) see?
 917 A review of vision in dogs and implications for cognition research. *Psychon. Bull. Rev.* 25,
 918 1798–1813. <https://doi.org/10.3758/s13423-017-1404-7>
- 919 Carlson, R.A., Avraamides, M.N., Cary, M., Strasberg, S., 2007. What do the hands externalize in
 920 simple arithmetic? *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 747–756.
 921 <https://doi.org/10.1037/0278-7393.33.4.747>
- 922 Chambers, J.M., Hastie, T.J., 1992. Linear models. Chapter 4 of statistical models in S. Wadsworth
 923 BrooksCole.
- 924 Chu, M., Kita, S., 2011. The nature of gestures’ beneficial role in spatial problem solving. *J. Exp.*
 925 *Psychol. Gen.* 140, 102–116. <https://doi.org/10.1037/a0021790>
- 926 Cooper, L.A., Shepard, R.N., 1973. Chronometric studies of the rotation of mental images, in: *Visual*
 927 *Information Processing*. Elsevier, pp. 75–176.
- 928 Delius, J.D., Hollard, V.D., 1995. Orientation invariant pattern recognition by pigeons (*Columba livia*)
 929 and humans (*Homo sapiens*). *J. Comp. Psychol.* 109, 278–290. <https://doi.org/10.1037/0735-7036.109.3.278>
- 930
- 931 Dobson, A.J., 2002. *An introduction to generalized linear models*. CRC press.
- 932 Edelman, S., Poggio, T., 1991. Models of object recognition. *Curr. Opin. Neurobiol.* 1, 270–273.
 933 [https://doi.org/10.1016/0959-4388\(91\)90089-P](https://doi.org/10.1016/0959-4388(91)90089-P)
- 934 Foster, D.H., 1978. Visual comparison of random-dot patterns: evidence concerning a fixed visual
 935 association between features and feature-relations. *Q. J. Exp. Psychol.* 30, 637–654.
- 936 Gauthier, I., Hayward, W.G., Tarr, M.J., Anderson, A.W., Skudlarski, P., Gore, J.C., 2002. BOLD Activity
 937 during Mental Rotation and Viewpoint-Dependent Object Recognition. *Neuron* 34, 161–171.
 938 [https://doi.org/10.1016/S0896-6273\(02\)00622-0](https://doi.org/10.1016/S0896-6273(02)00622-0)
- 939 Gilbert, S.J., 2015. Strategic use of reminders: Influence of both domain-general and task-specific
 940 metacognitive confidence, independent of objective memory ability. *Conscious. Cogn.* 33,
 941 245–260. <https://doi.org/10.1016/j.concog.2015.01.006>
- 942 Goldin-Meadow, S., Nusbaum, H., Kelly, S.D., Wagner, S., 2001. Explaining Math: Gesturing Lightens
 943 the Load. *Psychol. Sci.* 12, 516–522. <https://doi.org/10.1111/1467-9280.00395>
- 944 Hall, D.L., Friedman, A., 1994. Shape discriminations of three-dimensional objects depend on the
 945 number and location of bends. *Percept. Psychophys.* 56, 288–300.
 946 <https://doi.org/10.3758/BF03209763>

947 Hollard, V., Delius, J., 1982. Rotational invariance in visual pattern recognition by pigeons and
948 humans. *Science* 218, 804–806. <https://doi.org/10.1126/science.7134976>

949 Hopkins, W.D., Fagot, J., Vauclair, J., 1993. Mirror-image matching and mental rotation problem
950 solving by baboons (*Papio papio*): Unilateral input enhances performance. *J. Exp. Psychol.*
951 *Gen.* 122, 61.

952 Hsu, Y., Serpell, J.A., 2003. Development and validation of a questionnaire for measuring behavior
953 and temperament traits in pet dogs. *J. Am. Vet. Med. Assoc.* 223, 1293–1300.
954 <https://doi.org/10.2460/javma.2003.223.1293>

955 Janssens, L., Giemsch, L., Schmitz, R., Street, M., Van Dongen, S., Crombé, P., 2018. A new look at an
956 old dog: Bonn-Oberkassel reconsidered. *J. Archaeol. Sci.* 92, 126–138.

957 Köhler, C., Hoffmann, K.P., Dehnhardt, G., Mauck, B., 2005. Mental Rotation and Rotational
958 Invariance in the Rhesus Monkey (*Macaca mulatta*). *Brain. Behav. Evol.* 66, 158–166.
959 <https://doi.org/10.1159/000087156>

960 Kubinyi, E., Turcsán, B., Miklósi, Á., 2009. Dog and owner demographic characteristics and dog
961 personality trait associations. *Behav. Processes* 81, 392–401.
962 <https://doi.org/10.1016/j.beproc.2009.04.004>

963 Lamm, C., Windischberger, C., Moser, E., Bauer, H., 2007. The functional role of dorso-lateral
964 premotor cortex during mental rotation. *NeuroImage* 36, 1374–1386.
965 <https://doi.org/10.1016/j.neuroimage.2007.04.012>

966 Linn, M.C., Petersen, A.C., 1985. Emergence and characterization of sex differences in spatial ability:
967 A meta-analysis. *Child Dev.* 1479–1498.

968 Logothetis, N.K., Pauls, J., Poggio, T., 1995. Shape representation in the inferior temporal cortex of
969 monkeys. *Curr. Biol.* 5, 552–563. [https://doi.org/10.1016/S0960-9822\(95\)00108-4](https://doi.org/10.1016/S0960-9822(95)00108-4)

970 Matuschek, H., Kliegl, R., Vasisht, S., Baayen, H., Bates, D., 2017. Balancing Type I error and power in
971 linear mixed models. *J. Mem. Lang.* 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>

972 Mauck, B., Dehnhardt, G., 1997. Mental rotation in a California sea lion (*Zalophus californianus*). *J.*
973 *Exp. Biol.* 200, 1309–1316.

974 McCullagh, P., Nelder, J.A., 1989. Generalized linear models. Chapman & Hall. London.

975 Milgram, N.W., Head, E., Weiner, E., Thomas, E., 1994. Cognitive functions and aging in the dog:
976 Acquisition of nonspatial visual tasks. *Behav. Neurosci.* 108, 57–68.
977 <https://doi.org/10.1037/0735-7044.108.1.57>

978 Mongillo, P., Pitteri, E., Sambugaro, P., Carnier, P., Marinelli, L., 2017. Global bias reliability in dogs
979 (*Canis familiaris*). *Anim. Cogn.* 20, 257–265. <https://doi.org/10.1007/s10071-016-1044-8>

980 Müller, C.A., Mayer, C., Dörrenberg, S., Huber, L., Range, F., 2011. Female but not male dogs respond
981 to a size constancy violation. *Biol. Lett.* 7, 689–691.

982 Müller, C.A., Schmitt, K., Barber, A.L.A., Huber, L., 2015. Dogs Can Discriminate Emotional Expressions
983 of Human Faces. *Curr. Biol.* 25, 601–605. <https://doi.org/10.1016/j.cub.2014.12.055>

984 Navon, D., 1977. Forest before trees: The precedence of global features in visual perception. *Cognit.*
985 *Psychol.* 9, 353–383.

986 Nieuwenhuis, R., te Grotenhuis, M., Pelzer, B., 2012. influence. *ME: Tools for Detecting Influential*
987 *Data in Mixed Effects Models.*

988 Parsons, L.M., 1987. Imagined spatial transformations of one's hands and feet. *Cognit. Psychol.* 19,
989 178–241.

990 Pegado, F., Nakamura, K., Braga, L.W., Ventura, P., Nunes Filho, G., Pallier, C., Jobert, A., Morais, J.,
991 Cohen, L., Kolinsky, R., Dehaene, S., 2014. Literacy breaks mirror invariance for visual stimuli:
992 A behavioral study with adult illiterates. *J. Exp. Psychol. Gen.* 143, 887–894.
993 <https://doi.org/10.1037/a0033198>

994 Pitteri, E., Mongillo, P., Carnier, P., Marinelli, L., 2014a. Hierarchical stimulus processing by dogs
995 (*Canis familiaris*). *Anim. Cogn.* 17, 869–877. <https://doi.org/10.1007/s10071-013-0720-1>

996 Pitteri, E., Mongillo, P., Carnier, P., Marinelli, L., Huber, L., 2014b. Part-Based and Configural
997 Processing of Owner's Face in Dogs. *PLoS ONE* 9, e108176.
998 <https://doi.org/10.1371/journal.pone.0108176>

999 Queen, J.P., Keough, M.J., 2002. Experimental design and data analysis for biologists. Cambridge
1000 university press.

1001 Range, F., Aust, U., Steurer, M., Huber, L., 2008. Visual categorization of natural stimuli by domestic
1002 dogs. *Anim. Cogn.* 11, 339–347. <https://doi.org/10.1007/s10071-007-0123-2>

1003 Riesenhuber, M., Poggio, T., 2000. Models of object recognition. *Nat. Neurosci.* 3, 1199–1204.
1004 <https://doi.org/10.1038/81479>

1005 Risko, E.F., Gilbert, S.J., 2016. Cognitive Offloading. *Trends Cogn. Sci.* 20, 676–688.
1006 <https://doi.org/10.1016/j.tics.2016.07.002>

1007 Risko, E.F., Medimorec, S., Chisholm, J., Kingstone, A., 2014. Rotating With Rotated Text: A Natural
1008 Behavior Approach to Investigating Cognitive Offloading. *Cogn. Sci.* 38, 537–564.
1009 <https://doi.org/10.1111/cogs.12087>

1010 Schielzeth, H., Forstmeier, W., 2009. Conclusions beyond support: overconfident estimates in mixed
1011 models. *Behav. Ecol.* 20, 416–420. <https://doi.org/10.1093/beheco/arn145>

1012 Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis.
1013 *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>

1014 Serpell, J.A., Hsu, Y.A., 2005. Effects of breed, sex, and neuter status on trainability in dogs.
1015 *Anthrozoös* 18, 196–207. <https://doi.org/10.2752/089279305785594135>

1016 Shepard, R.N., Metzler, J., 1971. Mental rotation of three-dimensional objects. *Science* 171, 701–703.

1017 Sommese, A., Miklósi, Á., Pogány, Á., Temesi, A., Dror, S., Fugazza, C., 2021. An exploratory analysis
1018 of head-tilting in dogs. *Anim. Cogn.* <https://doi.org/10.1007/s10071-021-01571-8>

1019 Stich, K.P., Dehnhardt, G., Mauck, B., 2003. Mental rotation of perspective stimuli in a California sea
1020 lion (*Zalophus californianus*). *Brain. Behav. Evol.* 61, 102–112.

1021 Vanrie, J., Willems, B., Wagemans, J., 2001. Multiple Routes to Object Matching from Different
1022 Viewpoints: Mental Rotation versus Invariant Features. *Perception* 30, 1047–1056.
1023 <https://doi.org/10.1068/p3200>

1024 Wexler, M., Kosslyn, S.M., Berthoz, A., 1998. Motor processes in mental rotation. *Cognition* 68, 77–
1025 94. [https://doi.org/10.1016/S0010-0277\(98\)00032-8](https://doi.org/10.1016/S0010-0277(98)00032-8)

1026 Wohlschläger, A., 2001. Mental object rotation and the planning of hand movements. *Percept.*
1027 *Psychophys.* 63, 709–718.

1028 Wohlschläger, Andreas, Wohlschläger, Astrid, 1998. Mental and manual rotation. *J. Exp. Psychol.*
1029 *Hum. Percept. Perform.* 24, 397–412. <https://doi.org/10.1037/0096-1523.24.2.397>

1030 Zacks, 2008. Neuroimaging Studies of Mental Rotation: A Meta-analysis and Review | Journal of
1031 Cognitive Neuroscience | MIT Press Journals [WWW Document]. URL
1032 <https://www.mitpressjournals.org/doi/abs/10.1162/jocn.2008.20013> (accessed 5.7.20).
1033