

1 **Understanding imitation in *Papio papio*: the role of experience and the presence of a conspecific**
 2 **demonstrator**

3 **Abstract**

4 What factors affect imitation performance? Varying theories of imitation stress the role of experience,
 5 but few studies have explicitly tested its role in imitative learning in non-human primates. We tested
 6 several predictions regarding the role of experience, conspecific presence, and action compatibility
 7 using a stimulus-response compatibility protocol. Nineteen baboons separated into two experimental
 8 groups learned to respond by targeting on a touch screen the same stimulus as their neighbour
 9 (Compatible) or the opposite stimulus (Incompatible). They first performed the task with a conspecific
 10 demonstrator (Social phase) and then a computer demonstrator (Ghost phase). After reaching a
 11 predetermined success threshold, they were then tested in an opposite compatibility condition (i.e.,
 12 Reversal learning conditions). Seven baboons performed at least two reversals during the social phase,
 13 and we found no significant difference between the compatible and incompatible conditions, although
 14 we noticed slightly faster RTs in the compatible condition that disappeared after the first reversal.
 15 During the ghost phase, monkeys showed difficulties in learning the incompatible condition, and the
 16 compatible condition RTs tended to be slower than during the social phase. Together, these results
 17 suggest that 1) there is no strong movement compatibility effect in our task, and that 2) the presence
 18 of a demonstrator plays a role in eliciting correct responses but is not essential, as has been previously
 19 shown in human studies.

20 *Keywords*: automatic imitation, social learning, stimulus-response compatibility, mirror neuron,
 21 guinea baboon, ghost demonstrator

22 *Word count*: 11368 approx.

23 **Understanding imitation in *Papio papio*: the role of experience and the presence of a conspecific**
24 **demonstrator**

25 Imitation has been defined in various ways in the comparative and human literature (for a review see
26 O’Sullivan & Caldwell, 2018), with researchers emphasizing to various extents the importance of the
27 novelty of actions copied or an understanding of goals or intentions (Byrne, 2002; Subiaul, 2010).
28 However, in purely behavioral terms, imitation has been described simply as the matching of
29 topographical features of another’s behavior (O’Sullivan & Caldwell, 2018). To imitate someone’s
30 movements (for example, when someone sticks out their tongue), one must match the movements
31 being observed to one’s own body movements, often having little or no visual access to one’s own
32 actions. This is commonly referred to as the "correspondence problem" (Nehaniv et al., 2002) and a
33 number of theories have been proposed to explain how humans and animals might solve this
34 “problem”. Multi-level theories of imitation propose the capacity to imitate is multi-faceted, achieved
35 through different mechanisms depending on whether meaningful and meaningless actions are
36 performed (Rumiati & Tessari, 2002), or depending on the content to be imitated (Subiaul, 2016).
37 Other approaches emphasize a domain-general view of imitation, where the copying of actions can be
38 facilitated through the development of associations between motor actions and sensory information
39 (Brass & Heyes, 2005; Prinz, 1997; Ray & Heyes, 2011).

40

41 Issues of domain-general versus domain specific imitative mechanisms are often reflected in the
42 disagreements over the ontogeny of imitative behaviour. A strong nativist view assumes that the
43 correspondence problem is solved through an innate specialized mechanism for imitation (Meltzoff &
44 Moore, 1997); a view supported by evidence that newborns can imitate facial expressions (Meltzoff &
45 Moore, 1977). However, evidence of neonatal imitation has come under growing scrutiny over the last
46 few years (Oostenbroek et al., 2016; although see Meltzoff et al., 2017). Proponents of domain general
47 accounts of imitation have examined multiple avenues of research to support the notion that imitation

48 is largely driven by experience. Heyes & Ray (2000) have proposed the Associative Sequence Learning
49 (ASL) model, which suggests that the cortical connections mediating the observation of the action and
50 the motor activation (or vertical associations) are formed by correlated experiences of observing and
51 performing the same action. In other words, these cortical connections would be formed during
52 associative learning. An individual forms a sensorimotor link for a given action X when the observation
53 and execution of this action X occur in a congruent or correlated manner, i.e., when the sensory
54 representation of X is active at the same time as the motor representation of X, and not that of another
55 motor action Y. While both nativist and experiential theorists acknowledge the role of genetics and
56 experience in imitative behavior, they propose different pre-existing mechanisms. For example,
57 Meltzoff and Moore's (1997) Active Intermodal Mapping (AIM) theory describes an innate mechanism
58 specifically evolved to match observed movement to specific organs. This supramodal representation
59 system compares perceived actions to proprioceptive feedback from one's own actions to facilitate
60 imitation. Overall, key differences between AIM and ASL accounts, include the form of the mechanism
61 used to facilitate imitation (i.e., an amodal representation system versus a modal sensorimotor
62 system), and the origin of these systems (i.e., supramodal representation systems are thought to exist
63 from birth, while sensorimotor associations develop through experience). Further, while AIM proposes
64 a domain-specific system adapted for imitation, experiential approaches place domain general
65 learning mechanisms at the root of imitative behavior that may be scaffolded by species-specific
66 motivational and emotional predispositions (Heyes, 2018).

67

68 Studies of automatic imitation have shed light on the mechanisms involved in imitative behavior.
69 Automatic imitation is a kind of Stimulus-Response Compatibility (SCR) effect in which the
70 characteristics of a behavioral performance (speed, accuracy) are mediated by the congruence
71 between a stimulus and an individual's response (Eimer, 1995; Heyes, 2011): it is easier for an
72 individual to perform the same action as a demonstrator than the other way around, even if this may

73 compromise the completion of a task. This type of SRC protocol is commonly used in imitation studies
74 because it is thought to reflect an imitative mechanism. It also allows the isolation of certain processes
75 involved in the transferal of sensory information into a motor action, while minimizing the demands
76 on executive functions. A classic procedure with humans involves a cueing or imperative stimulus (i.e.
77 a green dot for an open hand; red dot for a closed hand) while simultaneously presenting distractor
78 images that are compatible or incompatible with the action to be performed. Using this framework,
79 Stürmer et al. (2000) showed that adult participants were faster when they had to respond to the color
80 stimulus when the same movement was simultaneously presented (compatible condition) than when
81 the opposite movement was present (incompatible condition). This difference in response time (RT)
82 suggests that observing a stimulus facilitates a compatible response and that one must inhibit this
83 response under incompatible conditions resulting in longer RTs.

84 Stimulus-response studies involving training or other types of experience have contributed to support
85 for domain general accounts of imitation. For example, recent studies have shown that children who
86 suffered from bilateral cataract at birth or a few months later, and therefore lacked the experience
87 necessary for the potential formation of sensorimotor connections, showed an impaired automatic
88 imitation effect in a SRC task a few months after regaining vision (McKyton et al., 2018). Further, the
89 ASL model predicts that training in situations involving incongruent sensory and motor representations
90 should cancel or even reverse this automatic imitation effect. In other words, if an individual is taught
91 to respond to an "open hand" stimulus with a closed hand, there should be a new connection between
92 an "open hand" visual representation and a "closed hand" motor representation. Incompatible training
93 suggests the emergence of "counter-mirror" neurons that may inhibit the bidirectional link between
94 open-handed sensory and open-handed motor representations, thereby reducing the bias of RT in
95 compatible trials (e.g., Catmur et al. 2008). In a study with adults, incompatible training (in which
96 participants had to open their hands in response to a closed hand stimulus and vice versa), abolished
97 the automatic imitation effect previously observed (Heyes et al., 2005). Twenty-four hours after this
98 training, responses to incompatible trials were as fast as responding to compatible trials for subjects

99 with incompatible training compared to subjects with compatible training. Gillmeister et al. (2008)
100 reported a similar abolition of the automatic imitation effect on foot or hand action following
101 incompatible training.

102 Stimulus-response compatibility paradigms have also been used to study imitation across a range of
103 non-human species. Mui et al. (2008) found that budgerigars that had to "counter" imitate to get a
104 reward, i.e., peck after seeing a stepping stimulus (and vice versa), made significantly fewer correct
105 responses than budgerigars that had to imitate to receive food, i.e., peck after seeing a pecking
106 stimulus (and vice versa). Range et al. (2011) found the same automatic effect but also showed that
107 dogs from the incompatible group (those rewarded for responding with muzzle use after observing
108 paw use and paw use after observing muzzle use) tended to make more errors when presented with
109 compatible trials (i.e., rewarded for using the same body part as a demonstrator in a trial) than dogs
110 from the compatible group. This suggested that dogs in the incompatible group developed a tendency
111 for "automatic counter-imitation", consistent with the sensory-motor models. Finally, although studies
112 involving non-human primates are particularly important because of their phylogenetic proximity to
113 humans, to our knowledge only one study has examined automatic imitation in an SRC task with non-
114 human primates, showing a stimulus-response compatibility effect in Capuchin monkeys (O'Sullivan et
115 al., 2017). Moreover, after counter-imitative training, the authors showed that sensorimotor learning
116 could eliminate the imitative bias, in line with what had already been found in previous experiments
117 with humans and dogs. However, the results of this study remain difficult to interpret, due to the small
118 number of individuals studied.

119 In general, experience dependent accounts of imitation are underexplored in primate literature, with
120 many studies focusing on identifying imitation at single timepoints to draw conclusions regarding the
121 presence/absence of this ability in a particular species (e.g., van de Waal & Whiten, 2012; Voelkl &
122 Huber, 2000). As described above, in humans, studies have used stimulus-response compatibility
123 paradigms to demonstrate the role of experience in imitative responses across a range of effectors and

124 context (e.g., Catmur et al., 2011; Press et al., 2007). Another method of examining domain general
125 versus domain specific accounts is to compare social and asocial learning on the same task. For
126 example, Blandin and Proteau (2000) found that experience obtained through social observation
127 produced similar results to conditions where participants were able to practice the task individually,
128 suggesting the mechanisms underlying imitative behavior were similar to those used in individual
129 learning. In the non-human primate literature, ghost controls are used to assess social learning
130 performance in comparison to asocial learning (i.e., affordance learning, goal emulation; (Hopper,
131 2010). In ghost-control studies, the social information that would normally be available in a social
132 learning trial (i.e., an animal moving a particular object) is instead performed by a non-social agent
133 (i.e., hidden string moving an object). In touchscreen tasks with rhesus macaques (*Macaca mulatta*),
134 ghost controls have been used to demonstrate that in some contexts the presentation of social
135 information by a conspecific lead to improved learning of a task, when compared to conditions where
136 the same information is presented automatically in some other asocial ways (Subiaul et al., 2004). In a
137 more recent study with capuchin monkeys (*Sapajus apella*), however, no difference was identified
138 between social and asocial condition in a touchscreen task (Renner et al., 2021). The impact of social
139 information in touchscreen tasks with non-human primates, then, is currently unclear.

140 The objective of the present experiment was to explore imitative biases in another non-human primate
141 species, Guinea baboons (*Papio papio*). To our knowledge, there is limited evidence of imitative
142 capacity in this species, and so one of our main objectives was to determine whether there is an
143 imitative bias in this species using a touch screen social learning task. Multiple forms of imitation have
144 been described in the literature (see Subiaul, 2016) but here our task overlaps in terms of goal, action,
145 and cognitive demands, and so we can only draw conclusions regarding general social learning
146 mechanisms (rather than specific forms of imitation – e.g., see Subiaul, 2010). Second, the role of
147 experience in imitation in non-human animals is under-explored and here we hope to further our
148 understanding of the impact of experience by examining both imitative and counter-imitative
149 behavior. To meet this goal we used a novel, automated touchscreen task to examine learning of both

150 rule types across thousands of trials. This method allows us to examine change in stimulus-response
151 learning across time, examining learning across both compatible and incompatible conditions. Finally,
152 we wish to examine whether the presence of another individual has an impact on imitative learning.
153 To address this goal we used a control condition in which there was no social component included in
154 the demonstration (i.e., the demonstrator of the action is replaced by a “Ghost” agent (Hopper, 2010);
155 see Method). Even though imitation is a form of social learning, and therefore necessarily involves the
156 presence of another individual, studies of imitation using stimulus-response compatibility in animals
157 have not attempted to show the importance of the presence of others. Even domain-general accounts
158 of imitative learning stress that imitative learning must be built upon a species’ capacity to form
159 specific associations between stimuli and responses which may be scaffolded by motivational biases
160 to attend to social stimuli. It may be possible to demonstrate that SRC effects could be observed in the
161 absence of another individual, which would strongly question the value of the SRC protocol for the
162 study of a special imitative capacity (Ferruci et al., 2019; Nougaret et al., 2019).

163

164

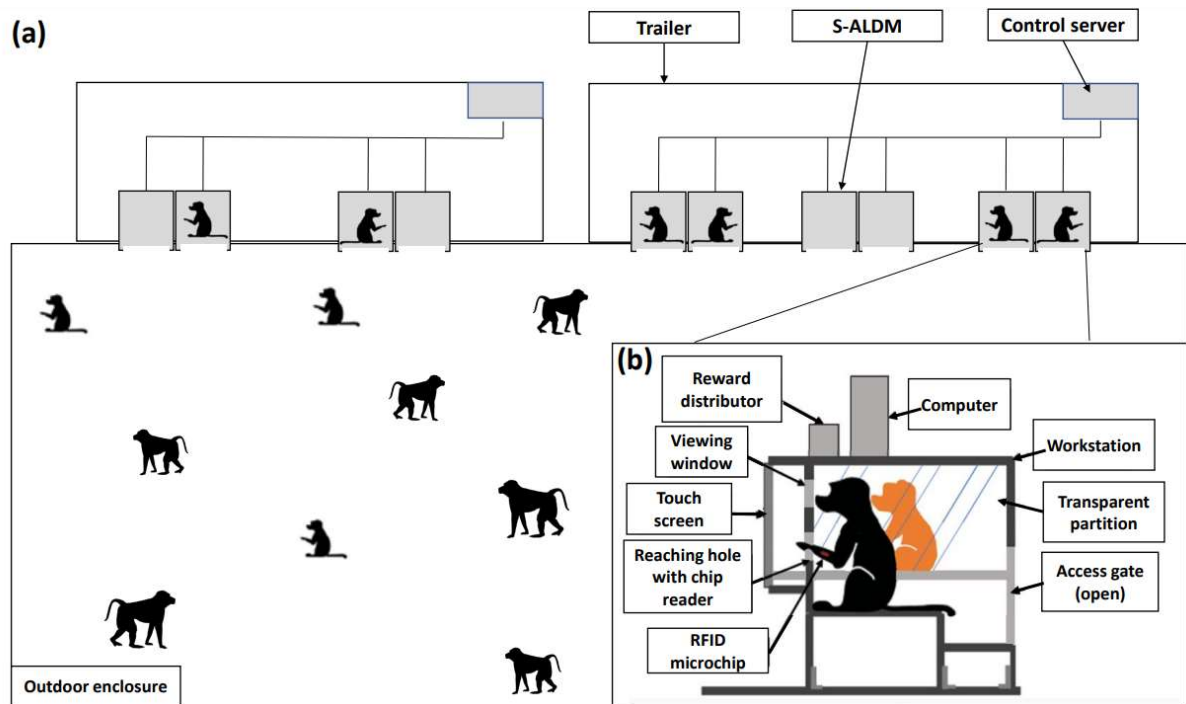
Method

165 Participants

166 Six males and 13 females Guinea baboons (*Papio papio*) were tested in this study (mean age: 11.5 years
167 [min=2.5; max=24.6]). They belonged to a social group of 19 individuals living in a 25x30m outdoor
168 enclosure connected to a 6x4m indoor enclosure and two 8x4m trailers (Fig.1a). The monkeys had *ad*
169 *libitum* access to ten Automated Computer Learning Devices for Monkeys (ALDMs: Fagot &
170 Paleressompoulle, 2009; Bonté et al, 2011), four in the first trailer and six in the second. The ALDM
171 used an automatic radio frequency identification device (RFID) implanted in each forearm of the
172 monkeys to recognize the individuals. This device makes it possible to test the animals without having
173 to capture and isolate them.

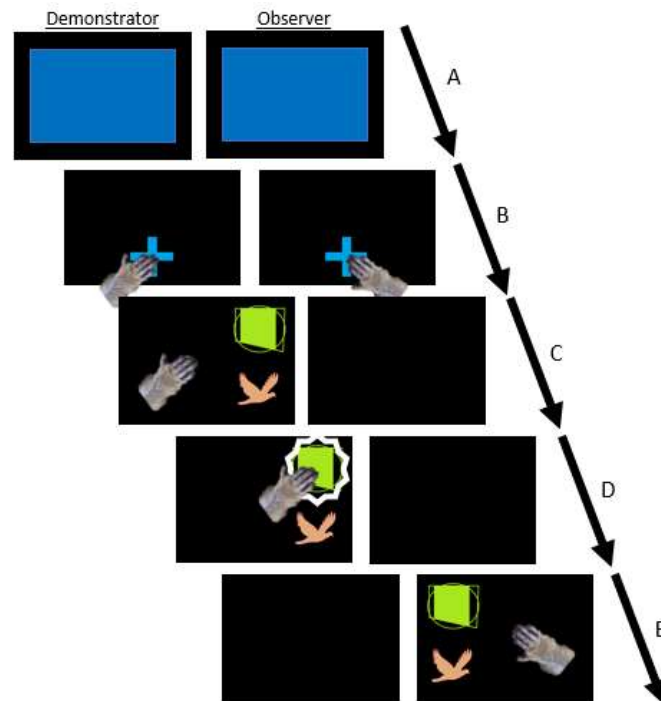
174 Social - Automated Computer Learning Devices for Monkeys (S - ALDMs)

175 The ALDMs were connected in pairs to allow an individual in one ALDM to see the touchscreen of the
176 neighboring ALDM (see Fig.1b; we dubbed this new version of the ALDM system the Social – ALDMs;
177 see videos). Two individuals could therefore see each other and their responses on the screen when a
178 transparent partition between the ALDM was used. All the monkeys had previously participated in
179 studies using this ALDM testing system, but the current task was entirely novel.



180
181 Figure 1: (a) 19 baboons living in an enclosure have ad libitum access to two trailers, each containing
182 several S-ALDM operating conditioning systems. The diagram shows the two bungalows and the five
183 units. (b) Detail of the organization of a S-ALDM, where two monkeys (1 and 2) can work side by side,
184 while seeing each other.

185 We used two different tasks, one when two individuals are present, the “dual task”, and one when an
186 individual is alone, the “single task”. These two tasks are presented below. When a monkey entered
187 an ALDM, a blue screen was displayed with a 4-second delay. If another individual was detected in the
188 neighboring ALDM during this delay, a blue screen was also displayed, announcing the synchronization
189 of the two machines and the start of the dual task (Fig.2). If the neighboring ALDM stayed empty for 4
190 secs, the single task started.



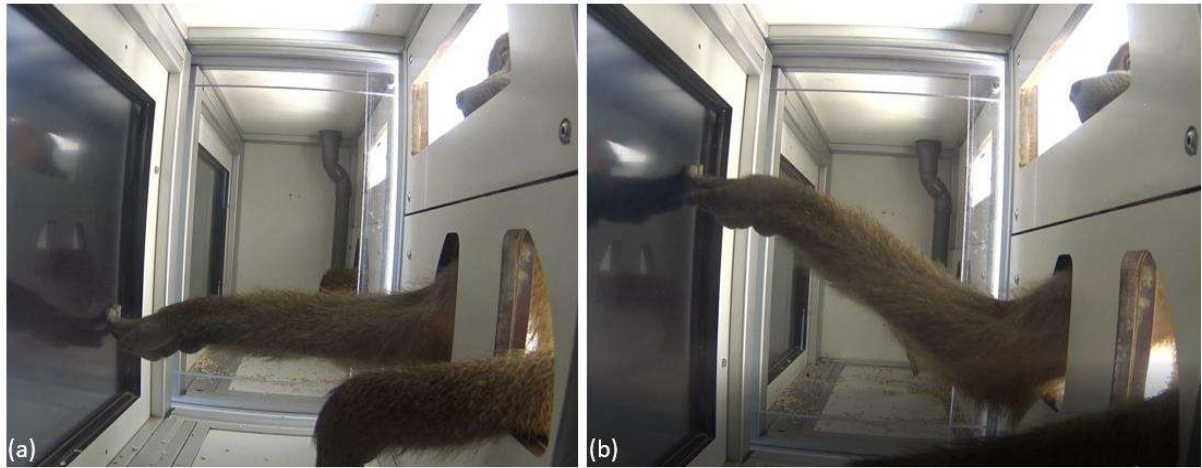
192

193 Figure 2: Schematic representation of the automated dual task on touch screen. The stimulus circled
 194 in white flashes briefly when selected. A to E: succession of events. (A) Each trial, one monkey is
 195 randomly selected to be the demonstrator, the other the observer. (B) Each monkey has to first touch
 196 a fixation cross. (C) Two randomly selected stimuli selected from a set of 70 appear on the side of the
 197 screen at a random (top or bottom) position. (D) when the demonstrator touches one of the two
 198 stimuli, it flashes briefly, then the stimuli disappear. (E) the same stimuli, in the same position,
 199 appeared on the observer's screen who had to touch the stimuli that was touched by the demonstrator
 200 to be rewarded in the compatible condition and the opposite stimuli in the incompatible condition.

201

202 The dual task (Fig. 2), which is the main focus of this paper, used a Stimulus-Response Compatibility
 203 (SRC) paradigm to test the capacity of baboons to learn from a demonstrating partner. In the social
 204 version of the SRC task, the test program identified the two monkeys when they entered a S-ALDM,
 205 synchronized the two computers, and displayed a blue fixation cross at the bottom center of the screen

206 (Fig.2B). The test started once both monkeys had pressed the fixation cross within 4-secs (otherwise
207 the trial was aborted and re-presented). Each trial, one monkey was randomly selected as the
208 “demonstrator” and would do the first phase of the trial Two stimuli were randomly selected from a
209 set of 70 stimuli and were displayed on the screen of the demonstrator monkey (Fig. 2C). These stimuli
210 were located on the side of the screen closest to the adjacent ALDM so that the second monkey had
211 the best possible visibility. Once the stimulus was selected by the demonstrator, it would flash twice
212 (Fig.2D), then both stimuli would disappear from the demonstrator’s screen and appear on the
213 observer’s screen, in the same location. The observer then had to choose one of the two stimuli
214 according to what the neighbor had chosen (Fig.2E). Monkeys were randomly assigned to the
215 Compatible or Incompatible condition before the start of the experiment (we balanced the assignment
216 of the monkeys in terms of age and sex). Monkeys in the Compatible condition had to select the same
217 stimulus as the demonstrator to get a reward. Monkeys in the Incompatible condition had to select
218 the opposite stimulus to obtain a reward. The selection of a stimulus towards the bottom of the screen
219 or towards the top of the screen thus led to the performance of two different arm movements (Figure
220 3). Note that two monkeys assigned to two different conditions could still take part together since the
221 condition was relevant only to the observer, not the demonstrator. The time in milliseconds (ms)
222 between the end of step (D) and the end of step (E) was used as response time (RT). The success of a
223 test was coded as a binary variable: correct response = 1 and incorrect response = 0. For the
224 demonstrator, any choice made by the observer triggered the delivery of a reward. For the observer,
225 a successful response triggered the delivery of a reward, a miss resulted in a 3 secs timeout.



226

227 Figure 3: Stills from a video recorded during the social reversal learning phase of a monkey choosing
228 the stimulus at the bottom (a) or at the top (b) of the screen.

229 Unlike previous studies on the automatic imitation effect in animals, we used a stimulus-response
230 compatibility protocol involving a visual device and the selection of a stimuli on a screen. Importantly,
231 the monkeys are used to work on these screens on visual tasks and therefore, according to the
232 experience dependent model of imitation (e.g., the ASL model; Ray & Heyes, 2011), they should have
233 formed bidirectional sensorimotor links associated with the use of their arm/hands to touch the
234 screen. Thus, we can expect that the observation of a monkey making a gesture towards a region of
235 the screen will lead to better performances in an observing monkey that has to make the same gesture
236 compared to a monkey that has to make the opposite gesture.

237 *Ghost task*

238 In addition to the social condition, we introduced a ghost condition that was identical to the social
239 condition as possible. In this condition, one ALDM for each pair of test systems (S-ALDM) was closed
240 and therefore inaccessible (one monkey could enter the pair and see the two screens but no one could
241 enter next to them). When a monkey entered the opened ALDM, the task proceeded exactly as in the
242 social phase, but all the actions of the neighboring device were automatically performed by the
243 computer. Each trial, the participant was randomly assigned to the observer or demonstrator role to
244 preserve the similarity between the social and the ghost phases. In other words, when the monkey

245 was selected as the observer, the neighbouring computer first selected a random stimulus that flashed
246 before the monkey made its choice. When the monkey was selected as the demonstrator, the
247 neighbouring computer waited for the demonstrator to choose a stimulus before selecting a random
248 response. The computer's response time was randomly selected from a uniform distribution based on
249 the median response time of the monkeys in the initial social phase (900 ms +/- 450, i.e. [450,1350];
250 see below).

251 *Stimuli and position bias avoidance procedure*

252 To avoid systematic responses from the demonstrator, such as always choosing the same kind of
253 stimuli (e.g., round vs. spiky) we selected 70 stimuli with similar salience from a set of 1000 (the
254 selection procedure for the stimuli is available in ESM). To avoid systematic responses based upon
255 position (i.e., bottom or top of the screen) we also limited to three the number of times the
256 demonstrator could choose the same position in consecutive trials. After the fourth choice in the same
257 location following trials were considered a miss on the part of the demonstrator until the position
258 changed and the demonstrator's screen immediately displayed a three second timeout.

259 *Single task*

260 If a monkey presented itself at the workstation alone and no other monkey participated in the task in
261 the neighboring workstation (at the time of synchronization of the blue screens, see Fig. 2A), a version
262 of the Wisconsin Card Sorting Task (WSCT, Berg, 1948) was presented. In the task, a monkey could
263 choose between three stimuli combining three colours and three shapes. Reward reinforcement
264 allowed the monkeys to deduce the rule associated with the properties of the stimulus (for instance,
265 round). When an individual reached 80% success in a block of 60 trials, the rule switched, and
266 perseverance was measured by the number of times the individual continued to choose according to
267 the old rule. Using this secondary task allowed the monkeys to maintain their motivation to participate
268 in the experiment while retrieving interesting data on cognitive flexibility. We collected 1 110 770 trials
269 using this task and since it was a well-known task, we did not expect any relationship between

270 performance on this “single task” and the main study. Post hoc inspection of the results did not reveal
271 any pattern and therefore these data will not be further analyzed in the context of this paper.

272 **Experimental timeline**

273 The 19 baboons were divided into two groups (compatible and incompatible), homogenized according
274 to their age, sex, and level of participation to the experiments (see detail in T1/ESM). All monkeys
275 followed the same experimental procedure, and all went through the same social and ghost phases
276 (following the experimental guide detailed in fig 4 and just below).

277 *Phase One and Two: Initial social and ghost test*

278 To determine any predisposition for a compatible or incompatible response, the two groups of
279 monkeys carried out a first social phase followed by a ghost phase without changing conditions. If the
280 monkeys had no prior bias, there should be no difference between compatible and incompatible
281 conditions. In these first two phases, we collected a comparable number of trials when the monkeys
282 were Observers (N =4850 in initial social phase and N =6900 in initial ghost phase for N=7 monkeys)
283 for a comparable experimental duration (5 days).

284 *Phase Three: Social reversal learning phase*

285 After these first phases, we introduced a "Social reversal-learning " phase, during which monkeys
286 alternated between compatible and incompatible conditions upon reaching a predefined criterion of
287 80% success on 50 trials. The objective of this phase was to evaluate the effect of long exposure to the
288 task on compatible and incompatible responses. Unfortunately, this phase began two weeks before
289 the announcement of the covid-19 pandemic containment measures. During this 16-day period, we
290 collected 46 350 observer trials (N= 7 monkeys). The experiment started again 3 months later, after
291 the confinement, and we collected 72 702 observer trials (N= 7 monkeys) and 86 788 demonstrator
292 trials over an additional 27 days.

293 *Phase Four: Ghost reversal learning phase*

294 After the social reversal phase, we performed a ghost reversal phase using the same criterion as
 295 before. The baboons performed 71 719 observer trials and 82 198 demonstrator trials over 20 days, a
 296 number comparable to the after-shutdown social reversal phase. All monkeys started the ghost phase
 297 in the condition they were in at the end of the social reversal phase. Another possibility would have
 298 been to use the condition they were in at the beginning of the social reversal phase, but since different
 299 monkeys had performed a different number of reversals, some monkeys would have started in the
 300 same condition and others in a different condition.

Name	Sex	Age	Initial Social Phase	Initial Ghost Phase	Social Reversal Learning Phase						Ghost Reversal Learning Phase	
					Number of reversal						Number of reversal	
					0	1	2	3	4	5	0	1
Bobo	M	162	Compatible (1)	(0)	(56)						(5)	
Atmosphère	F	263	Compatible (2)	(1)	(47)						(39)	
Feya	F	114	Compatible (14)	(1)	(28)	(65)	(77)	(4)			(101)	
Arielle	F	172	Compatible (10)	(2)	(239)						(167)	
Lips	F	51	Compatible (13)	(22)	(39)	(150)	(58)				(115)	(17)
Mako	M	43	Compatible (9)	(28)	(61)	(95)	(55)				(35)	(41)
Muse	F	44	Compatible (11)	(18)	(16)	(33)	(50)	(60)	(52)	(34)	(103)	
Felipe	M	117	Compatible (3)	(1)	(98)						(1)	
Petoulette	F	251	Compatible (2)	(0)	(13)						(25)	
Ewine	F	127	Compatible (23)	(7)	(198)	(140)					(180)	
Number of monkeys that reached each reversal phase					10	5	4	2	1	1		
Pipo	M	246	Incompatible (1)	(0)	(20)						(3)	
Kali	F	293	Incompatible (2)	(0)	(26)						(20)	
Fana	F	120	Incompatible (12)	(5)	(205)	(24)					(73)	
Violette	F	182	Incompatible (27)	(16)	(122)	(153)	(32)				(206)	
Mali	F	48	Incompatible (9)	(27)	(63)	(88)	(47)				(88)	
Lome	M	54	Incompatible (14)	(26)	(3)	(25)	(37)	(38)	(26)	(30)	(5)	(50)
Nekketsu	F	28	Incompatible (6)	(13)	(31)	(137)					(76)	
Harlem	M	91	Incompatible (6)	(13)	(136)						(23)	
Angele	F	177	Incompatible (4)	(3)	(73)						(23)	
Number of monkeys that reached each reversal phase					9	5	3	1	1	1		

301

302

303 Table 1: Course of the experiment for the 19 monkeys across the four phases (initial social, initial ghost,
 304 social reversal learning and ghost reversal learning). They started either in compatible condition (Blue)
 305 or in incompatible condition (Red). Numbers represent the number of 50 trials blocks performed in
 306 each condition. Empty cells mean they have not reached this stage of the experiment. In bold: monkeys
 307 that reached at least once the success threshold.

308 **Statistical analysis**

309 The main objective of the statistical analyses was to determine the influence of the experimental phase
 310 (social vs. ghost) and the condition (compatible vs. incompatible) on performance in the task (score

311 and reaction time were analysed separately). To take into account inter-individual variability and
312 repeated measurement, we used mixed models including a random intercept and slope (represented
313 by the number of blocks of 50 trials performed) for each individual (all the model details are presented
314 in the ESM). Thus, we limit our analysis to complete blocks of 50 trials. The score was coded as a binary
315 variable (success = 1, failure = 0), and we averaged it over each block of 50 trials to get the mean score.

316 Moreover, since the first objective of the experiment was to study variation in performance depending
317 on condition (compatible - incompatible) and how it would vary after training in the opposite
318 condition, we chose a priori to analyse the scores of monkeys having reached the threshold at least
319 twice (e.g., a monkey starts in compatible condition, reaches the threshold, switches in incompatible
320 condition and reaches the threshold again) (N = 7). Given the small number of individuals who
321 succeeded at the task, it was not possible to analyse the effect of age or sex in any detail but as can be
322 seen in table 1, there is no obvious pattern.

323 The effect of the predictors on Score was evaluated using a generalized mixed linear model (GLMM,
324 Baayen, 2008) with binomial error distribution and a logit link function. Models were developed using
325 the glmer function of R's lme4 package (Bates et al., 2015). We selected the compatibility condition,
326 and its interaction with the number of blocks performed, as fixed predictors of interest. We set the
327 intercept at trial block zero and at condition Incompatible. The effect of the predictors on the response
328 time (RT) was evaluated using a mixed linear model of similar structure to the one used for the score.
329 Because we were interested in how the action of a demonstrator could affect the speed with which
330 the observer did the same action, we only analysed RTs from successful trials, and we filtered RTs (+/-
331 2 SDs) to avoid biasing the models with overly large distributions.

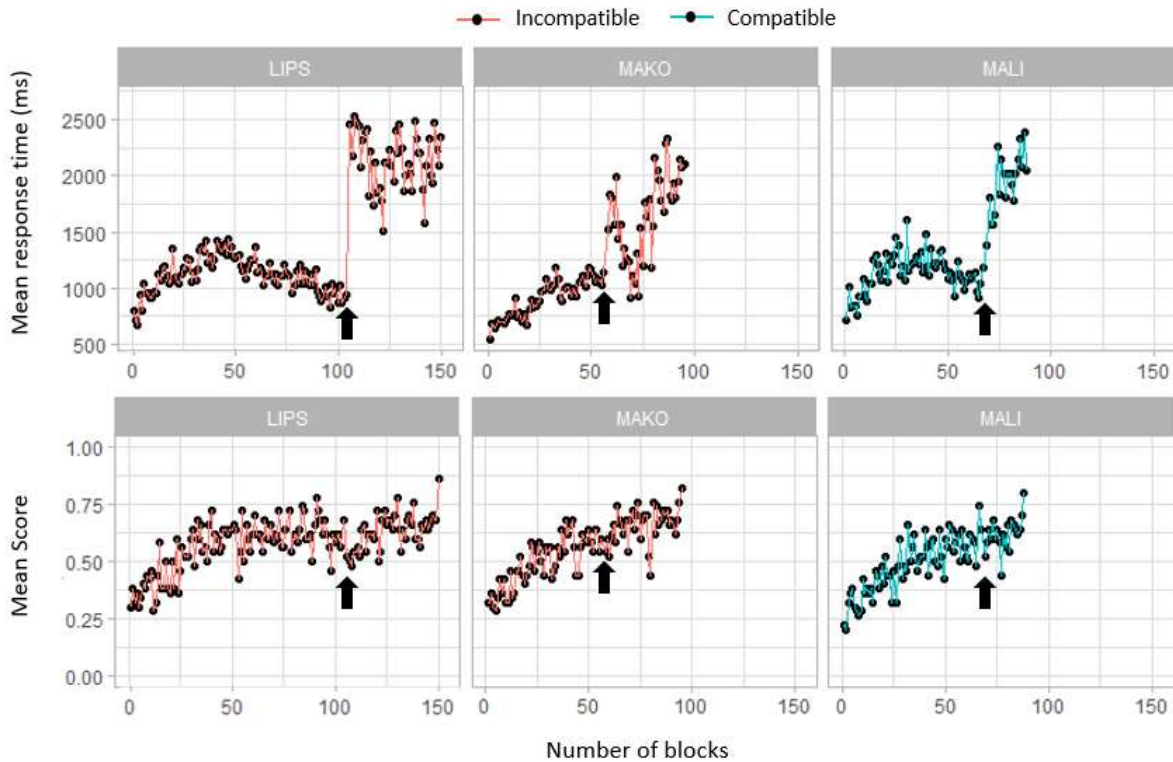
332 During our analysis, we faced singularity problems despite our relatively simple random-effects
333 structure containing only a random intercept and slope. There seems to be no agreed upon solution
334 to this problem and guidelines recommend achieving a balance between reducing the complexity of
335 random factors to achieve convergence and risking finding spurious effects (Bates et al. 2015). In our

336 case, the baboons show some systematic inter-individual variability (see Fig.6 below for instance) and
337 we found that model estimations were more realistic using the intercept plus slope random effects
338 rather than just a random intercept (despite the singularity). We therefore chose to present the results
339 of the models with the full random effects. All analyses were done with the statistical software R (R
340 Core Team 2020)

341 **Results**

342 The experiment lasted 73 days, during which the 19 baboons performed a total of 523 403 dual-task
343 trials (an average of 984 trials/monkeys/day, [min average =124 trials/day; max average = 2174
344 trials/day]), of which 240 619 were conducted as observers (on average 46% of the total number of
345 trials [min = 44%, max = 47%]). This is less than expected by chance (binomial test, $p < 0.001$) and
346 suggests that monkeys were more likely to give up on a trial when they were observers. During the
347 social reversal phase, ten monkeys passed the criterion of 80% success in a block of 50 successive trials.
348 After changing their condition, seven among them reached the threshold a second time.

349 Regarding RTs, we noticed that the laboratory's shutdown had little impact on the score but affected
350 RTs significantly (see Fig. 5 and we provide more details in S1/ESM). Since most of the monkeys had
351 already reached the success criterion once (first reversal), we decided to analyse the RTs of these
352 monkeys before and after the first reversal, but not beyond the break. Because we wanted to compare
353 the evolution of RTs over an identical period (same number of blocks) before and after the first
354 reversal, we didn't analyse the two entire period. We had to choose a number of blocks to analyse that
355 on the one hand would represent the number of blocks needed to reach the success threshold before
356 the reversal, but on the other hand would also avoid including the disruption of the RTs following the
357 shutdown. Therefore, we set the number of blocks at 39 (the median number of blocks for monkeys
358 to reach the first success threshold). Because we could only analyse RTs before and after the first
359 reversal, we decided to perform the following analyses on the ten monkeys that reached the success
360 threshold at least once.



361

362 Figure 5: Average RT and Score of three monkeys impacted by the closure of the platform after they
 363 had reached a first success threshold. Black arrow: closure of the platform. Red line: Incompatible
 364 condition; Blue line: Compatible condition.

365 *Phase one and two: Initial social and ghost test*

366 The purpose of the first phase was to estimate the presence of a bias between the compatible and the
 367 incompatible group.

368 The results of the analysis showed a slight interaction between the experimental condition (compatible
 369 - incompatible) and block number (GLMM, Condition x Number of Blocks, $\beta = -0.08$, SE= 0.04, $z = -1.98$,
 370 $p = 0.047$). The score increased in the incompatible condition ($\beta = 0.09$, SE= 0.03, $z = 3.05$, $p = 0.002$) but
 371 not in the compatible condition ($\beta = 0.01$, SE= 0.03, $z = 0.55$, $p = 0.58$), and the score in the incompatible
 372 group was significantly higher at the beginning of the phase ($\beta = -0.65$, SE= 0.23, $z = -2.832$, $p = 0.005$)
 373 than in the compatible group. Looking at reaction times, we found no interaction between condition
 374 and number of blocks (LMM, $\beta = -0.45$, SE= 14.98, $t = -0.03$, $p = 0.98$). Initially monkeys were faster in
 375 the incompatible condition ($\beta = 291.64$, SE= 125.68, $t = 2.32$, $p = 0.02$) and RT decreased with the number

376 of blocks in both conditions (Compatible: $\beta = -21.5$, $SE = 8.34$, $t = -2.58$, $p = 0.01$; Incompatible: $\beta = -25.81$,
377 $SE = 11.23$, $t = -2.3$, $p = 0.02$), a result typical of such experiments where monkeys become faster over
378 the time.

379 During the initial ghost condition, the results showed a small interaction between the condition and
380 block number on scores (GLMM, $\beta = 0.04$, $SE = 0.021$, $z = 2.03$, $p = 0.04$). The score in the incompatible
381 group decreased ($\beta = -0.04$, $SE = 0.015$, $z = 2.29$, $p = 0.02$) while it stayed stable in the compatible group
382 ($\beta = 0.002$, $SE = 0.014$, $z = 0.48$, $p = 0.63$) and they were no initial difference at the beginning of the phase
383 ($\beta = -0.59$, $SE = 0.43$, $z = -1.37$, $p = 0.17$). Regarding reaction times, we found a significant interaction
384 between condition and number of blocks ($\beta = -15.39$, $SE = 6.34$, $z = -2.43$, $p = 0.015$). RTs increased in the
385 incompatible group ($\beta = 13.26$, $SE = 4.5$, $t = 2.94$, $p = 0.003$), but not in the compatible group ($\beta = -2.13$,
386 $SE = 4.5$, $t = -0.48$, $p = 0.63$). Moreover, RTs were faster in the incompatible group at the beginning of
387 the phase compared to the compatible group ($\beta = 207.9$, $SE = 94.97$, $t = 2.19$, $p = 0.028$).

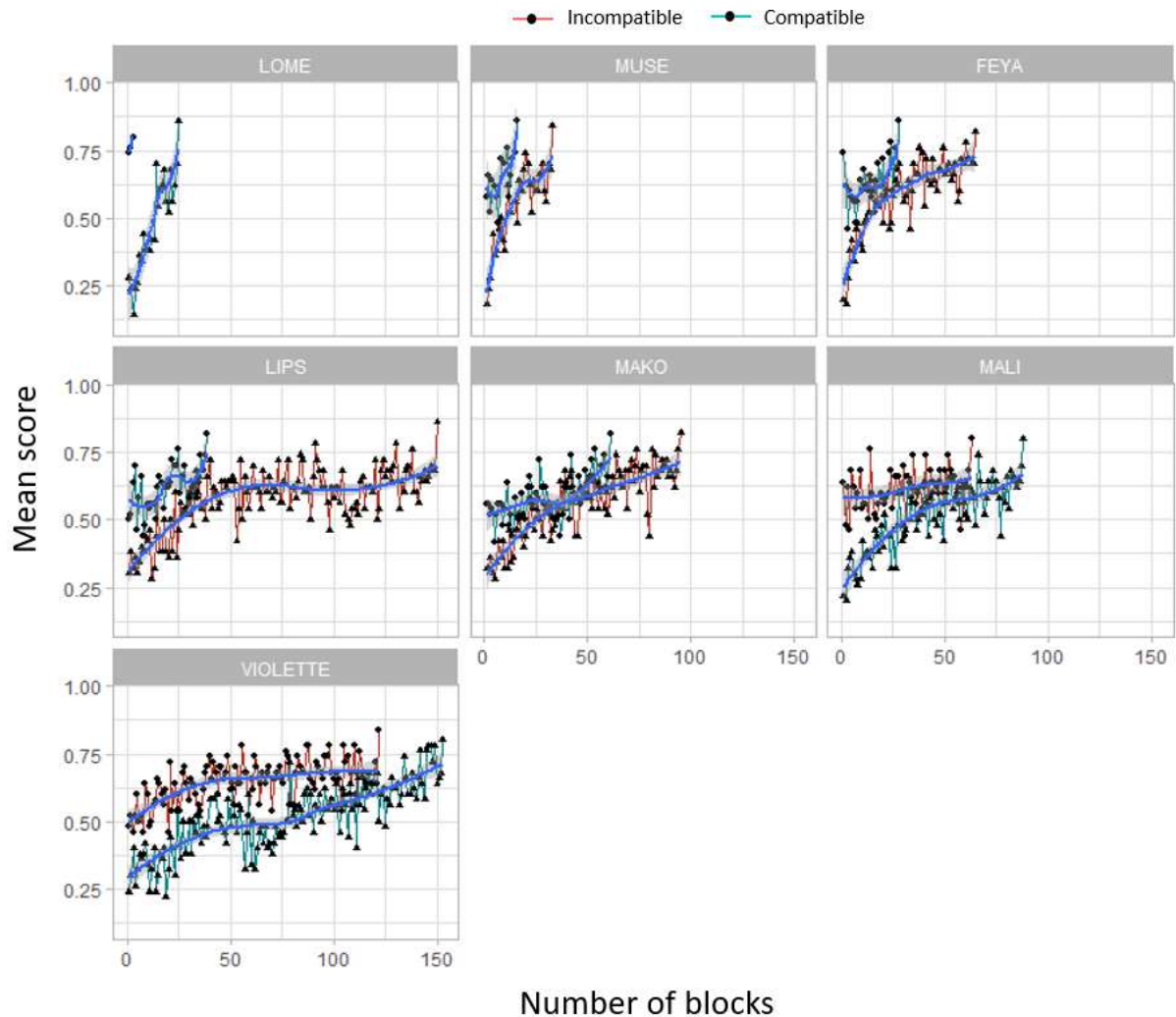
388

389 Taking a closer look at the individual results, we noticed that LOME, a monkey who had quickly reached
390 an accuracy score of 80% during the initial social phase (in incompatible condition), dropped to 20%
391 success during the ghost phase and MALI, who also reached a score of 80% during the initial social
392 phase (in incompatible condition), also dropped to 50% success during the ghost phase. The change in
393 behaviour of these two monkeys can explain the results of this first phase (further discussed in the
394 ESM/S2/S3). Regarding the other individuals, they showed no evidence of learning the task.

395 *Phase three: Social reversal learning*

396 During the social reversal learning phase, seven monkeys (among the group of 19) reached the success
397 threshold at least twice (i.e., performed at least two reversals). Four of them began in the compatible
398 condition. We found no significant interaction between condition and number of blocks in the first
399 reversal phase (Condition x Number of Blocks: $\beta = 0.008$, $SE = 0.005$, $z = 1.59$, $p = 0.11$), indicating that
400 the two groups learned at similar speed. There was also no evidence of an initial difference between

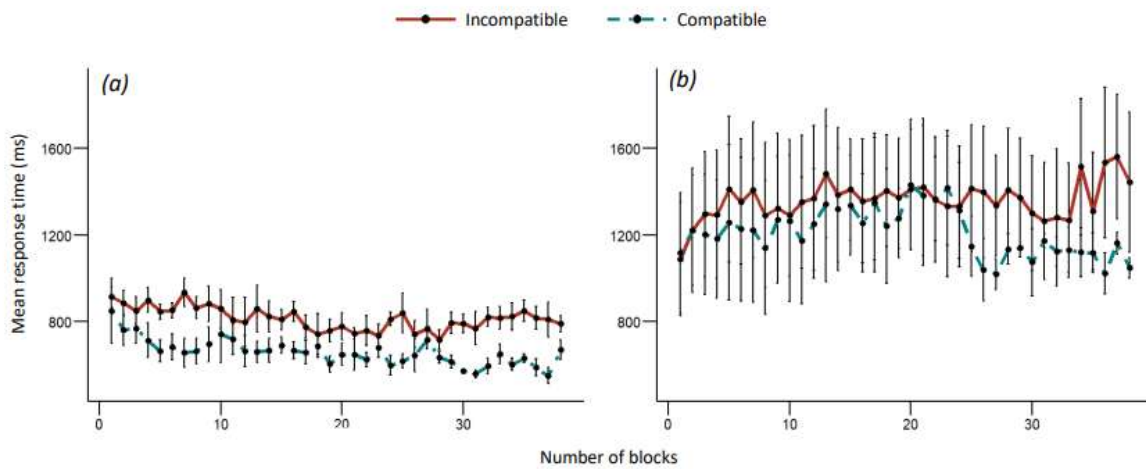
401 the condition ($\beta = -0.31$, $SE = 0.2$, $z = -1.53$, $p = 0.13$). The average score of monkeys in the Incompatible
 402 condition increased at an estimated rate of $\beta = 0.010$ ($SE = 0.006$, $z = 1.706$, $p = 0.09$) while it increased
 403 in the Compatible condition at an estimated rate of $\beta = 0.019$ ($SE = 0.005$, $z = 3.802$, $p < 0.001$). (Fig.6).



404
 405 Figure 6: Mean score in incompatible (red) and compatible (blue) condition, before (circle) and after
 406 (triangle) the first reversal, for the seven monkeys that performed at least two reversals. Blue lines
 407 represent the local regression (LOESS method).

408 Regarding reaction times (Fig. 7), we found no interaction between the condition and the number of
 409 blocks ($\beta = -2.28$, $SE = 3.6$, $t = -0.63$, $p = 0.53$), there were no initial difference between the RTs in both
 410 condition ($\beta = 106.36$, $SE = 78.4$, $t = 1.35$, $p = 0.18$). RTs were stable during this phase in incompatible
 411 condition ($\beta = -2.9$, $SE = 2.62$, $t = -1.12$, $p = 0.26$) and increased slightly in the compatible condition ($\beta = -$

412 5.2, SE= 2.48, t= -2.1, p= 0.04). Furthermore, when we averaged RTs for each monkey during this phase,
413 we found no clear evidence of a difference between conditions (N = 10, t = -1.86, df = 7.72, p = 0.10).



414

415

416 Figure 7: Mean response time in the incompatible (red solid line) and compatible (blue dotted line)
417 conditions during the social reversal learning phase before (a) and after (b) the first reversal. Error bars
418 represent the standard error of the mean.

419 After reaching the first 80% success threshold, seven monkeys changed condition and entered the
420 second reversal. We found no significant difference in the evolution of the monkeys' scores during the
421 second reversal (Condition x Number of Blocks, $\beta= 0.014$, SE=0.022, z=0.669, p=0.5). However, we
422 found an initially slightly higher average score in the Incompatible condition ($\beta=-0.43$, SE=0.2, z=-2.16,
423 p=0.03), and the average score of monkeys in the Incompatible condition increased at a rate of $\beta=0.024$
424 (SE=0.014, z=1.75, p=0.08) and at a rate of $\beta= 0.039$ (SE= 0.016, z= 2.37, p =0.018) in the Compatible
425 condition.

426 In the first 39 blocks after the first reversal (Fig 7b, we found no interaction between the condition and
427 the number of blocks ($\beta= 10.3$, SE= 6.19, t= 1.66, p= 0.1) and there was no initial difference between
428 the RTs in both condition ($\beta= -165.7$, SE= 429.2, t= -0.39, p= 0.7). Furthermore, RTs were stable during
429 this phase in both conditions (Compatible: $\beta= 8.07$, SE= 4.39, t= 1.84, p=0.066; Incompatible: $\beta= -2.2$,

430 SE= 4.36, t= -0.5, p=0.61). Similarly, we found no difference between averaged RTs (N = 10, t = -0.14,
431 df = 7.99, p = 0.89).

432 After the second reversal, the seven monkeys changed condition again and we found no significant
433 effect of the condition on the learning of the task ($\beta = -0.002$, SE=0.009, $z = -0.265$, p=0.79), we found
434 no initial difference between the two conditions ($\beta = -0.15$, SE=0.11, $z = -1.31$, p=0.19) and the score
435 progressed at a similar rate in the two conditions (Compatible: $\beta = 0.02$, SE=0.006, $z = 3.73$, p<0.001;
436 Incompatible: $\beta = 0.02$, SE= 0.007, $z = 3.35$, p<0.001).

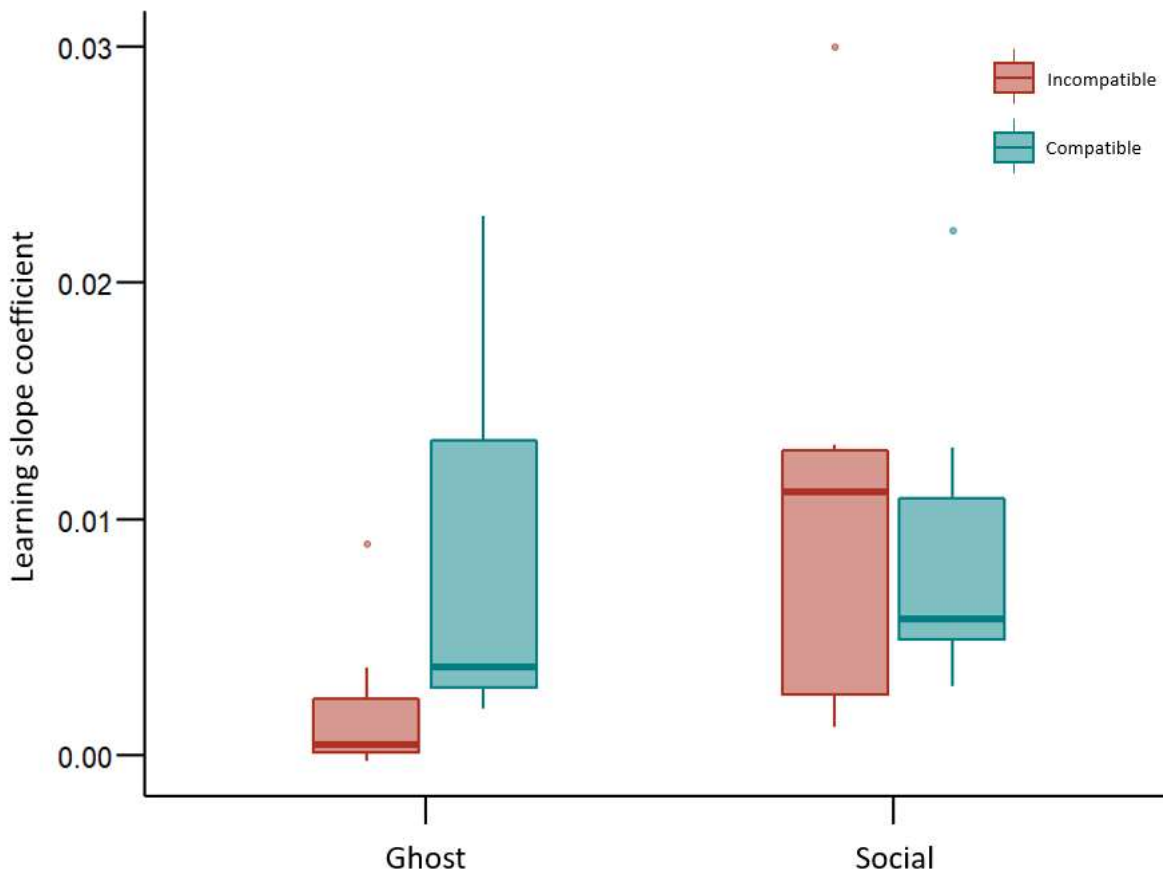
437 Note that we cannot directly compare the first reversal to the second (within individuals) because
438 during the first reversal the monkeys had to learn a condition, while during the second they had to
439 unlearn the first condition and learn the new one. Their learning rate during the second phase was
440 therefore different from the initial learning phase. Only three of the seven monkeys performed a third
441 reversal (FEYA, LOME and MUSE), and two continued until they reached five reversals (LOME and
442 MUSE; see Table 1). Then, we analysed separately the results of LOME and MUSE and did not find a
443 significant difference between the learning rate in the compatible and incompatible conditions (LOME,
444 $\beta = 0.045$, SE=0.03, $z = 1.40$, p=0.16 and MUSE, $\beta = -0.0028$, SE=0.011, $z = -0.258$, p=0.8; see Figure S4 and
445 S5).

446

447 *Phase four: Ghost reversal learning*

448 In the ghost reversal learning phase, three monkeys (LIPS, LOME, MAKO) among the seven monkeys
449 that reached this phase reached the success threshold once. None of the other 12 monkeys, who had
450 the opportunity, reached such criterion. The three monkeys that passed were all in the Compatible
451 condition, while none of the four monkeys in the Incompatible condition passed. Consistent with this
452 observation, we found a significant interaction between condition and block ($\beta = 0.006$, SE=0.002,
453 $z = 3.89$, p < 0.001) suggesting that monkeys in the compatible condition learned faster than monkeys
454 in incompatible condition. We found no initial difference between conditions ($\beta = -0.23$, SE=0.24, $z = -$

455 0.97, $p=0.33$), but the average score of monkeys in the Incompatible condition remained stable
456 ($\beta=0.0004$, $SE=0.0004$, $z=1.21$, $p=0.23$) while it increased in the Compatible condition ($\beta=0.006$,
457 $SE=0.015$, $z=4.30$, $p<0.001$). Furthermore, by comparing the ghost reversal phase with a comparable
458 number of trials at the beginning of each phase of the social reversal phase, we found a significant
459 difference between the two phases ($N=7$, $W=143$, $p=0.01$), monkeys learned faster in the social
460 reversal phase compared to the ghost reversal phase. Looking at each condition separately, this
461 difference was significant only for the Incompatible condition (Compatible: $W=17$, $p=0.6$;
462 Incompatible: $W=56$, $p=0.008$) (Fig. 8).

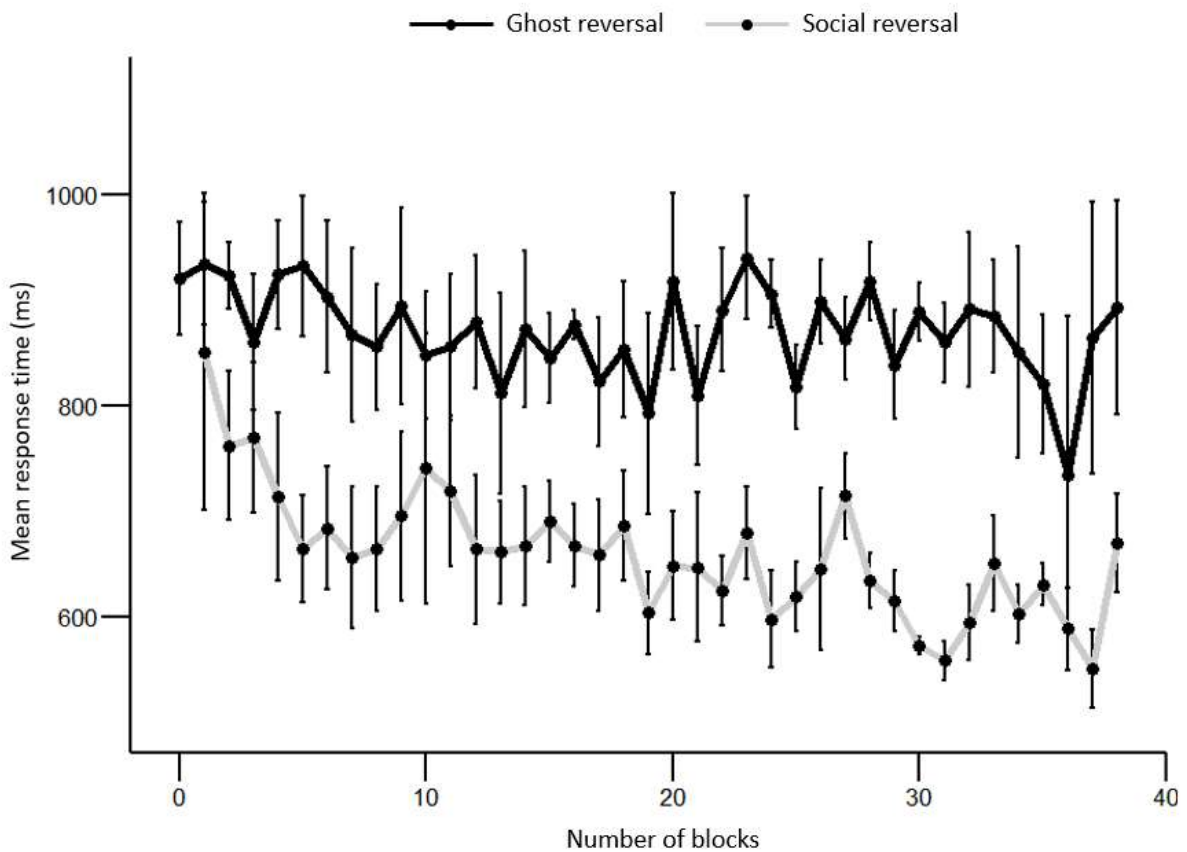


463

464 Figure 8: Distribution of the learning slopes of the seven monkeys that performed both the Social and
465 Ghost reversal learning phases in incompatible (red) and compatible (blue) condition.

466 Since none of the monkeys in the incompatible condition reached the success threshold, we could not
467 compare RTs between the compatible and incompatible conditions of the ghost phase. However, one

468 of the purposes of this study was to determine if the presence of a nearby individual was affecting the
469 response time of the focal participant. The comparison of the first blocks of the Ghost and Social
470 reversal phase in compatible condition, suggests that the monkeys in the Social condition have slightly
471 faster RTs than monkeys in Ghost phase (Fig. 9). We found no effect of the interaction between the
472 phase and the number of blocks on the monkey's RT ($\beta = -0.56$, $SE = 0.68$, $t = -0.82$, $p = 0.41$). We found
473 an initial difference between phases ($\beta = -290.69$, $SE = 15.62$, $t = -18.61$, $p < 0.001$) but the RTs progressed
474 at a similar rate in both reversal phases (Ghost: $\beta = -3.6$, $SE = 2.3$, $z = -1.57$, $p = 0.12$; Social: $\beta = -3.35$, $SE =$
475 1.93 , $z = -1.73$, $p = 0.08$). When we averaged RTs for each monkey during this phase, we found a small
476 non-significant difference between conditions ($N = 10$, $t = -1.98$, $df = 7.98$, $p = 0.08$).



477
478 Figure 9: Mean response time in compatible condition during the social reversal learning phase (light
479 grey) and the ghost reversal learning phase (dark) before the first reversal. Error bars represent the
480 standard error of the mean.

481 These results suggest that the ghost condition had a detectable effect on the incompatible condition,
482 not the compatible one. However, surprisingly, the learning rates of all three successful monkeys
483 remained high after they changed for the Incompatible condition. Their learning rates were not
484 significantly different from the rates observed during the social reversal phase ($W=7$, $p=0.28$). Despite
485 this lack of difference at the group level, we did, anecdotally, notice some interesting individual
486 patterns. We found that LOME, the monkey that achieved the success criterion the most times over
487 the entire experiment, took an average of 1100 trials (150,1850 and 1300) to do so in incompatible
488 condition during the social reversal phase. After reaching the criterion in the compatible condition in
489 the ghost phase, he failed to succeed in the incompatible condition after 2500 trials before the phase
490 ended.

491 **Discussion**

492 Domain general accounts of imitation (e.g., Heyes & Ray 2000; Prinz. 1997) propose that imitation is
493 the result of associative learning during sensory-motor experience. Domain specific views on imitation
494 argue that imitative capacities are adaptations to address specific problems (e.g., Subiaul, 2010), but
495 that experience will also guide the development of various social learning skills throughout
496 development (Meltzoff & Moore, 1997). Given the central role of experience under both of these
497 accounts, the main purpose of this study was to examine imitative behaviour in baboons in the context
498 of experience of both compatible and incompatible rules, across social and asocial conditions.

499 Specifically, we were interested in whether baboons might have initial preferences to imitate
500 behaviours and whether they may learn compatible rules more quickly than incompatible rules and/or
501 reach higher performance (in terms of score and/or response times). We also wanted to examine
502 whether any differences in performance stay stable over time (i.e., during reversal learning phases). In
503 addition to examining imitative biases, we also wanted to test the importance of the social component
504 of this SRC paradigm, namely the role of the demonstrator, by contrasting the social condition to a

505 ghost control. Both domain-general and domain-specific accounts predict that learning should be
506 impaired in the ghost condition for social species.

507 Our results showed that, during the initial social phase of testing, the score increased slightly in the
508 incompatible condition but not in the compatible condition. This is likely linked to the fact that two
509 monkeys of the incompatible group quickly reached an average score of 80%. The performance of
510 these two monkeys does not specifically point towards a bias for counter-imitation but is more likely
511 the result of chance factors (i.e., two monkeys that are particularly proficient at this task). In fact, ten
512 monkeys reached the success criterion in the whole experiment, five in each condition. Similarly, we
513 found that RTs were slightly faster in the incompatible condition at the beginning of the experiment,
514 but this small difference is difficult to interpret. Since the scores were very low, this difference is
515 unlikely to reflect a bias for one condition over another. During this phase RTs decreased similarly in
516 the two conditions as the monkeys gained more experience with the task. During the initial ghost
517 phase, we observed no systematic bias for either condition. We can therefore conclude that initially
518 the monkeys were not biased towards one condition over another in both social and asocial conditions.

519 During the reversal phases, we found no evidence of a difference in score between the compatible and
520 incompatible conditions, with learning progressing at a similar pace in the two conditions across the
521 reversal phases. Similarly, we did not find a systematic effect on response times. Monkeys in the
522 compatible condition had a small tendency to be faster than monkeys in the incompatible condition
523 (Fig. 7a) and the absence of a significant difference could be explained by a lack of power, given that
524 only 10 individuals participated in this phase and that the difference remains small. Faster RTs in the
525 compatible condition could suggest that seeing the action of a neighbouring monkey result in a quicker
526 response for the observer by eliciting the same motor response. However, we note that this difference
527 disappeared after the first reversal (Fig. 7b).

528 The absence of a difference in score and in response times between the compatible and incompatible
529 conditions stands in sharp contrast to previous studies. Since the first results that have highlighted the

530 “automatic imitation” effect (Brass et al., 2001), many studies have focused on the mechanisms
531 involved in its properties as well as on its neurological components (Bien et al., 2009; Catmur et al.,
532 2009). Heyes (2011), proposes that automatic imitation would result from a combination of effector
533 and movement compatibility. This is consistent with recent studies suggesting that the mirror neuron
534 system can integrate and represent an entire action, using the multiple components necessary for its
535 realization (Cracco & Brass, 2017; Cracco et al., 2016). Although in humans most studies on imitation
536 bias have focused on the body part component, faster RTs (between the observation of the action and
537 the motor response) in compatible condition is a common finding (Bertenthal et al., 2006; Gillmeister
538 et al., 2008; Otte et al., 2011) and studies have found reliable results for imitation using movement
539 compatibility (Brass et al., 2001; Heyes et al., 2005; Obhi & Hogeveen, 2010). In animals however, SRC
540 paradigms generally use different body parts (e.g., head or paw with dogs). Studies have shown that it
541 was easier to reach a success criterion in compatible trials in dogs (Range et al., 2010) and capuchins
542 (O’Sullivan et al., 2017) and that budgerigars would do more correct responses in compatible trials
543 (Mui et al., 2008). In contrast to previous research with non-human animals, we used the same body
544 part in both conditions but with different movement (hand raised up or down) and it is possible that
545 biases in copying certain movements are more difficult to capture than actions made with different
546 body parts, even with precise measures of reaction times. This is the first study in non-human animals
547 to examine social learning in a stimulus-response compatibility paradigm that hasn’t used different
548 effectors, and so it is possible that in previous studies the movements examined (peck versus step, Mui
549 et al, 2008; paw versus mouth movement, Range et al. 2010; hand versus mouth action, O’Sullivan et
550 al. 2017) activate different social learning biases than those activated during the copying of
551 spatial/stimulus matching tasks. Therefore, our null findings, when considered alongside successful
552 attempts to identify imitative biased may be consistent with a multi-level approach to copying
553 behaviour in animals (Rumiati & Tessari, 2007; Subiaul, 2010).

554 Nonetheless, methodologies examining the copying of movement in a particular direction or towards
555 certain stimuli have been used across a range of social learning contexts and species (e.g., Renner et

556 al. 2021; Aplin et al., 2015), and so we would expect that social learning biases in this species might
557 drive easier copying of movement in the same direction and towards the same stimulus. Demonstrator
558 and subject were encouraged to perform actions with different hands. This choice was made for
559 practical reasons to improve visibility of movements; however, it may have influenced the likelihood
560 of observing a compatibility effect. In human SRC studies, mirrored stimuli are often used (i.e.,
561 presenting an stimulus image of a left hand when responding with a right hand; e.g., Brass et al., 2000).
562 However, in primates, it may be more important to provide demonstrations that are consistent with
563 the animal's perspective of their own body. For example, it has been suggested that in humans, self-
564 observation of actions is one of the primary means through which the sensorimotor associations that
565 facilitate imitation are formed (Ray & Heyes, 2011; Wiggett et al., 2011). This may be more important
566 in non-human primates who do not have the same variety of sensorimotor experience to facilitate
567 development of an imitative capacity (e.g., synchronous behavioural routines, interactions with
568 mirrors, etc; Ray & Heyes, 2011).

569 In humans, studies have also shown that a short period of counter-imitative training was enough to
570 cancel or to reverse the automatic imitation effect (six blocks of 72 trials in the Heyes et al. (2005)'s
571 study were enough to cancel the automatic imitation effect). In children, when 3-7 years olds were
572 asked to perform a different response to a cuing action across only 10 trials (i.e. wave when they
573 observed hand clapping) automatic imitation effects in subsequent trials were reduced (O'Sullivan et
574 al., 2018). However, it is difficult to know the extent of rehearsal needed to bring about reorganization
575 of the sensorimotor representations of an action across different species. Interestingly, we found no
576 systematic evidence of an effect of rule-reversal linked to the compatible/incompatible conditions. The
577 learning speed before and after reversal were very similar in the two conditions (Fig. 6). This suggests
578 that monkeys in our experiment used the same (associative) learning mechanism across the two
579 conditions – whether copying the same movement or moving towards a different location. Another
580 explanation may be that the monkeys were not learning to associate the observed actions with
581 performed actions (i.e., Sensory-Motor associations, S-M), but learning associations between the

582 locations (or stimuli) selected on the demonstrator display and stimuli/locations on their own display
583 (i.e. Sensory-Sensory associations, S-S). We cannot rule this explanation out, and if S-S associations are
584 the primary driver of learning here, this might also explain the differences observed in our paradigm
585 when compared to other comparative work that has used different effectors (which necessitate
586 learning a rule between stimulus and specific motor responses).

587 It should also be noted that only 10 of the 19 monkeys tested reached the pre-specified threshold
588 once, and only seven twice. The introduction of the paired-workstations in which baboons can see
589 what another individual is doing is recent (2018) and it appears that some monkeys are still showing
590 difficulties at understanding that they must, in some tasks, look at what the neighbour is doing. This
591 could explain some variability in the initial phase of learning, with certain monkeys, such as LOME,
592 reaching the criterion during the initial phase, while others took much longer. However, all monkeys
593 were equal in this respect after the first reversal.

594 In general, the monkeys tested here are indeed highly trained in screen-based tasks. We believe that
595 the impact in this experiment may be because the monkeys were highly trained to perform individual
596 tasks, and thus to look at their own screen, and not at what their neighbour is doing. It is therefore
597 possible that untrained monkeys, or monkeys that we would firstly train to socio-cognitive tasks, would
598 take less time to learn this and the potential imitation bias would be quicker to pick up. However, no
599 age effect was observed, young monkeys with less experience with ALDM show identical results to
600 some adult monkeys.

601 Finally, we wanted to test the relevance of the presence of a demonstrator in the SRC paradigm using
602 a ghost control. Firstly, we note that during the initial testing phase, the scores of the two monkeys
603 that had managed to reach the criterion in the incompatible condition dropped when they entered the
604 ghost phase. Secondly, during the ghost reversal learning phase, our results show that only three
605 monkeys among nine reached the success threshold and that they all were in compatible condition.
606 The six monkeys in the incompatible condition showed no sign of improvement at all. The monkeys

607 therefore exhibited a real difficulty in learning the incompatible condition in the ghost condition (when
608 it is required to pick the opposite stimulus to the one that flashes), whereas they could all do so with
609 a demonstrator. Surprisingly, all three monkeys that learned the compatible condition during the ghost
610 phase started learning the incompatible condition without apparent difficulty after reversal. Although
611 this could be due to chance alone (maybe it just happens to be three monkeys that are good at this
612 task), it could also be that the individuals need to first associate the relevant cue (flashing) to their
613 response and they can do so more easily in the compatible condition. Once the monkeys have learned
614 that the correct response depends on stimulus flashing rather than on other features (such as colour,
615 shape, etc), they can then generalise this learning to the incompatible rule.

616 The fact that three monkeys learned the compatible rule in the ghost condition is informative. Studies
617 of imitation using the SRC paradigm assume that the presence of a demonstrator is a necessary
618 component of the process. Some studies have shown that human children, apes and monkeys are
619 capable of learning from a ghost display in touch-screen studies (Ferrucci et al., 2019; Renner et al.,
620 2020; Renner et al., 2021) or in other social learning studies (Huang & Charman, 2005; Tennie et al.,
621 2006; Thompson & Russell, 2004), however, the learning deficit we see in these conditions in
622 nonhumans primates in particular (see Hopper, 2010) suggests that monkeys and apes are at a
623 significant disadvantage without the scaffolding afforded by conspecific behaviour or mere presence.
624 In humans, a robotic hand can still elicit an automatic imitation effect, although one less potent than
625 a human hand (Press et al., 2005). If the SRC effect exists in the absence of a demonstrator, is it still
626 relevant to imitation? Noticeably, when we compared the social and ghost phases' RTs in the
627 compatible condition, we found a trend towards faster average RTs in the social condition (Fig. 9). This
628 suggests that the movement of the demonstrators elicits a faster response than the flashing cue, but
629 the flashing cue is still effective. It has been reported that guinea baboons in the wild are tolerant to
630 the presence of conspecifics (Fischer et al., 2017) and it could be that presence of a conspecific leads
631 to quicker actions in foraging and food processing behaviour so as to avoid potential competition. For
632 example, while much social learning literature assumes that animals will benefit from copying, in some

633 cases it is advantageous to perform different or complimentary behaviours. Hopewell et al. (2010)
634 found that grey squirrels find it easier to learn to choose a pot that a conspecific had not examined
635 over a pot where a conspecific had previously obtained a reward, a finding consistent with grey squirrel
636 foraging behaviour. Indeed, here, a bias towards an opposite behaviour (or location) may explain some
637 of the significant effects – that is, before taking part in this study participants may have learned that
638 food rewards are more likely to be obtained when interacting with locations that are different from
639 those explored by conspecifics. While beyond the scope of this study, imitation research driven by an
640 understanding of naturalistic foraging behaviours may shed light on both imitative and counter-
641 imitative biases.

642 Conclusion

643 Is it easier to do the same action compared to a different action? We used a stimulus-response
644 compatibility task with reversal learning phases in baboons to gain insight into the origins of automatic
645 imitation. We found no clear difference between the compatible and incompatible conditions,
646 although the monkeys tended to be slightly faster in the compatible condition. Remarkably, we found
647 that the learning rates between the two conditions were very similar, even after reversal. These results
648 contrast with what has been previously documented in humans and non-human animals and it may be
649 due to the use of different movements rather than different body parts. Using a ghost control, we
650 found that monkeys in the incompatible condition experienced important learning difficulties and that
651 monkeys in the compatible condition tended to be slower. These results suggest that the presence of
652 the demonstrator plays a role in eliciting a response but at the same time show that the presence of
653 the demonstrator is not essential, calling into question the relevance of this form of SRC task to study
654 automatic imitation. While the absence of any obvious differences between compatible and
655 incompatible responses makes it difficult to draw conclusions regarding the ontogeny of imitation, this
656 first study of conspecific “automatic imitation” in primates highlights the difficulty in observing
657 imitation in monkeys at all (whether automatic or not).

658 Imitation is pervasive in human social interactions from a young age fulfilling instrumental and
659 normative functions (Over & Carpenter, 2011), but in monkeys even the existing evidence of imitative
660 biases are not indicative of strong imitative tendencies (e.g., O’Sullivan, et al. 2017; Voelkl & Huber,
661 2000). Of course, the imitative faculty in primates may not represent one specific cognitive ability.
662 Instead, what has been defined generally as imitation could represent a range of cognitive skills
663 executed to different extents across species, context and development (Subiaul, 2016). The task we
664 used here, may not facilitate strong social learning effects in this species due to these animals’
665 evolutionary or individual histories. Here, we demonstrate the difficulty in observing the presence of
666 bodily imitation in monkeys adding to a literature of failed attempts at identifying imitations in
667 monkeys (see Visalberghi & Fragazy, 2001). In human research, evidence from dozens of studies
668 suggest that automatic imitation paradigms measure imitative tendencies (see Cracco et al. 2018) but
669 there is still some debate over how much automatic imitation is moderated by more general spatial
670 compatibility effect (e.g., Ramsey, 2018). Stimulus-response compatibility paradigms are a wonderful
671 means of determining compatibility effects under controlled experimental conditions, and here our
672 aim was to extend this paradigm to the question of imitation in monkeys. Across thousands of trials, we
673 observe little impact of imitative or spatial compatibility on task success demonstrating a clear
674 difference in the ways humans and monkeys perform in these paradigms.

675

676 Author contribution

677 Acknowledgment

678

679

680 Data availability:

681 Data analysed in this article are available at:

682 https://osf.io/x6krd/?view_only=c1ec2b441f1646ee8391fbf269bcd067

684 REFERENCES:

- 685 Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cockburn, A., Thornton, A., & Sheldon, B. C.
686 (2015). Experimentally induced innovations lead to persistent culture via conformity in wild birds.
687 *Nature*, 518(7540), 538–541. <https://doi.org/10.1038/nature13998>
688
- 689 Baayen R. (2008). Analyzing linguistic data. Cambridge, UK.
- 690 Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint
691 arXiv:1506.04967.
- 692 Bates D., Mächler M., Bolker B., Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4.
693 *Journal of Statistical Software* . 67(1):1-48
- 694 Berg E. A. (1948). A SIMPLE OBJECTIVE TECHNIQUE FOR MEASURING FLEXIBILITY IN THINKING. *The*
695 *Journal of General Psychology*, 39: 15-22
- 696 Bien N., Roebroek A., Goebel R., & Sack A. T. (2009). The Brain's Intention to Imitate: The
697 Neurobiology of Intentional versus Automatic Imitation. *Cerebral Cortex*, 19(10), 2338– 2351.
- 698 Blandin, Y., & Proteau, L. (2000). On the cognitive basis of observational learning: Development of
699 mechanisms for the detection and correction of errors. *Quarterly Journal of Experimental Psychology*
700 *Section A: Human Experimental Psychology*, 53(3), 846–867. <https://doi.org/10.1080/713755917>
- 701 Bonté E., Flemming T., Fagot J. (2011). Executive control of perceptual features and abstract relations
702 by baboons (*Papio papio*). *Behavioral Brain Research*. 222, 176–182
- 703 Bertenthal B. I., Longo M. R., & Kosobud A. (2006). Imitative response tendencies following observation
704 of intransitive actions. *Journal of Experimental Psychology Human Perception and Performance*, 32(2),
705 210–225.
- 706 Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence
707 problem? *Trends in Cognitive Sciences*, 9(10), 489–495. <https://doi.org/10.1016/j.tics.2005.08.007>
- 708 Brass M., Zysset S., von Cramon D. Y., (2001). The inhibition of imitative response
709 tendencies. *NeuroImage* 14, 1416–1423
- 710 Byrne, R. W. (2002). Imitation of novel complex actions: what does the evidence from animals mean?
711 *Advances in the Study of Behavior*, 31, 77–104. [https://doi.org/10.1016/S0065-3454\(02\)80006-7](https://doi.org/10.1016/S0065-3454(02)80006-7)
- 712 Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., & Heyes, C. (2008). Through the looking glass:
713 Counter-mirror activation following incompatible sensorimotor learning. *The European Journal of*
714 *Neuroscience*, 28(6), 1208–1215. <https://doi.org/10.1111/j.1460-9568.2008.06419.x>
- 715 Catmur, C., Mars, R. B., Rushworth, M. F., & Heyes, C. (2011). Making Mirrors: Premotor cortex
716 stimulation enhances mirror and counter-mirror motor facilitation. *Journal of Cognitive Neuroscience*,
717 23(9), 2352–2362.

718 Catmur C., Walsh V., & Heyes C. (2009). Associative sequence learning: The role of experience in the
719 development of imitation and the mirror system. *Philosophical Transactions of the Royal Society B:*
720 *Biological Sciences*, 364(1528): 2369–2380.

721 Claidière N., Gullstrand J., Latouche A., Fagot J. (2015). Using automated learning devices for Monkeys
722 (ALDM) to study social networks. *Behavior research methods*, 49: 24-34

723 Cracco E., Bardi L., Desmet C., Genschow O., Rigoni D., De Coster L., Radkova I., Deschrijver E., Brass
724 M., (2018). Automatic Imitation : A Meta-Analysis. *Psychological Bulletin*, 144(5)Cracco, E., & Brass, M.
725 (2017). Automatic Imitation of Multiple Agents: Simultaneous or Random Representation? *Journal of*
726 *Experimental Psychology: Human Perception and Performance*. 44(5)

727 Cracco, E., De Coster, L., Andres, M., Brass, M. (2016). Mirroring multiple agents: Motor resonance
728 during action observation is modulated by the number of agents. *Social Cognitive and Affective*
729 *Neuroscience*, 11(9), 1422–1427. <https://doi.org/10.1093/scan/nsw059>

730 Eimer M., (1995), Stimulus-Response Compatibility and Automatic Response Activation: Evidence
731 From Psychological Studies, *Journal of Experimental Psychology: Human Perception and Performances*
732 21(4): 837-854

733 Fagot, J., Gullstrand, J., Kemp, C., Defilles, C., & Mekaouche, M. (2014). Effects of freely accessible
734 computerized test systems on the spontaneous behaviors and stress level of Guinea baboons (*Papio*
735 *papio*). *American Journal of Primatology*, 76(1), 56-64. doi:10.1002/ajp.22193

736 Fagot, J., Marzouki, Y., Huguet, P., Gullstrand, J., & Claidière, N. (2015). Assessment of Social Cognition
737 in Non-human Primates Using a Network of Computerized Automated Learning Device (ALDM) Test
738 Systems. *JoVE*(99), e52798. doi:doi:10.3791/52798

739 Fagot J., Paleressompouille D., Automatic testing of cognitive performance in baboons maintained in
740 social groups (2009). *Behavioral Research Methods*, 41: 396-404

741 Fischer, J., Kopp, G. H., Dal Pesco, F., Goffe, A., Hammerschmidt, K., Kalbitzer, U., Klapproth, M., Maciej,
742 P., Ndao, I., Patzelt, A., & Zinner, D. (2017). Charting the neglected West: The social system of Guinea
743 baboons. *American Journal of Physical Anthropology*, 162(October 2016), 15–31.
744 <https://doi.org/10.1002/ajpa.23144>

745 Ferrucci L., Nougaret S., Brunamonti E., Genovesio A., (2019). Effects of reward size and context on
746 learning in macaques monkeys. *Behavioral Brain Research*, 372:111983

747 Gillmeister H., Catmur C., Liepelt R., Brass M., & Heyes C. (2008). Experience-based priming of body
748 parts: A study of action imitation. *Brain Research*, 1217: 157–170.

749 Heyes C. (2011). *Automatic Imitation*. *Psychological bulletin* 137(3) :463–483.

750 Heyes, C. M. (2018). *Cognitive Gadgets: The Cultural Evolution of Thinking*. Harvard University Press.

751 Heyes C., Bird, G., Johnson H., & Haggard P. (2005). Experience modulates automatic imitation.
752 *Cognitive Brain Research* 22: 233–240.

753 Heyes C. M., & Ray E. D. (2000). What Is the Significance of Imitation in Animals? *Advances in the Study*
754 *of Behavior*, 29(C): 215–245.

755 Hopewell, L. J., Leaver, L. A., Lea, S. E. G., & Wills, A. J. (2010). Grey squirrels (*Sciurus carolinensis*) show
756 a feature-negative effect specific to social learning. *Animal Cognition*, 13(2), 219–227.
757 <https://doi.org/10.1007/s10071-009-0259-3>

758 Hopper L. M. (2010). “Ghost” experiments and the dissection of social learning in humans and animals.
759 *Biological Reviews*, 85(4): 685–701.

760 Huang C.T., Charman T., (2005). Gradations of Emulation Learning in Infants’ Imitation of Actions on
761 Objects. *Journal of Experimental Child Psychology*, 92(3): 276-302

762 McKyton A., Ben-Zion I., & Zohary E. (2018). Lack of Automatic Imitation in Newly Sighted Individuals.
763 *Psychological Science*, 29(2): 304–310.

764 Meltzoff A. N., & Moore M. K. (1977). Imitation of Facial and Manual Gestures by Human Neonates.
765 *Science*, 198(4312): 75–78.

766 Meltzoff, A. N., & Moore, M. (1997). Explaining facial imitation: A theoretical model. *Early Development*
767 *and Parenting*, 6, 179-192. [https://doi.org/10.1002/\(SICI\)1099-0917\(199709/12\)6:3/4<179::AID-](https://doi.org/10.1002/(SICI)1099-0917(199709/12)6:3/4<179::AID-EDP157>3.0.CO;2-R)
768 [EDP157>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-0917(199709/12)6:3/4<179::AID-EDP157>3.0.CO;2-R)

769 Meltzoff A. N., Murray L., Simpson E., Heimann., Nagy E., Nadel J., Pedersen E. J., Brooks R., Messinger
770 D. S., Duncan De Pascalis L.L., Subiaul F., Paukner A., Ferrari P. F., (2017). Re-examination of

771 Mui R., Haselgrove M., Pearce J., & Heyes C. (2008). *Automatic imitation in budgerigars. Proceedings*
772 *of the royal society B: Biological Sciences*, 275 (1654): 2547–2553.

773 Nehaniv, C. L., & Dautenhahn, K. (2002). The Correspondence Problem. In C. L. Nehaniv & K.
774 Dautenhahn (Eds.), *Imitation in animals and artifacts* (pp. 41–61). MIT Press.

775 Nougaret S., Ferrucci L., Genovesio A., (2019). Role of the social actor during social interaction and
776 learning in human-monkey paradigms. *Neurosciences and Biobehavioral Reviews*, 102: 242-250

777 Obhi S. S., & Hogeveen J. (2010). Incidental action observation modulates muscle activity. *Experimental*
778 *Brain Research*, 203(2), 427–435.

779 Oostenbroek J., Suddendorf T., Nielsen M., Redshaw J., Kennedy-Costantini S., Mengersen Davis, J. T.,
780 Clark S., Slaughter V., (2016). Comprehensive Longitudinal Study Challenges the Existence of Neonatal
781 Imitation in Humans. *Current Biology*, 26(10)

782 O’Sullivan, E. P., Bijvoet-van den Berg, S., & Caldwell, C. A. (2018). Automatic imitation effects are
783 influenced by experience of synchronous action in children. *Journal of Experimental Child Psychology*,
784 171, 113–130. <https://doi.org/10.1016/j.jecp.2018.01.013>

785 O’Sullivan, E. P., & Caldwell, C. A. (2018). Imitation. In *Encyclopedia of Animal Cognition and Behavior*.
786

787 O’Sullivan E. P., Claidière N., & Caldwell C. A. (2017). Action-Matching Biases in Monkeys (*Sapajus spp*
788 .) in a Stimulus – Response Compatibility Task : Evaluating Experience-Dependent Malleability. *Journal*
789 *of Comparative Psychology*, 131(4): 337–347.

790

791 Otte E., Habel U., Schulte-Rüther M., Konrad K., Koch I., (2010). Interference in simultaneously
792 perceiving and producing facial expressions-Evidence from electromyography. *Neuropsychologia*,
793 49(1):124-30

794 Over, H., & Carpenter, M. (2011). Putting the social into social learning: Explaining both selectivity and
795 fidelity in children's copying behavior. *Journal of Comparative Psychology*, 126(2), 182–192.
796 <https://doi.org/10.1037/a0024555>

797 Press C., Bird G., Flach R., Heyes C., (2005). Robotic movement elicits automatic imitation. *Cognitive*
798 *Brain Research*, 25 (3) : 632-640

799 Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic
800 imitation of robotic action. *Proceedings of the Royal Society B: Biological Sciences*, 274(1625), 2509–
801 2514. <https://doi.org/10.1098/rspb.2007.0774>
802

803 Prinz, W. (1997). Perception and action planning. *The European Journal of Cognitive Psychology*,
804 9(2), 129–154. <https://doi.org/10.1080/713752551>
805

806 Ramsey, R. (2018). What are reaction time indices of automatic imitation
807 measuring?. *Consciousness And Cognition*, 65, 240-254.
808 <https://doi.org/10.1016/j.concog.2018.08.006>
809

810 Ray, E., & Heyes, C. (2011). Imitation in infancy: The wealth of the stimulus. *Developmental*
811 *Science*,14(1), 92–105. <https://doi.org/10.1111/j.1467-7687.2010.00961.x>
812

813 R Core Team (2020). R: A language and environment for statistical computing. R Foundation for
814 Statistical Computing, Vienna, Austria.
815

816 Range F., Huber L., & Heyes C. (2011). Automatic imitation in dogs. *Proceedings of the Royal Society B:*
817 *Biological Sciences*, 278(1703): 211–217.

818 Renner, E., Kean, D., Atkinson, M., & Caldwell, C. A. (2021). The use of individual, social, and animated
819 cue information by capuchin monkeys and children in a touchscreen task. *Scientific Reports*, 11(1), 1–
820 11. <https://doi.org/10.1038/s41598-020-80221-4>

821 Renner, E., Patterson, E. M., & Subiaul, F. (2020). Specialization in the vicarious learning of novel
822 arbitrary sequences in humans but not orangutans. *Philosophical transactions of the Royal Society of*
823 *London. Series B, Biological sciences*, 375(1805), 20190442. <https://doi.org/10.1098/rstb.2019.0442>

824 Rumiati, R. I., & Tessari, A. (2002). Imitation of novel and well-known actions: The role of short-term
825 memory. *Experimental Brain Research*, 142(3), 425–433. <https://doi.org/10.1007/s00221-001-0956-x>

826 Rumiati R. I., Tessari A. (2007). Automatic and strategic effects in human imitation. In *Sensorimotor*
827 *foundations of human cognition: attention and performance* (eds Haggard P., Kawato, Rossetti I.),
828 vol. XXII Oxford, UK: Oxford University Press

829 Stürmer B., Aschersleben G., & Prinz W. (2000). Correspondence Effects with Manual Gestures and
830 Postures: A Study of Imitation. *Journal of Experimental Psychology: Human Perception and*
831 *Performance*, 26(6):1746–1759.

832 Subiaul, F. (2010). Dissecting the imitation faculty: the multiple imitation mechanisms (MIM)
833 hypothesis. *Behavioural Processes*, 83(2), 222–234. <https://doi.org/10.1016/j.beproc.2009.12.002>
834

835 Subiaul, F. (2016). What's Special about Human Imitation? A Comparison with Enculturated Apes.
836 *Behavioral Sciences*, 6(3), 13. <https://doi.org/10.3390/bs6030013>
837

838 Subiaul, F., Cantlon, J. F., Holloway, R. L., & Terrace, H. S. (2004). Cognitive imitation in rhesus
839 macaques. *Science*, 305(5682), 407–410. <https://doi.org/10.1126/science.1099136>
840

841 Tennie C., Call J., Tomasello M., (2006). Push or Pull: Imitation vs. Emulation in Great Apes and Human
842 Children. *Ethology*, 112(12):1159-1169

843 Thompson D. E., Russell J., (2004). The Ghost Condition: Imitation Versus Emulation in Young Children's
844 Observational Learning. *Developmental Psychology*, 40(5): 882-9

845 van de Waal, E., & Whiten, A. (2012). Spontaneous emergence, imitation and spread of alternative
846 foraging techniques among groups of vervet monkeys. *PLoS ONE*, 7(10), e47008.
847 <https://doi.org/10.1371/journal.pone.0047008>

848 Visalberghi, E., & Fragaszy, D. M. (2001). "Do Monkeys Ape?" — Ten Years After. In E. Dupoux (Ed.),
849 *Language, Brain, and Cognitive Development* (pp. 472–499). MIT Press.

850 Voelkl, B., & Huber, L. (2000). True imitation in marmosets. *Animal Behaviour*, 60(2), 195–202.
851 <https://doi.org/10.1006/anbe.2000.1457>

852 Wiggett, A. J., Hudson, M., Tipper, S. P., & Downing, P. E. (2011). Learning associations between
853 action and perception: effects of incompatible training on body part and spatial priming. *Brain and*
854 *Cognition*, 76(1), 87–96. <https://doi.org/10.1016/j.bandc.2011.02.014>
855

856

857