

1 **Tool-use training temporarily enhances cognitive performance in long-tailed macaques**
2 **(*Macaca fascicularis*)**

3
4 Banty Tia^a, Riccardo Viaro^{a,b}, Luciano Fadiga^{a,b}

5
6 *^aCenter for Translational Neurophysiology of Speech and Communication, Istituto Italiano di*
7 *Tecnologia, 44121 Ferrara, Italy; ^bDepartment of Biomedical and Specialty Surgical Sciences,*
8 *Section of Human Physiology, University of Ferrara, 44121 Ferrara, Italy.*

9
10 **Corresponding author:**

11 Banty Tia, Ph.D.,
12 UMR 7179, MECADEV,
13 Centre National de la Recherche Scientifique,
14 Muséum National d'Histoire Naturelle,
15 1 avenue du Petit Château,
16 91800 Brunoy (France).
17 Phone: +33-1-60-47-92-03
18 Fax: +33-1-60-47-92-18
19 E-mail: banty.tia@mnhn.fr

20

21 **Abstract**

22 Tool use relies on numerous cognitive functions, including sustained attention and
23 understanding of causality. In this study, we investigated the effects of tool-use training on
24 cognitive performance in primates. Specifically, we applied the Primate Cognition Test Battery to
25 three long-tailed macaques (*Macaca fascicularis*) at different stages of a training procedure that
26 consisted of using a rake to retrieve out-of-reach food items. In addition, we evaluated a control
27 group (n=3) performing a grasping task, in order to account for possible effects related to a simple
28 motor act. Our results showed that tool-use training enhances mean performance in the physical
29 cognition domain, i.e. the understanding of spatial relations, numerosity and causality. In particular,
30 causal cognition (evaluating noise- and shape-related causality and understanding of tool properties)
31 showed significant improvement after training. Also, spatial cognition (evaluating spatial memory,
32 object permanence, rotation and transposition) showed a trend to improvement. Despite these
33 findings, none of our trained monkeys succeeded in the tool-use task of the Primate Cognition Test
34 Battery, which involved a differently shaped tool. Some training-related effects did not persist after
35 a 35-day resting period, suggesting that continuous practice may be necessary, or that a longer
36 training period before resting may be needed to better maintain cognitive performance. In contrast
37 with the training group, the control group did not display any change in cognitive performance. This
38 finding paves the way to further investigation into the link between tool-use behaviour and the
39 evolution of primate cognition.

40

41 **Keywords:** macaque, physical cognition, social cognition, tool use.

42

43 **Abbreviation:** PCTB, Primate Cognition Test Battery.

44

45 **Introduction**

46 A large body of work has examined species-specific physical and social cognition in
47 primates (Herrmann et al. 2007; Sabbatini et al. 2012; Schmitt et al. 2012). It has been argued that
48 Old World monkeys possess physical cognition abilities comparable to apes, with clear differences
49 emerging only in tasks of spatial understanding and tool use (Schmitt et al. 2012). However, the
50 development of physical and social cognition in a given species can be modulated by various
51 factors. For instance, animals raised in environments providing abundant opportunities for social
52 communication can develop better communicative abilities than in standard laboratory settings, as
53 shown in chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*; Russell et al. 2011).
54 Moreover, previous investigations support the idea that physical enrichment of a habitat promotes
55 the development of cognitive functions (Celli et al. 2003; Ventura and Buchanan-Smith 2003).

56 In the present study, we examined whether the acquisition of tool-use behaviour could affect
57 the cognitive functions in place. Previous meta-analyses of studies on primates indicate that several
58 cognitive skills may be involved in tool-use behaviour, including causal reasoning related to the
59 understanding of objects' physical properties and spatial relations (Seed and Byrne 2010). To
60 propagate tool use in a population, group members must also be able to rely on social transmission
61 mechanisms such as emulation or imitation (Van Schaik and Pradhan 2003). Furthermore, there is
62 evidence that tool-use learning is associated with the development of novel, complex abilities (Iriki
63 2006; Iriki and Sakura 2008). Tool use in primates has been found to promote their understanding
64 of physical interactions between objects, which can be generalised to different contexts (Fujita et al.
65 2011; Macellini et al. 2012). In other words, after learning to use a specific tool, monkeys can apply
66 their newfound ability to unfamiliar tools (Macellini et al. 2012), and more easily acquire
67 proficiency in sequential tool use (Hihara et al. 2003). In addition, previous work suggests that the
68 use of tools alters the representation of peripersonal space (the space within arm's reach), which is
69 associated with object localisation and movement execution, as well as extrapersonal space (the
70 space beyond arm's reach), which is associated with object identification (Gamberini et al. 2008;

71 Heber et al. 2010; Maravita and Iriki 2004). Tool-use acquisition also enhances the processing of
72 visual stimuli around the tool (Holmes and Calmels 2008), and, more generally, modifies
73 multisensory spatial attention (Holmes et al. 2007; Seed and Byrne 2010), which could affect a
74 wide range of tasks.

75 In order to exhaustively assess tool-related effects in the physical and social cognition
76 domains, we applied the Primate Cognition Test Battery (PCTB) to three long-tailed macaques
77 (*Macaca fascicularis*) at different stages of a tool-use training procedure. The PCTB was originally
78 designed by Herrmann et al. (2007) and adapted for testing on Old World monkeys by Schmitt et al.
79 (2012). It enables assessment of physical cognition skills related to the understanding of spatial,
80 numerical and causal relations between objects, and social cognition skills related to social learning,
81 communication ability and intention understanding. The tool-use task selected for the training
82 procedure has been used extensively in this field, and consists of using a rake to retrieve out-of-
83 reach food items (Maravita and Iriki 2004). As controls, we assessed a group of animals
84 accustomed to perform repeated food grasping tasks, in order to account for any effects related to
85 the performance of a simple motor task (Hillman et al. 2009).

86 Although previous work has evaluated the effects of tool-use training on specific functions
87 such as causal reasoning or spatial attention (Fujita et al. 2011; Hihara et al. 2003; Holmes et al.
88 2007; Macellini et al. 2012; Sabbatini et al. 2012), none have yet used a full test battery to explore
89 tasks in the physical and social domains. In particular, some skills, such as those related to
90 communication and theory of mind, have not previously been assessed in relation to tool-use
91 training. However, previous work supports the idea of a common network for tool use and
92 communicative gestures (Króliczak and Frey 2009; Steele et al. 2012), and therefore the
93 hypothetical effect of training on communication skills warrants investigation. In this work, we
94 performed periodic assessment throughout the training procedure in order to highlight any changes
95 in performance at different stages of the learning process. This paradigm was designed to test the
96 following hypotheses: 1) neither tool use nor repeated food grasping affect cognitive performance;

97 2) tool use enhances performance in specific skills (e.g. those related to causality, see also Fujita et
98 al. 2011; Macellini et al. 2012; Sabbatini et al. 2012); 3) both tool use and repeated grasping are
99 associated with better performance due to increased interaction with the experimenters and repeated
100 PCTB testing.

101

102 **Methods**

103

104 ***Subjects***

105 The data presented here was obtained from six long-tailed macaques (*Macaca fascicularis*;
106 two females and four males; 10.2 ± 2.0 years), housed at the Department of Biomedical and
107 Specialty Surgical Sciences, University of Ferrara, Italy. The macaques were individually housed in
108 cages (1.8 x 1.3 x 1.9 m). In order to maintain a high motivation for the task, macaques were mildly
109 food deprived, receiving part of their daily food at the end of each testing session. Water was
110 always available *ad libitum*. Experimental protocols were approved by the Animal Care Ethics
111 Committee of the University of Ferrara, authorised by the Italian Ministry of Health (research
112 permission n. 1139/2016-PR) and complied with the European laws on the use of laboratory
113 animals.

114

115 ***Procedure***

116 In this study, monkeys were evaluated using the PCTB at different stages of a tool-use
117 training program (training group) or grasping task (control group; Fig. 1). Macaques in the training
118 group (three monkeys; one female and two males; 12.3 ± 4.2 years) were trained to use a rake to
119 retrieve out-of-reach food items. Macaques in the control group (three monkeys; one female and
120 two males; 8.3 ± 1.8 years) had to grasp food items placed at different locations on a tray. The PCTB
121 was performed four times: (i) before the training or control procedure (baseline); (ii) after
122 completion of the first two training stages for the training group or after 25 grasping sessions for the

123 control group (1st period); (iii) after completion of the third and fourth training stages for the
124 training group or after 25 additional grasping sessions for the control group (2nd period); (iv) after a
125 35-day resting period post-training or post-grasping sessions (resting period). The PCTB was used
126 to assess each animal separately once it had achieved the completion criteria for the relevant
127 training stages. Figure 1b shows a schematic representation of the experimental procedure. All
128 datasets generated and/or analysed during the current study are available from the corresponding
129 author upon request.

130

131 *Tool-use training*

132 The macaques were trained to use a rake-shaped tool to retrieve out-of-reach food items
133 (raisins, peanuts, pieces of fruits and vegetables). All training sessions were conducted with the
134 monkeys in their individual home cage at the rate of 1 session/day, 5 days/week. Each session lasted
135 20–30 min. The food and tool were presented on a tray (length, 60 cm; width, 50 cm) attached to a
136 sliding table that could be moved horizontally from cage to cage. The rake tool was composed of a
137 wooden plate (16.5 x 11.5 x 0.5 cm) attached to a handle (diameter, 1.5 cm; length, 38 cm). The
138 tray was adjusted at cage ground level, so that the animals could see the food item from above,
139 thereby preventing the rake from obstructing their view of the food.

140 The step-by-step training protocol was similar to that previously described by Yamazaki et
141 al. (2011; Fig. 1a). In brief, the protocol comprised four main stages corresponding to different
142 locations of the food item relative to the tool. In stage a, the food item was placed close to the plate
143 of the rake so that the macaques were rewarded by simply pulling the tool. In stage b, the macaques
144 were required to move the tool to the left or right in addition to pulling. In stages c and d, monkeys
145 had to learn to move the tool forward from the original position in order to retrieve the food item.
146 Completion of a training stage was defined as five successive sessions in which the monkey
147 executed at least 80% successful trials.

148

149 <Figure 1>

150

151 ***Control grasping task***

152 Macaques in the control group were habituated to grasp food items placed at different
153 locations on a tray, using their preferred hand. Similar to the tool-use training procedure, all
154 sessions were conducted with the monkeys in their individual home cage, at the rate of 1
155 session/day, 5 days/week. Each session lasted 10–20 min. The same tray and food items were used
156 as for the tool-use training group, but, unlike the tool-use group, food items were placed within
157 reach of the animal, at six possible locations (anteroposterior: 10 cm or 20 cm away from the
158 animal; mediolateral: at the midpoint, and at 15 cm to the left or right of the midpoint). Forty trials
159 were performed in a single session. In order to maintain a high motivation for the task, food items
160 grasped during the session constituted part of the animals' daily food ration, with the macaques
161 receiving the remainder at the end of the session.

162

163 ***Primate Cognition Test Battery***

164 The PCTB was originally designed by Herrmann et al. (2007), and was adapted for testing
165 on Old World monkeys by Schmitt et al. (2012). It consists of different tasks designed to examine
166 physical and social cognition skills. It subdivides the physical domain into three scales, namely
167 space (4 tasks), quantities (2 tasks) and causality (4 tasks). The spatial cognition scale was designed
168 to assess the monkeys' ability to remember object locations and to infer the location of an object
169 after occluded lateral displacement or spatial rotation. It comprises four tasks: spatial memory (6
170 trials), object permanence (18 trials), rotation (18 trials) and transposition (18 trials). The quantities
171 scale evaluates monkeys' ability to distinguish between different amounts, and consists of two
172 tasks: relative numbers (16 trials) and additions (14 trials). Finally, the causality scale tests
173 monkeys' understanding of causal relations between objects. It includes four different tasks: noise
174 (12 trials), shape (12 trials), tool use (1 trial) and tool properties (30 trials). In particular, the tool-

175 use task assesses whether the animals are able to retrieve a food reward by using a wooden stick,
176 while the tool properties task examines, via five different sub-tasks (side, bridge, ripped cloth,
177 broken wool and tray circle), whether the animals can distinguish between a functional and a non-
178 functional tool based on the physical properties of the tool and the respective locations of the tool
179 and reward.

180 The social domain is also subdivided into three scales, namely social learning (1 task),
181 communication (3 tasks), and theory of mind (2 tasks). The social learning scale consists of testing
182 whether monkeys can learn to reproduce actions by observing an experimenter (3 trials). The
183 communication scale, on the other hand, tests their ability to understand and produce
184 communicative cues related to the experimenter. It comprises three tasks: comprehension (18 trials),
185 pointing cups (8 trials) and attentional state (4 trials). The theory of mind scale is designed to
186 evaluate monkeys' aptitude for acknowledging the mental states of the experimenter. It comprises
187 two tasks: gaze following (9 trials) and intentions (12 trials).

188 All testing procedures were similar to the ones previously published by Schmitt et al.
189 (2012). When giving a correct response, monkeys were rewarded with a small food item (raisins,
190 peanuts, pieces of fruits and vegetables) and were given a score of 1. A PCTB was generally
191 completed over 12 to 15 sessions. In order to ensure that the macaques' performance was
192 maintained throughout all PCTB sessions, a 10-min training/grasping period was provided at the
193 beginning of each PCTB session. All experiments were conducted with the monkeys in their
194 individual home cage (1.8 x 1.3 x 1.9 m) and were video-recorded with a digital video camera (QV-
195 IPC12B07, Qualvision Technology). A second observer independently scored 20% of all videotapes
196 to assess inter-observer agreement, which reached 94% for all the tasks combined.

197

198 *Data analysis*

199 The proportion of correct responses was calculated for each task (spatial memory, object
200 permanence, rotation, transposition, relative numbers, additions, noise, shape, tool use, tool

201 properties, comprehension, pointing cups, attentional state, gaze following, intentions). We assessed
202 whether individual performance exceeded chance level using Binomial tests (significance, $p < 0.05$;
203 see also Schmitt et al. 2012). Detailed explanation on the calculation of chance level was included
204 as Supplementary Material. Next, scores were averaged across tasks for each scale (space,
205 quantities, causality, social learning, communication, theory of mind) and across scales for each
206 domain (physical cognition, social cognition).

207 To explore whether tool use induced significant changes in cognitive performance, we
208 analysed scores calculated for each domain, scale and task using two-way repeated-measures
209 analysis of variance (RM-ANOVA), with GROUP (two levels: training, control) as between-
210 subjects factor, and TIME (baseline, 1st period, 2nd period, resting period) as within-subjects
211 factor. When a significant GROUP \times TIME interaction was detected, we conducted separate one-
212 way RM-ANOVAs for each group, with TIME as within-subjects factor. When a significant TIME
213 effect was found, within-subjects contrasts were investigated to compare each period with the
214 baseline value (significance, $p < 0.05$). Finally, using unpaired t-tests we compared our results with
215 those obtained by Schmitt et al. (2012), who evaluated a larger sample of untrained long-tailed
216 macaques ($n=13$).

217

218 **Results**

219

220 *Characterisation of tool-use learning*

221 In order to assess the learning progress across the different training stages, we plotted the
222 success rate over training sessions for each animal (Fig. 2). The plots showed that the three
223 monkeys in the training group needed similar time periods to achieve the success rate threshold in
224 stage a (range: 6–8 sessions; mean \pm SD: 7.0 \pm 1.0 sessions), but they differed from each other in
225 learning time for the subsequent training stages (range: 8–17, 5–19 and 6–14 sessions; mean \pm SD:
226 13.7 \pm 4.9, 10.3 \pm 7.6 and 10.0 \pm 4.0 sessions, respectively for stages b, c and d). Taken together, stages

227 a–b were performed successfully in 15–24 sessions (mean±SD: 20.7±4.9 sessions), and stages c–d
228 in 13–29 sessions (mean±SD: 20.3±8.1 sessions). Animals from the control group performed 25
229 grasping sessions during the period corresponding to training stages a–b, and another 25 sessions
230 during the period corresponding to stages c–d. Overall, animals in the training group performed a
231 total of 28–52 training sessions (mean±SD: 41.0±12.1 sessions), whereas animals in the control
232 group executed a total of 50 grasping sessions. In the training group, each animal executed 27.7±8.4
233 trials per session on average (mean±SD). In the control group, each animal executed a total of 40
234 successful grasping actions per session (data not shown).

235

236

<Figure 2>

237

238 *Effect of tool-use training on physical and social cognition domains*

239 To assess changes in cognitive performance at the domain level, we conducted separate
240 analyses for physical and social cognition scores. The two-way RM-ANOVA on physical cognition
241 values (Fig. 3a) confirmed a significant effect for GROUP ($F_{1,4}=25.97$, $p=0.0070$, partial $\eta^2=0.867$),
242 TIME ($F_{3,12}=4.30$, $p=0.0281$, partial $\eta^2=0.518$), and GROUP \times TIME interaction ($F_{3,12}=6.37$,
243 $p=0.0079$, partial $\eta^2=0.614$). Separate one-way RM-ANOVA conducted for the training group
244 showed a significant effect for TIME ($F_{3,6}=7.32$, $p=0.0198$, partial $\eta^2=0.786$). Tests of simple
245 contrasts showed significant differences from baseline at 1st period ($F_{1,2}=21.41$, $p=0.0437$, partial
246 $\eta^2=0.915$) and 2nd period ($F_{1,2}=88.42$, $p=0.0111$, partial $\eta^2=0.978$), but not resting period
247 ($F_{1,2}=8.12$, $p=0.1043$, partial $\eta^2=0.802$). On the opposite, the analysis conducted for the control
248 group showed a non-significant effect of TIME ($F_{3,6}=0.23$, $p=0.8725$, partial $\eta^2=0.103$). Overall,
249 these results indicate that the training group displayed improved performance from baseline to 1st
250 and 2nd periods, whereas the control group displayed no change in performance across periods.

251 Considering the social cognition domain, two-way RM-ANOVA showed no change in
252 performance (Fig. 3b). Non-significant effects were found for GROUP ($F_{1,4}=1.21$, $p=0.3330$, partial

253 $\eta^2=0.232$), TIME ($F_{3,12}=1.77$, $p=0.2055$, partial $\eta^2=0.307$) and GROUP \times TIME interaction
254 ($F_{3,12}=1.14$, $p=0.3717$, partial $\eta^2=0.222$).

255

256 <Figure 3>

257

258 *Effect of tool-use training on cognition scales*

259 To better assess the effects of tool-use training on cognition, we performed additional
260 analyses on the scores calculated for each scale. The two-way RM-ANOVA applied to the space
261 scale (Fig. 4a) showed a non-significant (albeit close to significance) effect for GROUP ($F_{1,4}=7.13$,
262 $p=0.0558$, partial $\eta^2=0.641$) and TIME ($F_{3,12}=2.00$, $p=0.1672$, partial $\eta^2=0.334$), but a significant
263 GROUP \times TIME interaction ($F_{3,12}=6.03$, $p=0.0096$, partial $\eta^2=0.601$). Separate one-way RM-
264 ANOVA conducted for the training group showed a significant effect for TIME ($F_{3,6}=4.76$,
265 $p=0.0500$, partial $\eta^2=0.704$). Simple contrasts comparing each period with the baseline showed a
266 trend for 1st period ($F_{1,2}=12.92$, $p=0.0695$, partial $\eta^2=0.866$), but non-significant differences for 2nd
267 period ($F_{1,2}=5.96$, $p=0.1347$, partial $\eta^2=0.749$) and resting period ($F_{1,2}=3.90$, $p=0.1871$, partial
268 $\eta^2=0.661$). The analysis conducted for the control group showed a non-significant effect of TIME
269 ($F_{3,6}=1.55$, $p=0.2955$, partial $\eta^2=0.437$). These results illustrate a trend to an increase in
270 performance for the tool-use training group, whereas the control group's performance remained
271 constant throughout periods.

272 The two-way RM-ANOVA applied to the quantities scale showed no change in performance
273 (Fig. 4b). Finally, the two-way RM-ANOVA conducted on the causality scale (Fig. 4c), showed a
274 non-significant effect for GROUP ($F_{1,4}=5.78$, $p=0.0741$, partial $\eta^2=0.591$), but a significant effect
275 for both TIME ($F_{3,12}=12.25$, $p=0.0006$, partial $\eta^2=0.754$) and GROUP \times TIME interaction
276 ($F_{3,12}=4.90$, $p=0.0190$, partial $\eta^2=0.550$). Separate one-way RM-ANOVA conducted for the training
277 group showed a significant effect of TIME ($F_{3,6}=11.00$, $p=0.0075$, partial $\eta^2=0.846$). Simple
278 contrasts showed significant differences from baseline at 1st period ($F_{1,2}=46.95$, $p=0.0206$, partial

279 $\eta^2=0.959$), 2nd period ($F_{1,2}=37.80$, $p=0.0255$, partial $\eta^2=0.950$), and resting period ($F_{1,2}=19.77$,
280 $p=0.0470$, partial $\eta^2=0.908$). On the contrary, the analysis conducted for the control group showed a
281 non-significant effect of TIME ($F_{3,6}=2.68$, $p=0.1401$, partial $\eta^2=0.573$). In other words, the tool-use
282 training group improved performance at 1st, 2nd and resting periods, while the control group
283 showed constant performance throughout PCTBs ($p<0.05$).

284 Regarding the social cognition scales (Fig. 4d–e), two-way RM-ANOVAs failed to detect
285 any effect on communication or theory of mind. As regards the communication scale, non-
286 significant effects were found for GROUP ($F_{1,4}=2.39$, $p=0.1970$, partial $\eta^2=0.374$), TIME
287 ($F_{3,12}=0.20$, $p=0.8964$, partial $\eta^2=0.047$), and GROUP \times TIME interaction ($F_{3,12}=2.04$, $p=0.1618$,
288 partial $\eta^2=0.338$). Likewise, for the theory of mind scale, non-significant effects were found for
289 GROUP ($F_{1,4}=0.52$, $p=0.5103$, partial $\eta^2=0.115$), TIME ($F_{3,12}=2.85$, $p=0.0820$, partial $\eta^2=0.416$),
290 and GROUP \times TIME interaction ($F_{3,12}=0.03$, $p=0.9937$, partial $\eta^2=0.007$). The social learning scale
291 was excluded from analysis, since all animals obtained null scores.

292

293 <Figure 4>

294

295 ***Effect of tool-use training on single tasks***

296 For a more precise assessment, we conducted separate two-way RM-ANOVAs on scores
297 obtained for each task in the PCTB. The tool-use task was excluded from the analysis, since all
298 animals obtained a null score at this exercise. Few single-task analyses yielded significant effects
299 (Tables 1 and 2).

300 Specifically, in the physical domain, two-way RM-ANOVA conducted on the transposition
301 task revealed only a trend for GROUP ($F_{1,4}=5.45$, $p=0.0799$, partial $\eta^2=0.576$), but significant
302 effects for TIME ($F_{3,12}=6.33$, $p=0.0081$, partial $\eta^2=0.613$) and GROUP \times TIME interaction
303 ($F_{3,12}=7.22$, $p=0.0050$, partial $\eta^2=0.644$). Separate one-way RM-ANOVA conducted for the training
304 group showed a significant effect of TIME ($F_{3,6}=13.67$, $p=0.0043$, partial $\eta^2=0.872$). Simple

305 contrasts showed significant differences from baseline at 1st period ($F_{1,2}=18.82$, $p=0.0492$, partial
306 $\eta^2=0.904$), 2nd period ($F_{1,2}=42.43$, $p=0.0228$, partial $\eta^2=0.955$), and resting period ($F_{1,2}=49.57$,
307 $p=0.0196$, partial $\eta^2=0.961$). On the contrary, the analysis conducted for the control group showed a
308 non-significant effect of TIME ($F_{3,6}=0.42$, $p=0.7453$, partial $\eta^2=0.174$). These findings illustrate
309 that the tool-use training group improved performance at 1st, 2nd and resting periods with respect to
310 the baseline, whereas the performance of the control group remained unchanged.

311 Furthermore, for the additions task, the ANOVA indicated a significant effect for GROUP
312 ($F_{1,4}=9.84$, $p=0.0349$, partial $\eta^2=0.711$), but a non-significant effect for TIME ($F_{3,12}=0.32$,
313 $p=0.8097$, partial $\eta^2=0.074$), and only a trend for GROUP \times TIME interaction ($F_{3,12}=2.99$,
314 $p=0.0732$, partial $\eta^2=0.428$). As regards the shape task, the same analysis revealed a trend for
315 GROUP ($F_{1,4}=4.79$, $p=0.0939$, partial $\eta^2=0.545$), and a significant effect for both TIME
316 ($F_{3,12}=10.02$, $p=0.0014$, partial $\eta^2=0.715$) and GROUP \times TIME interaction ($F_{3,12}=4.90$, $p=0.0190$,
317 partial $\eta^2=0.550$). Separate one-way RM-ANOVA conducted for the training group showed a
318 significant effect of TIME ($F_{3,6}=23.93$, $p=0.0010$, partial $\eta^2=0.923$). Simple contrasts showed
319 significant differences from baseline at 1st period ($F_{1,2}=120.59$, $p=0.0082$, partial $\eta^2=0.984$), 2nd
320 period ($F_{1,2}=47.81$, $p=0.0203$, partial $\eta^2=0.960$), and resting period ($F_{1,2}=26.57$, $p=0.0356$, partial
321 $\eta^2=0.930$). On the contrary, the analysis conducted for the control group showed a non-significant
322 effect of TIME ($F_{3,6}=1.44$, $p=0.3222$, partial $\eta^2=0.418$). These results indicate that the tool-use
323 group improved performance in the shape task at 1st, 2nd and resting periods compared to baseline;
324 once again, the performance of the control group did not change.

325 Finally, as regards the tool properties task, the ANOVA showed significant effects for
326 GROUP ($F_{1,4}=10.72$, $p=0.0307$, partial $\eta^2=0.728$) and TIME ($F_{3,12}=3.73$, $p=0.0421$, partial
327 $\eta^2=0.482$), but a non-significant GROUP \times TIME interaction ($F_{3,12}=2.75$, $p=0.0890$, partial
328 $\eta^2=0.407$). In an attempt to better characterise which aspect of tool understanding might have been
329 influenced by training, additional analyses were conducted on each of the tool properties sub-tasks.
330 All sub-tasks comprised a functional tool that could be used to retrieve a food reward, and a non-

331 functional one. The first two sub-tasks emphasised the role of the location of the tool with respect to
332 the reward. Specifically, in the side sub-task, the animal had to choose between pulling a piece of
333 cloth with a reward on it, or a piece of cloth placed next to a reward. In this sub-task, the ANOVA
334 showed a significant effect for GROUP ($F_{1,4}=46.24$, $p=0.0024$, partial $\eta^2=0.920$), but a non-
335 significant effect for TIME ($F_{3,12}=1.66$, $p=0.2292$, partial $\eta^2=0.293$) and GROUP \times TIME
336 interaction ($F_{3,12}=0.76$, $p=0.5375$, partial $\eta^2=0.160$). Similarly, in the bridge sub-task, in which the
337 reward was placed either on a bridge over the piece of cloth, or directly on the piece of cloth under
338 the bridge, the analysis showed a trend for GROUP ($F_{1,4}=6.00$, $p=0.0705$, partial $\eta^2=0.600$), but a
339 non-significant effect for TIME ($F_{3,12}=2.38$, $p=0.1207$, partial $\eta^2=0.373$) and GROUP \times TIME
340 interaction ($F_{3,12}=1.23$, $p=0.3408$, partial $\eta^2=0.236$).

341 The last three exercises in the tool properties task focused on the intrinsic properties of the
342 tool. In the ripped sub-task, the animal had to select between pulling an intact piece of cloth with a
343 reward on it, or a ripped cloth with similar reward. In this sub-task, the ANOVA showed a non-
344 significant effect for GROUP ($F_{1,4}=3.13$, $p=0.1518$, partial $\eta^2=0.439$), a significant effect for TIME
345 ($F_{3,12}=3.77$, $p=0.0406$, partial $\eta^2=0.485$) and a non-significant GROUP \times TIME interaction
346 ($F_{3,12}=2.06$, $p=0.1590$, partial $\eta^2=0.340$). The broken wool sub-task was based on the same principle
347 as the ripped cloth, but used a length of wool instead of a piece of cloth. However, in this sub-task,
348 the analysis showed a trend for GROUP ($F_{1,4}=7.23$, $p=0.0547$, partial $\eta^2=0.644$), but non-significant
349 effects for TIME ($F_{3,12}=0.43$, $p=0.7341$, partial $\eta^2=0.097$) and GROUP \times TIME interaction
350 ($F_{3,12}=1.40$, $p=0.2899$, partial $\eta^2=0.260$). In the tray circle sub-task, the monkey had to select
351 between a tray with a circular hole in it, or a tray with a U-shaped hole. Rewards were placed inside
352 both holes, but only the tray with the circular hole would allow to retrieve the reward. Here, the
353 ANOVA showed a non-significant effect for GROUP ($F_{1,4}=1.43$, $p=0.2985$, partial $\eta^2=0.263$), a
354 significant effect for TIME ($F_{3,12}=3.52$, $p=0.0491$, partial $\eta^2=0.468$) and a non-significant GROUP
355 \times TIME interaction ($F_{3,12}=1.45$, $p=0.2763$, partial $\eta^2=0.267$). Overall, none of the tool properties

356 sub-tasks disclosed a significant GROUP \times TIME interaction that could specify the effects of
357 training on tool-properties understanding.

358 Regarding the social domain, the two-way RM-ANOVA performed on the comprehension
359 task showed a non-significant effect for both GROUP ($F_{1,4}=3.86$, $p=0.1210$, partial $\eta^2=0.491$) and
360 TIME ($F_{3,12}=1.23$, $p=0.3433$, partial $\eta^2=0.234$), but a significant GROUP \times TIME interaction
361 ($F_{3,12}=3.69$, $p=0.0431$, partial $\eta^2=0.480$). Separate one-way RM-ANOVA conducted for the training
362 group showed a significant effect of TIME ($F_{3,6}=5.57$, $p=0.0361$, partial $\eta^2=0.736$). Simple contrasts
363 showed significant differences from baseline at 1st period ($F_{1,2}=25.00$, $p=0.0377$, partial $\eta^2=0.926$)
364 and resting period ($F_{1,2}=49.77$, $p=0.0195$, partial $\eta^2=0.961$), but not 2nd period ($F_{1,2}=4.96$,
365 $p=0.1557$, partial $\eta^2=0.713$). On the contrary, the analysis conducted for the control group showed a
366 non-significant effect of TIME ($F_{3,6}=1.62$, $p=0.2812$, partial $\eta^2=0.447$). In other words, the training
367 group improved scores at 1st period and resting period, while 2nd period did not reach significance.
368 The control group did not display any significant changes in comprehension performance across
369 periods. Finally, for the intentions task, the ANOVA did show a significant GROUP effect
370 ($F_{1,4}=8.68$, $p=0.0421$, partial $\eta^2=0.684$), but non-significant effects for TIME ($F_{3,12}=1.64$, $p=0.2317$,
371 partial $\eta^2=0.291$) and GROUP \times TIME interaction ($F_{3,12}=0.24$, $p=0.8685$, partial $\eta^2=0.056$).

372

373 *Single task performance with respect to chance level*

374 To further investigate performance in single tasks, we performed Binomial tests in order to
375 detect tasks in which individual performance exceeded the chance level. For the training group, the
376 number of individuals performing above chance increased from baseline to 1st period in the tasks:
377 rotation, transposition, relative numbers, additions, shape, tool properties, comprehension and
378 intentions, but decreased for the pointing cups task ($p<0.05$; Tables 1 and 2). Individual animals
379 performing above chance further increased from 1st to 2nd period in tasks: spatial memory, object
380 permanence, noise, shape, and pointing cups, but decreased in the comprehension task. From 2nd
381 period to resting period, the number of individuals performing above chance level increased only in

382 the relative numbers and comprehension tasks, whereas it decreased for spatial memory, additions,
383 noise, shape, tool properties, and pointing cups.

384

385 <Table 1>

386

387 In contrast, for the control group, only minor variations were found. Specifically, between
388 baseline and 1st period, the number of individuals performing above chance decreased for the
389 additions and pointing cups task ($p < 0.05$; Tables 1 and 2). Between 1st and 2nd periods, this
390 number increased for the tool properties task but decreased for object permanence, transposition and
391 comprehension tasks. Between 2nd and resting periods, the number of individuals performing above
392 chance increased for object permanence and comprehension, while it decreased for relative numbers
393 and tool properties tasks.

394

395 <Table 2>

396

397 *Comparison with previous work*

398 In order to identify possible differences related to study site, we conducted unpaired t-tests
399 comparing our sample at baseline with the sample described by Schmitt et al. (2012). The macaques
400 tested by Schmitt et al. (2012) lived in a large social group ($n=28$) and had access to both indoor
401 and outdoor areas, unlike our macaques ($n=6$) who were only housed indoors and had regular
402 interactions with the experimenters. Although these differences could have reasonably been
403 expected to affect performance in the physical and social domains, unpaired t-tests conducted on
404 space, quantities, communication and theory of mind scales showed no difference between the two
405 groups. That being said, our animals achieved significantly lower scores in causality than those
406 described by Schmitt et al. (mean \pm SD: 0.39 ± 0.04 as compared to 0.46 ± 0.05 ; $t_{17}=3.06$, $p=0.0071$).

407 Considering single tasks of the physical domain, our macaques displayed lower scores in
408 spatial memory (mean±SD: 0.47±0.07 as compared to 0.68±0.22; $t_{17}=2.24$, $p=0.0391$), rotation
409 (mean±SD: 0.36±0.08 as compared to 0.46±0.09; $t_{17}=2.27$, $p=0.0363$) and shape (mean±SD:
410 0.49±0.06 as compared to 0.66±0.17; $t_{17}=2.40$, $p=0.0281$) than those studied by Schmitt et al.
411 (2012). However, in the social domain, our animals obtained higher scores in the attentional state
412 task (mean±SD: 0.63±0.44 compared to 0.23±0.26; $t_{17}=2.47$, $p=0.0242$). Noteworthily, for the
413 majority of tasks, no difference was found between our subjects at baseline and those tested by
414 Schmitt et al. (2012).

415 In the second step of our comparative analysis, we conducted unpaired t-tests to compare the
416 performance of our subjects at 2nd period, when training was complete, with those of Schmitt et al.
417 (2012). This analysis, designed to identify skills that could be improved by tool use, revealed that,
418 at 2nd period, our training group animals displayed improved performance, surpassing the results
419 described by Schmitt et al. (2012), in scales of space (mean±SD: 0.64±0.09 as compared to
420 0.54±0.06; $t_{14}=2.30$, $p=0.0376$), quantities (mean±SD: 0.86±0.05 as compared to 0.67±0.08;
421 $t_{14}=3.82$, $p=0.0019$), causality (mean±SD: 0.57±0.10 as compared to 0.46±0.05; $t_{14}=3.00$, $p=0.0095$)
422 and communication (mean±SD: 0.80±0.17 as compared to 0.53±0.13; $t_{14}=3.12$, $p=0.0075$). In
423 contrast, our control group scores did not differ from those reported by Schmitt et al. (2012).

424 Considering single tasks of the physical domain, at 2nd period, the tool-use training group
425 displayed higher scores than those described by Schmitt et al. (2012) in terms of transposition
426 (mean±SD: 0.65±0.09 as compared to 0.39±0.07; $t_{14}=5.57$, $p<0.0001$), additions (mean±SD:
427 0.88±0.04 as compared to 0.64±0.11; $t_{14}=3.64$, $p=0.0030$) and tool properties (mean±SD: 0.88±0.13
428 as compared to 0.64±0.08; ($t_{14}=4.22$, $p=0.0009$). Similarly, in the social domain, our training group
429 animals achieved higher scores in the attentional state task (mean±SD: 0.83±0.29 as compared to
430 0.23±0.26; $t_{14}=3.56$, $p=0.0031$). Unlike the training group, control group scores did not differ from
431 those reported by Schmitt et al. (2012).

432

433 **Discussion**

434 In this study, we investigated the consequences of tool-use learning on physical and social
435 cognition skills in long-tailed macaques (*Macaca fascicularis*). The tool-use task consisted of using
436 a rake to retrieve out-of-reach food items at different locations. Performance on the PCTB was
437 assessed after pairs of training stages (a-b; c-d), rather than after each stage, in order to focus on the
438 most salient effects. Indeed, completion of a PCTB was achieved over 12–15 sessions, whereas
439 single training stages could reach completion criteria in as few as 5 sessions, which may have been
440 insufficient to detect great changes in performance.

441 Our results show that learning to use a tool enhances performance in two out of the three
442 measures of physical cognition, but zero of the three measures of social cognition. Causal cognition,
443 and to a lesser extent, spatial cognition, emerged as having particularly improved during training.
444 Despite these improvements, none of our monkeys succeeded in the tool-use task of the PCTB,
445 showing therefore an absence of generalisation to an unfamiliar tool. This finding justified more in-
446 depth investigation of single tasks of the PCTB to identify which aspects of physical/social
447 cognition were affected by training. Furthermore, some effects related to tool-use training did not
448 persist after a 35-day resting period, indicating that sustained practice may be necessary, or that a
449 longer period of training before resting may be needed to better maintain cognitive performance or
450 generalize to a different tool use. In contrast, the control task of repeated object grasping did not
451 affect performance on the PCTB.

452

453 ***Task acquisition***

454 Overall, the animals learnt to correctly perform the tool-use task of the training procedure
455 over 28–52 sessions. Albeit stage “a” of training was learned almost immediately, stages “b”, “c”
456 and “d” required a greater number of sessions for the macaques to become proficient. Indeed, these
457 latter stages required a more advanced understanding of spatial relations in order to adapt the
458 trajectory of the rake to the location of the reward (see also Yamazaki et al. 2011). Large inter-

459 individual differences were found in the time necessary to complete task acquisition (MK1: 52;
460 MK2: 43; MK3: 28 sessions), which echoes previous works documenting variability in manual
461 dexterity (long-tailed macaques, *Macaca fascicularis*; Kaeser et al. 2014) and tool-use training
462 (common marmosets, *Callithrix jacchus*; Yamazaki et al. 2011) in primates. Overall, it is important
463 to note that all monkeys eventually became successful in performing the tool-use task of the
464 training procedure, which consisted of using a rake to retrieve out-of-reach food items, despite
465 being unsuccessful at the PCTB tool task, which consisted of using a stick (differing from the
466 familiar rake shape) to retrieve food items.

467

468 ***Comparison with previous work***

469 Compared to a previous study evaluating long-tailed macaques in the PCTB (Schmitt et al.
470 2012), our animals at baseline achieved largely similar scores in quantities, social learning,
471 communication and theory of mind scales. However, they performed significantly worse in the
472 causal cognition scale, as well as in the spatial memory and rotation tasks from the spatial cognition
473 scale. In the study by Schmitt et al. (2012), animals had access to in- and outdoor areas, in contrast
474 with our animals, who were housed in laboratory conditions. This distinction could explain different
475 initial performance in the physical domain (Celli et al. 2003; Kozorovitskiy et al. 2005; Sanchez et
476 al. 1998; Ventura and Buchanan-Smith 2003). Despite lower initial performance, our animals, after
477 training, demonstrated improved performance in spatial, numerical and causal cognition, surpassing
478 the scores reported by Schmitt et al. (2012). Animals from the control group did not show
479 significant difference in performance throughout testing. This evidence strongly suggests that
480 improvements in physical cognition could be attributed to tool-use training.

481 Considering the social domain, at baseline our animals achieved significantly higher scores
482 in the attentional state task, in which the monkeys had to draw the attention of the experimenter
483 towards a reward. In general, the communication tasks of the PCTB tested interaction with a human
484 experimenter, rather than between monkeys. Our animals, raised in settings which frequently led

485 them to interact with experimenters, may therefore have developed better aptitudes to solve these
486 tasks (see also Russell et al. 2011). Interestingly, however, animals trained in tool use displayed
487 further improved performance on the communication scale, significantly exceeding that reported by
488 Schmitt et al. (2012). Therefore, it appears that training also affected some aspects of social
489 cognition, although to a minor extent than physical cognition.

490 Overall, the comparison of our results with those of Schmitt et al. (2012) unveiled lower
491 initial scores of our animals in some aspects of physical cognition, but higher initial scores in
492 aspects of social cognition, likely attributable to different raising environments. Moreover, in our
493 study, tool-use training strongly improved performance in the physical domain, while it more
494 weakly did so in the social domain. Altogether, these findings suggest that physical and social
495 cognition present some degree of independence and can be differently affected by raising conditions
496 or training procedures (Herrmann et al. 2007; Jack et al. 2013), although both domains could be
497 positively affected by training.

498

499 ***Tool-use training temporarily improves skills related to physical cognition***

500 Our results indicate that tool-use training induces significant improvement in cognitive tests
501 evaluating the macaques' comprehension of the physical world. The physical cognition domain
502 explored spatial memory and understanding of spatial relationships, ability to differentiate between
503 quantities, understanding of causality in the visual and auditory domains, and comprehension of
504 tool use (Herrmann et al. 2007; Schmitt et al. 2012). Training-related improvement was observed
505 mainly in the scales of causal cognition, and to a lesser extent, spatial cognition. Generally
506 speaking, significant effects were detected at both the scale and domain levels, while a few effects
507 arose in single tasks (transposition, shape). Of note, our tool-use group displayed higher scores in
508 all scales of the physical domain (space, quantities and causality) than those described by Schmitt et
509 al. (2012). Furthermore, the number of tool-use trained animals performing above chance level
510 increased in several tasks, from baseline to 1st period (rotation, transposition, relative numbers,

511 additions, shape and tool properties tasks), and from 1st to 2nd period (spatial memory, object
512 permanence, noise and shape tasks). As a whole, these findings suggest that, besides skills related to
513 spatial and causal cognition, tool-use training affected more general information-processing
514 functions (e.g. attentional), leading to enhanced performance in a variety of tasks of the physical
515 domain.

516 Nonetheless, it is worth noting that, despite evidence that monkeys can generalise tool-use
517 skills to unfamiliar tools (Macellini et al. 2012), none of our macaques succeeded in the PCTB tool
518 task, that consisted of using a stick to retrieve a food item. In fact, subjects rarely attempted to use
519 the stick as a tool and were more inclined to bite/play with the stick itself. A few attempts, albeit
520 unsuccessful, were however made by monkeys from the tool-use training group (2nd period: MK1;
521 resting period: MK1 and MK2). One possible explanation for this finding is that training was
522 insufficient to foster the macaques' ability to generalise their skill to different tasks/tools. Another
523 resides in the different properties of the stick and rake-shaped tool; unlike the rake used in our
524 training procedure, the stick lacked the frontal plate that facilitated retrieval of the reward.
525 Therefore, it may have been perceived as non-functional or not recognized as a tool (Macellini et al.
526 2012; Santos et al. 2003; Visalberghi et al. 2009)—a hypothesis compatible with a certain degree of
527 the macaques' understanding of a tool's functional properties. Indeed, statistical comparison of our
528 results with those reported by Schmitt et al. (2012) in a larger sample showed that tool-use trained
529 monkeys obtained higher scores in the understanding of tool properties immediately after training
530 (2nd period), supporting the idea that training promotes the understanding of task-relevant
531 properties of unfamiliar tools.

532 In this study we recorded performance enhancement in PCBT scores only after the first
533 stages of training, without further progress after greater familiarisation with the tool (i.e. from 1st to
534 2nd period). This suggests that training effects could have been triggered in the initial stage by
535 spatial/physical interactions between the tool and reward, whereas the subsequent stages only
536 served to maintain cognitive performance. Another possibility is that handling a new object

537 represented some enrichment sufficient to affect skills tested in the PCTB at the initial stages only
538 (see also Ventura and Buchanan-Smith 2003). Moreover, improvement in the average physical
539 cognition scores did not persist after a 35-day interruption (resting period), indicating that
540 continuous practice may be necessary, or that a longer training period before resting may be needed
541 to better maintain improvements in cognitive performance.

542 Overall, our findings are in line with previous works showing that tool-use training
543 promotes abilities related to spatial information processing (Gamberini et al. 2008; Maravita and
544 Iriki 2004) and understanding of physical interactions between objects (Fujita et al. 2011; Macellini
545 et al. 2012). Brain imaging experiments in Japanese macaques (*Macaca fuscata*) showed that 13–14
546 intensive tool-use training sessions (90 min/session, held over a 21-day period) induce an expansion
547 of grey matter in the temporal and parietal cortex (Quallo et al. 2009), suggesting that changes in
548 cortical structure may underlie performance enhancement following tool-use acquisition. In
549 humans, brain structure is known to be flexible enough to display changes in cortical grey matter
550 after only three months of self-exercising, which can recede after a similar period without training
551 (Boyke et al. 2008; Draganski et al. 2004). In long-tailed macaques, we found changes in
552 performance after relatively short time periods (e.g. one month of rest), which may be considered
553 insufficient for structural changes. Another possibility is that performance
554 enhancement/deterioration relies on training-induced reconfiguration of connectivity within brain
555 networks, as previously shown for humans trained in motor tasks (Taubert et al. 2011).

556

557 ***Tool-use training weakly affects skills related to social cognition***

558 In contrast with the results obtained in the physical domain, tool-use training seemed to have
559 a very small effect on skills in the social domain. Indeed, laboratory trained macaques (*Macaca*
560 *nemestrina* and *Macaca mulatta*) do not typically show signs of improvement in social cognition,
561 such as transmission of tool-use behaviour (Macellini et al. 2012). However, tool use and
562 communicative gestures are thought to depend on similar cortical networks (Króliczak and Frey

563 2009; Steele et al. 2012), which could suggest some influence of training on communication skills.
564 Furthermore, when comparing our data with that pertaining to Schmitt et al.'s (2012) larger sample,
565 statistical tests showed that tool-use trained monkeys obtained significantly higher scores in the
566 communication scale at the end of training (2nd period). When single tasks were considered, the
567 tool-use group displayed higher comprehension performance at 1st period and resting period, while
568 2nd period did not reach significance. The number of individuals performing above chance level in
569 the comprehension task increased from zero to three between baseline and 1st period for the
570 training group, while it remained unchanged for the control group. Together, these results suggest
571 that the training procedure did have a small influence on communication skills.

572 In fact, the lack of further significant effects of tool-use training on social cognition may be
573 due to insufficient task complexity and/or training duration. It should also be noted that as
574 mentioned, our macaques were raised and housed singly in a laboratory environment, which may
575 have affected their performance in socio-cognitive tasks (Russell et al. 2011; Sanchez et al. 1998).
576 However, as stated earlier, our animals obtained similar (social learning, theory of mind scales) or
577 higher (attentional state task) baseline scores than animals from a larger social group (28 animals;
578 Schmitt et al. 2012), which is likely due to the nature of the PCTB tasks, mainly evaluating
579 communication between monkeys and the human experimenter. Macaques raised in the laboratory
580 have frequent opportunities for animal-experimenter interaction, which may have contributed to
581 them developing these specific skills (see also Russell et al. 2011). That being said, a further
582 increase in performance was noted in our training group, which achieved higher scores in the
583 communication scale than described in Schmitt et al. (2012), whereas the performance of our
584 control group remained unchanged. This difference between groups suggests that tool-use training
585 strongly promoted interaction between monkeys and experimenters, leading to better performance
586 in communication tasks (i.e. comprehension and attentional state). Indeed, previous work suggests
587 that greater interaction with human caregivers during development (for complex communicative
588 activity, exploration, cognitive testing) can improve cognitive abilities in chimpanzees and bonobos

589 (Russell et al. 2011). Thus, enculturated apes can perform similarly to 2.5-year-old human children,
590 and outperform standard reared apes in the social domain (comprehension, production, attentional
591 state) as well as in the physical domain (rotation, relative numbers and noise tasks). Our results hint
592 that these findings in apes (Russell et al. 2011) could, to some degree, be applicable to macaques.
593 However, given the different time scales of the two studies (28–52 sessions of tool-use training vs.
594 9–30 years of socially enriched rearing environment), it is likely that increased human–macaque
595 interaction was not the only factor to promote social cognition in our animals, and that complex task
596 learning also played a role.

597 Although the PCTB did not include an evaluation of inter-individual communication, it is
598 worth noting that social interaction among monkeys could also have been influenced by training. In
599 macaques, competitive and cooperative social interactions are mediated by body postures, facial
600 expressions, gestures and vocalisations (Maestriperi 1997), which were not assessed in the present
601 study. Furthermore, there is evidence that affiliative and aggressive behaviours are not fixed in
602 time, but can be modified with procedures such as reinforcement training (Bloomsith et al. 1994;
603 Schapiro et al. 2001).

604 Finally, we recorded substantial inter-individual variability in social cognition tests (see
605 Tables 1 and 2). When considering each animal from the training group separately, all individuals
606 increased mean performance in the social domain, from baseline (MK1: 0.304; MK2: 0.503; MK3:
607 0.495) to 1st period (MK1: 0.340; MK2: 0.523; MK3: 0.613) and 2nd period (MK1: 0.367; MK2:
608 0.546; MK3: 0.647), albeit with considerable variation in scores among monkeys. Coherent with
609 our results, previous investigations have shown greater variability in social cognition (in particular,
610 social learning and communication) as compared to physical cognition skills (Herrmann et al. 2007;
611 Schmitt et al. 2012).

612

613 ***Conclusions***

614 In this study, we showed that tool use enhances performance in the physical cognition
615 domain, and more specifically in causal and spatial cognition, in adult long-tailed macaques.
616 Surprisingly, tool-use training did not improve performance in the tool-use task of the PCTB. These
617 results confirm that learning to use a tool promotes the comprehension of physical relations between
618 objects, despite possible absence of generalisation to an unfamiliar tool. Nonetheless, further
619 investigation is needed to shed more light on the apparent link between tool-use behaviour and the
620 evolution of primate cognition.

621

622 **Acknowledgements:** This work was supported by a grant from the Italian Ministry of Education,
623 University and Research (PRIN-2015, n. F52F16000840001).

624

625 **Conflict of Interest:** The authors declare that they have no conflict of interest.

626

627 **Ethical approval:** All applicable international, national, and/or institutional guidelines for the care
628 and use of animals were followed. All procedures were performed in accordance with the Animal
629 Care Ethics Committee of the University of Ferrara, authorised by the Italian Ministry of Health
630 (research permission n. 1139/2016-PR) and complied with the European laws on the use of
631 laboratory animals.

1 **References**

- 2 Bloomsmith MA, Laule G, Alford P, Thurston R (1994) Using training to moderate chimpanzee
3 aggression during feeding. *Zoo Biol* 13:557-566.
- 4 Boyke J, Driemeyer J, Gaser C, Buchel C, May A (2008) Training-induced brain structure changes
5 in the elderly. *J Neurosci* 28:7031-7035.
- 6 Celli ML, Tomonaga M, Udono T, Teramoto M, Nagano K (2003) Tool use task as environmental
7 enrichment for captive chimpanzees. *Appl Anim Behav Sci* 81:171-182.
- 8 Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, May A (2004) Neuroplasticity: changes
9 in grey matter induced by training. *Nature* 427:311-312.
- 10 Fujita K, Sato Y, Kuroshima H (2011) Learning and generalization of tool use by tufted capuchin
11 monkeys (*Cebus apella*) in tasks involving three factors: Reward, tool, and hindrance. *J Exp*
12 *Psychol Anim Behav Process* 37:10.
- 13 Gamberini L, Seraglia B, Priftis K (2008) Processing of peripersonal and extrapersonal space using
14 tools: evidence from visual line bisection in real and virtual environments.
15 *Neuropsychologia* 46:1298-1304.
- 16 Heber IA, Siebertz S, Wolter M, Kuhlen T, Fimm B (2010) Horizontal and vertical pseudoneglect
17 in peri- and extrapersonal space. *Brain Cogn* 73:160-166.
- 18 Herrmann E, Call J, Hernández-Lloreda MV, Hare B, Tomasello M (2007) Humans have evolved
19 specialized skills of social cognition: The cultural intelligence hypothesis. *Science*
20 317:1360-1366.
- 21 Hihara S, Obayashi S, Tanaka M, Iriki A (2003) Rapid learning of sequential tool use by macaque
22 monkeys. *Physiol Behav* 78:427-434.
- 23 Hillman CH, Pontifex MB, Raine LB, Castelli DM, Hall EE, Kramer AF (2009) The effect of acute
24 treadmill walking on cognitive control and academic achievement in preadolescent children.
25 *Neuroscience* 159:1044-1054.
- 26 Holmes NP, Sanabria D, Calvert GA, Spence C (2007) Tool-use: capturing multisensory spatial
27 attention or extending multisensory peripersonal space? *Cortex* 43:469-489.
- 28 Holmes P, Calmels C (2008) A neuroscientific review of imagery and observation use in sport. *J*
29 *Mot Behav* 40:433-445.
- 30 Iriki A (2006) The neural origins and implications of imitation, mirror neurons and tool use. *Curr*
31 *Opin Neurobiol* 16:660-667.
- 32 Iriki A, Sakura O (2008) The neuroscience of primate intellectual evolution: natural selection and
33 passive and intentional niche construction. *Philos Trans R Soc Lond B Biol Sci* 363:2229-
34 2241.

35 Jack AI, Dawson AJ, Begany KL, Leckie RL, Barry KP, Ciccia AH, Snyder AZ (2013) fMRI
36 reveals reciprocal inhibition between social and physical cognitive domains. *Neuroimage*
37 66:385-401.

38 Kaeser M et al. (2014) Variability of manual dexterity performance in non-human primates
39 (*Macaca fascicularis*). *Int J Comp Psychol* 27

40 Kozorovitskiy Y, Gross CG, Kopil C, Battaglia L, McBreen M, Stranahan AM, Gould E (2005)
41 Experience induces structural and biochemical changes in the adult primate brain. *Proc Natl*
42 *Acad Sci U S A* 102:17478-17482.

43 Króliczak G, Frey SH (2009) A common network in the left cerebral hemisphere represents
44 planning of tool use pantomimes and familiar intransitive gestures at the hand-independent
45 level. *Cereb Cortex*:bhn261.

46 Macellini S, Maranesi M, Bonini L, Simone L, Rozzi S, Ferrari PF, Fogassi L (2012) Individual and
47 social learning processes involved in the acquisition and generalization of tool use in
48 macaques. *Philos Trans R Soc Lond B Biol Sci* 367:24-36.

49 Maestripietri D (1997) Gestural communication in macaques: usage and meaning of nonvocal
50 signals. *Evol Comm* 1:193-222.

51 Maravita A, Iriki A (2004) Tools for the body (schema). *Trends Cogn Sci* 8:79-86.

52 Quallo MM, Price CJ, Ueno K, Asamizuya T, Cheng K, Lemon RN, Iriki A (2009) Gray and white
53 matter changes associated with tool-use learning in macaque monkeys. *Proc Natl Acad Sci*
54 *U S A* 106:18379-18384.

55 Russell JL, Lyn H, Schaeffer JA, Hopkins WD (2011) The role of socio - communicative rearing
56 environments in the development of social and physical cognition in apes. *Dev Sci* 14:1459-
57 1470.

58 Sabbatini G, Truppa V, Hribar A, Gambetta B, Call J, Visalberghi E (2012) Understanding the
59 functional properties of tools: chimpanzees (*Pan troglodytes*) and capuchin monkeys (*Cebus*
60 *apella*) attend to tool features differently. *Anim Cogn* 15:577-590.

61 Sanchez MM, Hearn EF, Do D, Rilling JK, Herndon JG (1998) Differential rearing affects corpus
62 callosum size and cognitive function of rhesus monkeys. *Brain Res* 812:38-49.

63 Santos LR, Miller CT, Hauser MD (2003) Representing tools: how two non-human primate species
64 distinguish between the functionally relevant and irrelevant features of a tool. *Anim Cogn*
65 6:269-281.

66 Schapiro SJ, Perlman JE, Boudreau BA (2001) Manipulating the affiliative interactions of group-
67 housed rhesus macaques using positive reinforcement training techniques. *Am J Primatol*
68 55:137-149.

69 Schmitt V, Pankau B, Fischer J (2012) Old world monkeys compare to apes in the primate
70 cognition test battery. PLoS One 7:e32024.

71 Seed A, Byrne R (2010) Animal tool-use. Curr Biol 20:R1032-R1039.

72 Steele J, Ferrari PF, Fogassi L (2012) From action to language: comparative perspectives on
73 primate tool use, gesture and the evolution of human language. Philos Trans R Soc Lond B
74 Biol Sci 367:4-9.

75 Taubert M, Lohmann G, Margulies DS, Villringer A, Ragert P (2011) Long-term effects of motor
76 training on resting-state networks and underlying brain structure. Neuroimage 57:1492-
77 1498.

78 Van Schaik CP, Pradhan GR (2003) A model for tool-use traditions in primates: implications for
79 the coevolution of culture and cognition. J Hum Evol 44:645-664.

80 Ventura R, Buchanan-Smith HM (2003) Physical environmental effects on infant care and
81 development in captive *Callithrix jacchus*. Int J Primatol 24:399-413.

82 Visalberghi E, Addessi E, Truppa V, Spagnoletti N, Ottoni E, Izar P, Fragaszy D (2009) Selection
83 of effective stone tools by wild bearded capuchin monkeys. Curr Biol 19:213-217.

84 Yamazaki Y, Echigo C, Saiki M, Inada M, Watanabe S, Iriki A (2011) Tool-use learning by
85 common marmosets (*Callithrix jacchus*). Exp Brain Res 213:63-71.

86

87

88

89 **Figure legends**

90

91 **Fig. 1** Experimental setup. **a** Rake-shaped tool composed of a 16.5 x 11.5 x 0.5 cm rectangular plate
92 attached to a 38 cm handle. The protocol consisted of four main stages, corresponding to different
93 locations of the food item relative to the tool. In stage a, the food item was placed on the side of the
94 plate facing the monkey. In stage b, the food item was placed to the left or right of the side facing
95 the monkey. In stage c, the food item was placed to the left or right, on the opposite side of the
96 plate. In stage d, the food item was placed on the side of the plate facing away from the monkey.
97 Distances between plate and food items are indicated in the Figure. **b** Time course of the
98 experimental procedure for control (upper panel) and training (lower panel) groups. Both groups
99 were tested before training/control procedures (PCTB-baseline), after completion of stages a–b for
100 the training group or 25 grasping sessions for the control group (PCTB-1st period), after completion
101 of stages c–d for the training group or 25 additional grasping sessions for the control group (PCTB-
102 2nd period), and after a 35-day resting period (PCTB-resting period).

103

104 **Fig. 2** Progress at tool-use training stages a, b, c and d for MK1, MK2 and MK3. Completion of a
105 stage was defined as five successive sessions in which the monkey executed at least 80% successful
106 trials. The horizontal dotted line indicates the 80% threshold.

107

108 **Fig. 3** Physical (**a**) and social (**b**) cognition scores obtained for control and training groups. All
109 values are expressed as mean±SD. * $p < 0.05$, different from baseline, as assessed by RM-ANOVAs
110 followed by within-subjects simple contrasts. For each domain, the dotted line represents mean
111 chance level calculated by averaging chance levels (or baseline values, where applicable) across
112 scales.

113

114 **Fig. 4** Scores obtained in space **(a)**, quantities **(b)**, causality **(c)**, communication **(d)** and theory of
115 mind **(e)** scales by control and training groups. For the social learning scale, no statistical analysis
116 was possible as performance was zero for all monkeys. All values are expressed as mean \pm SD. *
117 $p < 0.05$, different from baseline, as assessed by RM-ANOVAs followed by within-subjects simple
118 contrasts. For each scale, the dotted line represents mean chance level calculated by averaging
119 chance levels (or baseline values, where applicable) across tasks.

120

Tasks	Trials	n	Baseline			1st period			2nd period			Resting period			Chance
			M	(SD)	Ind	M	(SD)	Ind	M	(SD)	Ind	M	(SD)	Ind	
Physical			0.55	(0.02)		0.54	(0.02)		0.54	(0.02)		0.55	(0.05)		
			0.51	(0.05)		0.67	(0.04)		0.69	(0.05)		0.64	(0.06)		
Space			0.51	(0.05)		0.45	(0.06)		0.47	(0.04)		0.46	(0.08)		
			0.44	(0.09)		0.61	(0.04)		0.64	(0.09)		0.57	(0.02)		
Spatial memory	6	3	0.44	(0.10)		0.39	(0.10)		0.44	(0.10)		0.50	(0.17)		0.33
		3	0.50	(0.00)		0.61	(0.10)		0.67	(0.17)	<u>1</u>	0.44	(0.10)		
Object permanence	18	3	0.70	(0.13)	<u>2</u>	0.53	(0.05)	<u>2</u>	0.57	(0.11)	<u>1</u>	0.54	(0.20)	<u>2</u>	0.33
		3	0.56	(0.22)	<u>2</u>	0.69	(0.10)	<u>2</u>	0.73	(0.17)	<u>3</u>	0.70	(0.06)	<u>3</u>	
Rotation	18	3	0.41	(0.08)		0.37	(0.03)		0.39	(0.06)		0.37	(0.08)		0.33
		3	0.32	(0.06)		0.44	(0.20)	<u>1</u>	0.50	(0.22)	<u>1</u>	0.46	(0.12)	<u>1</u>	
Transposition	18	3	0.48	(0.06)	<u>1</u>	0.50	(0.10)	<u>1</u>	0.46	(0.03)		0.44	(0.06)		0.33
		3	0.41	(0.12)		0.69	(0.14)	<u>3</u>	0.65	(0.08)	<u>3</u>	0.67	(0.06)	<u>3</u>	
Quantities			0.76	(0.04)		0.75	(0.02)		0.72	(0.07)		0.74	(0.12)		
			0.70	(0.03)		0.86	(0.10)		0.86	(0.05)		0.83	(0.12)		
Relative numbers	16	3	0.79	(0.04)	<u>3</u>	0.81	(0.00)	<u>3</u>	0.81	(0.00)	<u>3</u>	0.69	(0.13)	<u>1</u>	0.50
		3	0.71	(0.10)	<u>1</u>	0.85	(0.16)	<u>2</u>	0.83	(0.07)	<u>2</u>	0.92	(0.10)	<u>3</u>	
Additions	14	3	0.74	(0.08)	<u>2</u>	0.69	(0.04)	<u>1</u>	0.62	(0.15)	<u>1</u>	0.79	(0.12)	<u>1</u>	0.50
		3	0.69	(0.08)	<u>1</u>	0.86	(0.07)	<u>3</u>	0.88	(0.04)	<u>3</u>	0.74	(0.15)	<u>2</u>	
Causality			0.38	(0.03)		0.41	(0.05)		0.42	(0.05)		0.44	(0.01)		
			0.40	(0.05)		0.54	(0.02)		0.57	(0.10)		0.53	(0.09)		
Noise	12	3	0.47	(0.05)		0.53	(0.05)		0.56	(0.05)		0.58	(0.14)		0.50
		3	0.53	(0.05)		0.58	(0.08)		0.58	(0.14)	<u>1</u>	0.61	(0.10)		
Shape	12	3	0.47	(0.05)		0.56	(0.10)		0.50	(0.17)		0.61	(0.10)		0.50
		3	0.50	(0.08)		0.81	(0.10)	<u>1</u>	0.83	(0.17)	<u>2</u>	0.75	(0.17)	<u>1</u>	
Tool use	1	3	0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		n.a.
		3	0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		
Tool properties	30	3	0.59	(0.02)		0.56	(0.05)		0.63	(0.09)	<u>1</u>	0.58	(0.05)		0.50
		3	0.56	(0.08)		0.79	(0.07)	<u>3</u>	0.88	(0.13)	<u>3</u>	0.74	(0.20)	<u>2</u>	
Side	5	3	0.50	(0.17)		0.61	(0.19)		0.61	(0.10)		0.50	(0.17)		0.50
		3	0.50	(0.17)		0.78	(0.19)	<u>1</u>	0.89	(0.10)	<u>1</u>	0.83	(0.29)	<u>2</u>	
Bridge	5	3	0.50	(0.00)		0.44	(0.10)		0.61	(0.35)	<u>1</u>	0.61	(0.25)		0.50
		3	0.44	(0.10)		0.78	(0.19)	<u>1</u>	0.89	(0.19)	<u>2</u>	0.78	(0.10)		
Ripped cloth	5	3	0.56	(0.10)		0.56	(0.10)		0.78	(0.10)		0.67	(0.00)		0.50
		3	0.56	(0.25)		0.94	(0.10)	<u>2</u>	0.89	(0.19)	<u>2</u>	0.72	(0.25)	<u>1</u>	
Broken wool	5	3	0.67	(0.17)		0.56	(0.10)		0.39	(0.19)		0.61	(0.10)		0.50
		3	0.67	(0.17)		0.61	(0.10)		0.78	(0.25)	<u>1</u>	0.72	(0.19)		
Tray circle	5	3	0.72	(0.10)		0.61	(0.10)		0.78	(0.25)	<u>1</u>	0.50	(0.17)		0.50
		3	0.61	(0.10)		0.83	(0.17)	<u>1</u>	0.94	(0.10)	<u>2</u>	0.67	(0.29)	<u>1</u>	

122

123

124 **Table 1.** Mean proportion of correct responses in each task and scale of the physical domain of the
125 PCTBs, for control and training groups. Training group values are highlighted in grey. Mean scores
126 for scales and domain were obtained by averaging scores of all animals at each scale and domain.
127 Significant differences from baseline are in bold (RM-ANOVAs followed by within-subjects simple

128 contrasts, $p < 0.05$). Trials: number of trials for each task; n: number of tested individuals; Ind:

129 number of individuals performing above chance level; n.a.: not applicable.

130

131

132

Tasks	Trials	n	Baseline			1st period			2nd period			Resting period			Chance
			M	(SD)	Ind	M	(SD)	Ind	M	(SD)	Ind	M	(SD)	Ind	
Social			0.37	(0.12)		0.38	(0.08)		0.38	(0.06)		0.41	(0.08)		
			0.43	(0.11)		0.49	(0.14)		0.52	(0.14)		0.48	(0.12)		
Social learning	3	3	0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		n.a.
			0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		0.00	(0.00)		
Communication			0.60	(0.21)		0.57	(0.15)		0.50	(0.13)		0.63	(0.13)		
			0.70	(0.21)		0.79	(0.11)		0.80	(0.17)		0.73	(0.10)		
Comprehension	18	3	0.65	(0.08)	<u>1</u>	0.59	(0.16)	<u>1</u>	0.46	(0.12)		0.63	(0.22)	<u>1</u>	0.50
			0.63	(0.03)		0.81	(0.08)	<u>3</u>	0.78	(0.15)	<u>2</u>	0.76	(0.06)	<u>3</u>	
Pointing cups	8	3	0.67	(0.19)	<u>1</u>	0.63	(0.13)		0.54	(0.07)		0.58	(0.07)		0.50
			0.71	(0.19)	<u>1</u>	0.71	(0.07)		0.79	(0.19)	<u>1</u>	0.67	(0.07)		
Attentional state	4	3	0.50	(0.50)		0.50	(0.43)		0.50	(0.25)		0.67	(0.14)		n.a.
			0.75	(0.43)		0.83	(0.29)		0.83	(0.29)		0.75	(0.25)		
Theory of mind			0.50	(0.19)		0.56	(0.08)		0.65	(0.08)		0.61	(0.12)		
			0.61	(0.13)		0.69	(0.30)		0.76	(0.28)		0.70	(0.26)		
Gaze following	9	3	0.59	(0.23)		0.63	(0.17)		0.78	(0.11)		0.74	(0.23)		n.a.
			0.63	(0.26)		0.74	(0.45)		0.74	(0.45)		0.74	(0.36)		
Intentions	12	3	0.42	(0.14)		0.50	(0.00)		0.53	(0.05)		0.47	(0.05)		0.50
			0.58	(0.08)		0.64	(0.24)	<u>1</u>	0.78	(0.13)	<u>1</u>	0.67	(0.17)	<u>1</u>	

134

135

136

137

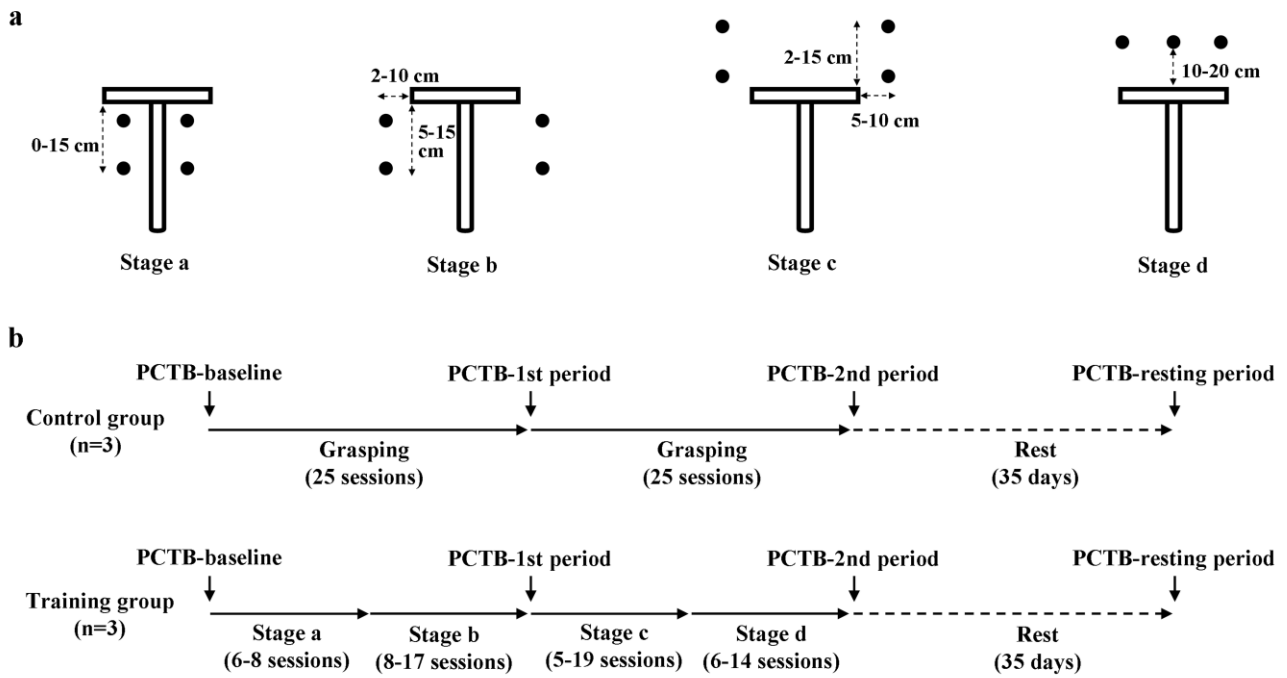
138

139

140

141

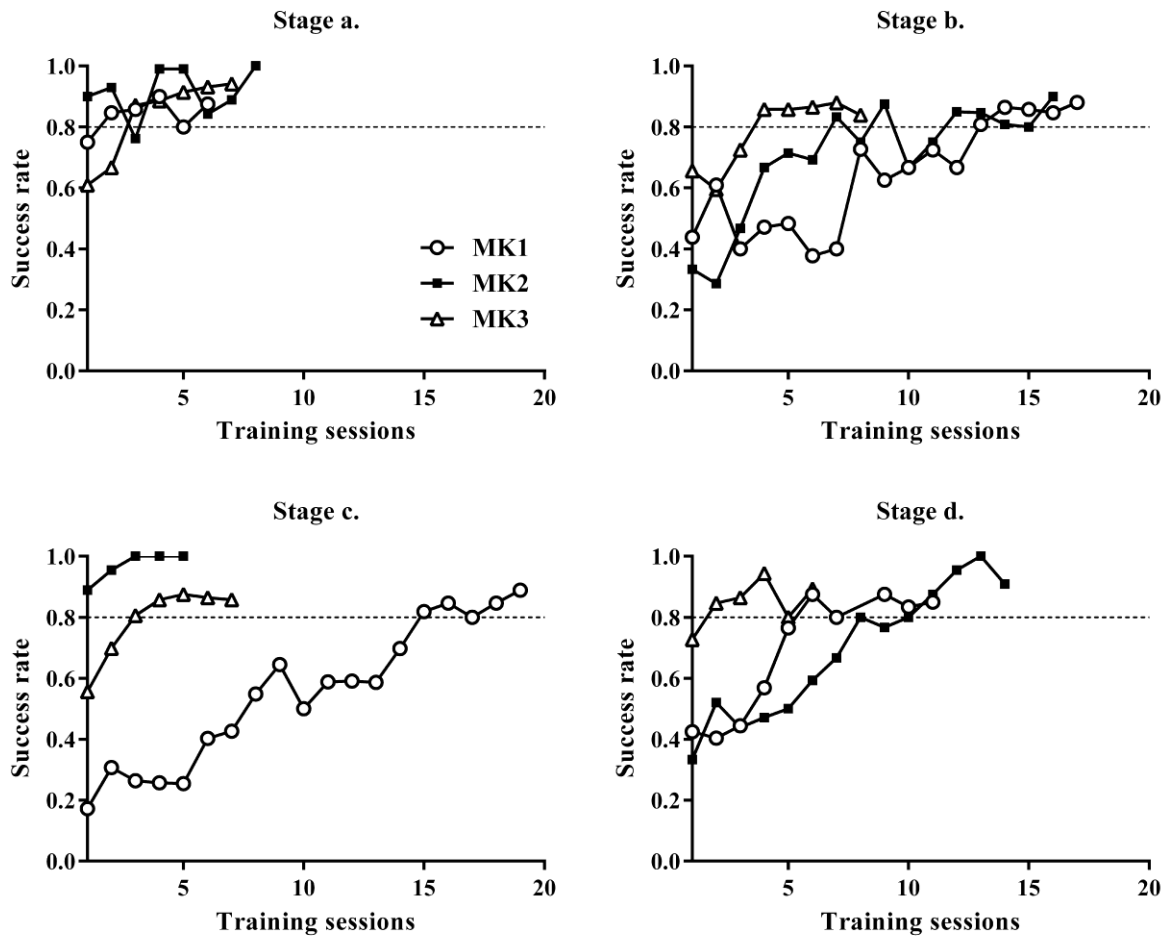
Table 2. Mean proportion of correct responses in each task and scale of the social domain of the PCTBs, for control and training groups. Training group values are highlighted in grey. Mean scores for scales and domain were obtained by averaging scores of all animals at each scale and domain. Significant differences from baseline are in bold (RM-ANOVAs followed by within-subjects simple contrasts, $p < 0.05$). Trials: number of trials for each task; n: number of tested individuals; Ind: number of individuals performing above chance level; n.a.: not applicable.



142

143

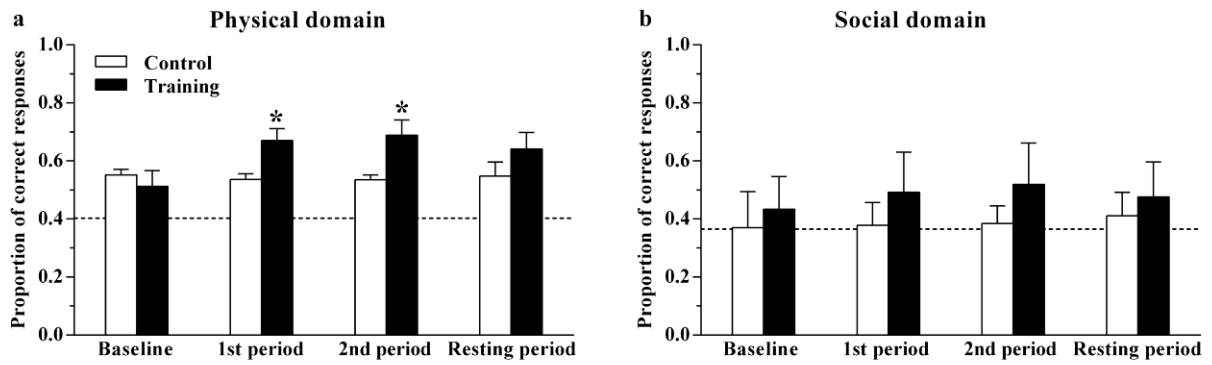
144 **Fig. 1** Experimental setup. **a** Rake-shaped tool composed of a 16.5 x 11.5 x 0.5 cm rectangular plate
 145 attached to a 38 cm handle. The protocol consisted of four main stages, corresponding to different
 146 locations of the food item relative to the tool. In stage a, the food item was placed on the side of the
 147 plate facing the monkey. In stage b, the food item was placed to the left or right of the side facing
 148 the monkey. In stage c, the food item was placed to the left or right, on the opposite side of the
 149 plate. In stage d, the food item was placed on the side of the plate facing away from the monkey.
 150 Distances between plate and food items are indicated in the Figure. **b** Time course of the
 151 experimental procedure for control (upper panel) and training (lower panel) groups. Both groups
 152 were tested before training/control procedures (PCTB-baseline), after completion of stages a–b for
 153 the training group or 25 grasping sessions for the control group (PCTB-1st period), after completion
 154 of stages c–d for the training group or 25 additional grasping sessions for the control group (PCTB-
 155 2nd period), and after a 35-day resting period (PCTB-resting period).



156

157 **Fig. 2** Progress at tool-use training stages a, b, c and d for MK1, MK2 and MK3. Completion of a
 158 stage was defined as five successive sessions in which the monkey executed at least 80% successful
 159 trials. The horizontal dotted line indicates the 80% threshold.

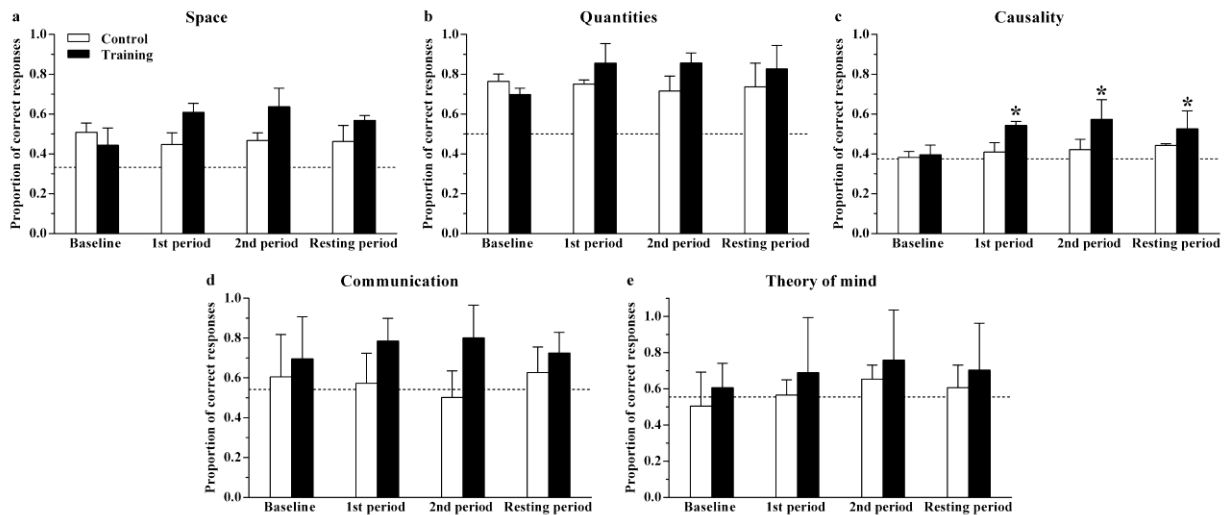
160



161

162 **Fig. 3** Physical (a) and social (b) cognition scores obtained for control and training groups. All
 163 values are expressed as mean±SD. * p<0.05, different from baseline, as assessed by RM-ANOVAs
 164 followed by within-subjects simple contrasts. For each domain, the dotted line represents mean
 165 chance level calculated by averaging chance levels (or baseline values, where applicable) across
 166 scales.

167



168

169 **Fig. 4** Scores obtained in space (a), quantities (b), causality (c), communication (d) and theory of
 170 mind (e) scales by control and training groups. For the social learning scale, no statistical analysis
 171 was possible as performance was zero for all monkeys. All values are expressed as mean±SD. *
 172 $p < 0.05$, different from baseline, as assessed by RM-ANOVAs followed by within-subjects simple
 173 contrasts. For each scale, the dotted line represents mean chance level calculated by averaging
 174 chance levels (or baseline values, where applicable) across tasks.