1 **Stereotypic horses *(Equus caballus)* are not cognitively impaired**

2

3 S. Briefer Freymond[1], A. Ruet[1], M. Grivaz[1], C. Fuentes[1], K. Zuberbühler[2,4], I. Bachmann[1],

4 E. F. Briefer[3]

5 [1] Agroscope, Swiss National Stud Farm, Les Longs-Prés, 1580 Avenches, Switzerland

6 [2] Institute of Biology, University of Neuchatel, rue Emile-Argand 11, 2000 Neuchâtel,

7 Switzerland

8 [3] Institute of Agricultural Sciences, ETH Zürich, Universitätsstrasse 2, 8092 Zürich,

9 Switzerland

10 [4] School of Psychology & Neuroscience, University of St Andrews, Scotland

11

12 **Corresponding author:**

13 Sabrina Briefer Freymond; Email: sabrina.briefer@agroscope.admin.ch; Tel: +41 (0)58

14 484 61 01; Address: Swiss National Stud Farm SNSTF, Les Longs Prés, CH-1580

15 Avenches. Switzerland

16

17 **ABSTRACT**

18

19 Stereotypies in animal behaviour are thought to be an interaction between genetic

20 predisposition and sub-optimal housing conditions. In domestic horses, a well-studied

21 stereotypy is crib-biting, an abnormal behaviour that appears to help individuals to cope

22 with stressful situations. One prominent hypothesis states that animals affected by

23 stereotypies are cognitively less flexible compared to healthy controls, due to

24 sensitization of a specific brain area, the *basal ganglia*. The aim of this study was to test

25 this hypothesis, by using a cognitive task, reversal learning, which has been used as a

26 diagnostic for *basal ganglia* dysfunction, in crib-biting and healthy controls. The

27 procedure consisted of exposing subjects to four learning tasks; first and second

28 acquisition, and their reversals. For each task, we measured the number of trials to reach

29 criterion and heart rate and heart rate variability. Importantly, we did not try to prevent

30 crib-biters from executing their stereotypic behaviour. We found that the first reversal

31 learning task required the largest number of trials, confirming its challenging nature.

32 Interestingly, the second reversal learning task required significantly fewer trials to reach

33 criterion, suggesting generalisation learning. However, we did not find any performance

34 differences across groups; both stereotypic and control animals required similar numbers

35 of trials and did not differ in their physiological responses. Our results thus challenge the

36 widely held belief that crib-biting horses, and stereotypic animals more generally, are

37 cognitively impaired. We conclude that cognitive underperformance may occur in

38 stereotypic horses if they are prevented from crib-biting to cope with experienced stress.

39

40 **Keywords:** crib-biting; basal ganglia; learning capacity

41 **INTRODUCTION**

42

43 Stereotypies exist in humans and non-human animals and have been defined as

44 repetitive, relatively invariant patterns of behaviour with no apparent goal or function

45 (Mason and Latham 2004). In animals, they seem to be artefacts of a captive

46 environment involving restricted and sub-optimal housing conditions (McGreevy et al.

47 1995). They exist in various forms, the most common forms being oral and locomotor

48 stereotypies (Mason and Rushen 2006). One classic ethological model of motivation

49 suggests that restrictive environments can prevent the animal from reaching the

50 "consummatory phase" of a behaviour (e.g. feeding) (Hughes and Duncan 1988). As a

51 consequence, a number of appetitive behaviours (e.g. locomotor action to access food)

52 are being attempted in an effort to reach unobtainable goal states (e.g. elevated blood

53 glucose). When the goal is available, its consummation results in a negative feedback

54 that decreases motivation. In contrast, in cases where the goal is unachievable, the lack

55 of consummation and resulting absence of negative feedback increases the level of

56 motivation to perform appetitive behaviours. If such frustration-inducing situations occur

57 repeatedly, stereotypies can develop (Mason and Rushen 2006; McBride and Parker

58 2015). Along similar lines of reasoning, it has been suggested that privation (as

59 experienced in captivity) increases the desire for rewards, which in turn enhances

60 appetitive behaviours such as locomotor or feeding behaviours (Spruijt et al. 2001).

61 Excessive use of appetitive behaviours, in turn, can lead to a loss of regulatory control,

62 which means that these behaviour will become environmentally irreversible (inflexible)

63 and can develop into stereotypies (Toates 2004).

64

65 Other studies have focussed on the impact of chronic stress on the brain, and in

66 particular, on parts of the basal ganglia (the mesoaccumbens dopaminergic system)

67 (Spruijt et al. 2001). The reasoning behind this research is that stressful negative

68 experiences lead to an increased tendency to seek rewards via a neurobiological process

69 of sensitisation (van der Harst et al. 2003), a possible common principle underlying

70 stereotypies in evolutionarily distant species such as mice and human (Cabib et al.

71 1998). For example, some research as shown an association between stressors,

72 stereotypy development and dopamine receptor function in the *basal ganglia* in rodents

73 (Cabib et al. 1998) and crib-biting horses, *Equus caballus* (McBride and Hemmings

74 2005), suggesting that the *basal ganglia* plays a general role in the development of

75 stereotypic behaviour (McBride et al. 2017). However, the exact causal relationship

76 between *basal ganglia* alterations and the emergence of stereotypic behaviour is

77 currently unknown (McBride and Hemmings 2005).

78

79 Each part of the basal ganglia plays a specific role during instrumental learning

80 (Dickinson 1985). The first stage ("acquisition") is mediated by the ventral and

81 dorsomedial striatum of the basal ganglia (McBride et al. 2017; Parker et al. 2008; Yin

82 and Knowlton 2006), the second stage ("action-outcome") by the dorsomedial striatum

83 (McBride et al. 2017; Parker et al. 2008; Yin and Knowlton 2006), and the final stage

84 ("habit formation") by a shift in activation from the dorsomedial towards the dorsolateral

85 striatum of the basal ganglia. To summarize, the ventral striatum and dorsomedial

86 striatum seem to be important in the acquisition and execution of goal-directed actions,

87 that is, in establishing the link between stimulus response and outcome (McBride et al.

88 2017; Yin et al. 2008). By contrast, the dorsolateral striatum seems to control

89 subsequent habit formation, independently of the response outcome (Yin and Knowlton

90 2006). Therefore, due to its primary functional role, dysfunctions in the basal ganglia

91 might lead to behavioural abnormalities linked to impaired action selection and also to

92 impairments in controlling instrumentally learned behaviour (McBride and Hemmings

93 2005; Vickery and Mason 2005).

94

95 Although there are strong reasons to assume that an impaired *basal ganglia* function is

96 related to the development of stereotypic behaviour, research is challenging for financial,

97 logistical, and ethical reasons, which has led to the development and use of indirect and

98 non-invasive methods (McBride et al. 2017). For instance, stereotypy levels (frequency of

99   stereotypy performance) have been shown to constitute strong predictors of the latency

100   to extinguish conditioned responses or of the tendency to inappropriately repeat

101   responses, both of which constitute indirect measures of *basal ganglia* dysfunction in

102   several species such as bears, *Ursus thibetanus* and *Helarctos malayanus* (Vickery and

103   Mason 2005), Orange-Wing Amazon Parrots, *Amazona amazonica* (Garner et al. 2003)

104   and bank voles, *Clethrionomys glareolus* (Garner and Mason 2002). In humans, poor

105   abilities to suppress learnt behaviour (perseveration) have been shown in autistic

106   patients who are prone to stereotypic behaviour (Boyd et al. 2009; Lopez et al. 2005).

107

108   Domesticated horses are subject to management practices that make them prone to

109   developing stereotypic behaviours. Understanding the nature of stereotypies and their

110   impact on learning abilities is therefore of considerable importance for horse owners.

111   Crib-biting, an oral stereotypy, is one of the most common forms of stereotypy in horses

112   (Luescher et al. 1991; Wickens and Heleski 2010). The performance of this behaviour

113   varies between horses in terms of the percentage of time occupied by the stereotypic

114   behaviour (Houpt 1993). Crib-biting has been linked to learning impairments in extinction

115   paradigms (Hemmings et al. 2007; Roberts et al. 2015). In particular, stereotypic horses

116   need more trials compared to healthy individuals before extinction of a previously learnt

117   action occurred, and this may be linked to a *basal ganglia* dysfunction. In one study,

118   crib-biting horses appeared to exhibit altered dopamine receptor sensitivity in the *basal*

119   *ganglia* (McBride and Hemmings 2005) due to a higher numbers of dopamine receptors in

120   the *ventral striatum* and lower numbers of receptors in the *dorsomedial striatum.* Since

121   the *dorsomedial striatum* mediates action-outcome learning, it is possible that crib-biting

122   horses are simply unable to maintain this type of learning and show an accelerated shift

123   from action-outcome learning to habit formation (Parker et al. 2009; Parker et al. 2008;

124   Roberts et al. 2015) and a reduced ability to learn about outcomes (Schwabe and Wolf

125   2011). Additionally, another study including many different kind of stereotypies in horses

126   (locomotor and oral) showed that stereotypic horses need more time to learn an

5

127 instrumental task (opening a chest by raising the lid using the nose) compared to non-

128 stereotypic horses (Hausberger et al. 2007).

129

130 Previous studies have found differences in the learning capacities of crib-biting and

131 control horses (Hemmings et al. 2007; Parker et al. 2009; Parker et al. 2008; Roberts et

132 al. 2015). Other work has shown that crib-biting horses appear to have alterations in the

133 dopaminergic system (McBride and Hemmings 2005). However, link between these

134 alterations and cognitive performances has remained unclear (Roberts et al. 2017).

135 Reversal learning paradigms are of particular relevance, as they have been used as a

136 diagnostic tool for dopaminergic dysfunction and as general measure of cognitive

137 flexibility in rodents, nonhuman primates and humans (Izquierdo et al. 2017; McBride et

138 al. 2017). To our knowledge, reversal learning has not been investigated in crib-biting

139 horses and has been shown to pose a challenge to this species when based on visual

140 cues (Brubaker and Udell 2016; Hothersall et al. 2010; Martin et al. 2006; McBride et al.

141 2017; Sappington et al. 1997; Voith 1975), unlike reversal learning tasks based on

142 spatial cues, which seem to be fairly easy for horses due to their ecological relevance

143 (e.g. finding natural food sources) (Brubaker and Udell 2016; Fiske and Potter 1979;

144 Martin et al. 2006; Voith 1975; Warren and Warren 1962). We therefore tested crib-

145 biting and control, non-stereotypic horses in two subsequent reversal learning tasks

146 based on visual cues. First, we predicted that crib-biting horses would need less trials

147 than controls to perform the first and second acquisition task, because they might be

148 more prone to habit learning than to response-outcome learning (Hemmings et al. 2007;

149 Parker et al. 2009; Parker et al. 2008; Roberts et al. 2015). Second, we predicted that

150 crib-biting horses would need more trials than controls to perform the reversal learning

151 tasks, suggesting learning disabilities, if they suffered from dopaminergic dysfunction. By

152 contrast, similar performances between crib-biting and control horses would suggest that

153 the stereotypic horses are not suffering from such a dysfunction. Importantly, we did not

154 try to prevent crib-biters from executing their stereotypic behaviour, based on our

155 previous finding that crib-biting reduces stress (Briefer Freymond et al. 2015) in order to

156 avoid the confounding influence of stress on learning (Schwabe and Wolf 2010).

157

158 **METHODS**

159

160 **Subjects and management conditions**

161 The study was carried out on six crib-biters ("CB") and seven control horses ("C") (N =

162 13) of various breeds, sexes (mares, geldings and stallions), and ages (10 to 25 years

163 old), housed in five different farms in Switzerland, between January and May 2016 (Table

164 1). Eight horses were privately owned, and five horses were owned by the Swiss National

165 Stud Farm. All the horses had been at their respective farms, for at least one year. To be

166 eligible for inclusion in the study, crib-biters were required to have demonstrated crib-

167 biting behaviour for a minimum of 1 year, as reported by their owners. All the crib-biters

168 eventually included in the study had been crib-biting for at least 4 years.Controls were

169 horses that had never been observed crib-biting or perform other kinds of stereotypies

170 (e.g., weaving or box-walking). All but two animals participated in a previous study,

171 which involved a spatial learning task (Briefer Freymond et al., in preparation; one crib-

172 biter, one control). Each crib-biter was matched with a control horse of similar breed

173 (except for one pair), sex, age, and housing conditions (individual or group, single box or

174 box with paddock, and if possible in the same farm) (Table 1). One supplementary

175 control horse was tested to lower the average age of the controls, which was originally

176 higher than the age of the crib-biters (final mean age (years old): controls = 17.6; crib-

177 biters = 13.5). Routine care was provided by the owner.

178

179 **Experimental design**

180 *Experimental protocol*

181 Before the start of the learning experiment, all the CB horses were filmed in their home

182 pen, while undisturbed, during 48-hour periods (excluding periods when horses were

183 ridden or in pasture), in order to assess their crib-biting frequency per hour (see in

184 *Stereotypy level* below and Table 1). For the learning experiment, each horse was led,

185 individually, to a delimited (8/10 m) familiar arena, after equipping it with a heart rate

186 monitor in its home pen (see in *Physiological measures* below). The arena was divided

187 into a waiting area and a test area, separated by a start rope. The learning apparatus

188 was placed at one end of the test area (Fig 1). The horses were filmed from the back

189 with a video camera fixed on a pole in order to score their behaviour (see in *Behavioural*

190 *measures* below). Two experimenters were present during the study sessions.

191 Experimenter 1 was located in the arena and handled the subjects (see in *Discrimination*

192 *procedure* below); Experimenter 2 located outside the arena was entering comments on

193 the collected physiological data (see in *Physiological measures* below) and preparing the

194 learning apparatus for the next trial (see in *Discrimination procedure* below).

195

196 *Apparatus*

197 Because horses are very skilled at reading subtle unintentional human behaviour

198 (Ringhofer and Yamamoto 2017), we built an apparatus which allowed the experimenters

199 to remain in the back of the area (Fig 1). Following Gabor and Gerken (2010), the

200 apparatus consisted of a wooden box measuring 1 m (height) X 1.6 m (length) X 0.4 m

201 (width) with two flaps (45 cm X 61 cm) on the front side (Fig 1). The horse could reach

202 the food through these flaps. To prevent the horses from using olfactory cues,

203 Experimenter 2 always filled both bowls situated at the back side of the apparatus with

204 food (20 g of commercial concentrate), even though only one side was rewarded each

205 time. This was achieved by closing the unrewarded flap automatically using an

206 electromagnet that could be activated with an infrared remote control. In addition, in

207 case of an incorrect choice, the positively reinforced flap was immediately closed from

208 the other side by activating the electromagnet in order to prevent the horse from being

209 rewarded for an incorrect choice. A vertical piece of wood was added in the middle of the

210 apparatus between the two flaps to better separate the two sides of the apparatus and

211 facilitate the scoring of the horse's choices.

212

213   The visual stimuli (see *Two-choice visual discrimination tasks* for more details) were

214   inserted on the front side of the apparatus inside a plastic window fixed on the wooden

215   boards. The stimuli consisted of sheets of laminated paper (21 cm X 29.5 cm) on which

216   either a black cross on top of a white background, a white cross on top of a black

217   background (first set of stimuli), a black circle on top of a white background, or a white

218   circle on top of a black background (second set of stimuli) were drawn. The same amount

219   of sheet area was covered by the cross (13.5 cm X 13.5 cm) and by the circle (13 cm in

220   diameter).
221
222

223   *Acclimation and pre-training (2 – 6 days)*

224   During the acclimation and pre-training phases, the horses were habituated to the

225   experimental arena and trained to move from the starting point to the apparatus and to

226   open the flaps in the absence of any stimuli. The horses were trained during two 10-min

227   sessions each day, for two to six days. The horses were first trained to touch a target (a

228   tennis ball fixed on a stick) with their noses and then the flaps on the apparatus, using a

229   shaping procedure. This shaping procedure is also called "successive approximations",

230   which consisted of reinforcing behaviours directed towards the desired response

231   (McGreevy 2010). The first step of the pre-training phase lasted until each horse was

232   acclimated to the apparatus and touched both flaps easily with its nose. The second step

233   consisted in shaping the horse to open the flaps, by rewarding it each time it pushed the

234   flaps. When the horse had learnt to open the flaps with its nose, Experimenter 1 led the

235   horse to the waiting area, and it was trained to go alone from the waiting area and to

236   open the flaps (third step). The pre-training lasted until the horses opened the flaps at

237   least five times on both sides of the apparatus.

238

239   *Discrimination procedure*

240   The discrimination procedure took place during the following three to seven weeks. The

241   horses were tested each Monday, Tuesday, Thursday and Friday. They were tested one

242   by one with two sessions per day (15-20 min), each containing between 20 and 23 trials,

243   with a break between the respective sessions of about 20 min, during which time another

244   horse from the same farm was tested. Before each trial, Experimenter 2 inserted the two

245   stimulus sheets into the plastic windows in a "pseudo-randomized" order (established a-

246   priori). This order ensured that each given stimulus was not presented for more than

247   three consecutive trials at the same position (left or right). After inserting the stimuli

248   sheets, the feeding bowls at the back side of the apparatus were filled and the

249   unrewarded flap was remotely blocked. During this time, Experimenter 1 led the horse to

250   the waiting area and released it after closing the waiting area with a rope. As soon as the

251   setting for the next trial was ready, Experimenter 1, who was blind to the correct

252   stimulus at the beginning of each session, opened the start rope while facing and looking

253   away from the horse and the test area. In case of a correct choice, the horse was led

254   back to the waiting area after reaching and eating the reward. In case of an incorrect

255   choice, the horse was led back to the waiting area and allowed to choose again with the

256   same arrangement of stimuli ("correction trial")(Flannery 1997). After three wrong

257   decisions, the horse was led to the correct stimulus where it could open the flap to reach

258   the reward (Flannery 1997). In this case, if necessary, Experimenter 1 pointed at the

259   correct flap with the hand. Each trial was limited to a 90-s duration, during which all

260   horses made a choice (i.e. no trial had to be stopped before a choice was made).

261

262   *Two-choice visual discrimination task*

263   The learning procedure consisted of four different phases ("Phase"). At the start of the

264   learning phase 1 ("Acq1"), two initial stimuli (a black versus a white cross or a black

265   versus a white circle) were presented to the horses. The learning criterion was set at six

266   consecutive correct responses in one learning session (i.e. probability of doing this by

267   chance = 0.02) (McBride et al. 2016). Once the horse had reached the learning criterion

268   for Acq1, the colour of the stimuli were reversed ("Rev1") and the next session started.

269   Once the horses had reached the learning criterion for Rev1, they were then tested with

270   a second set with novel stimuli (i.e. a black versus a white circle or a black versus a

271   white cross; second acquisition phase, "Acq2"). Acq 2 was then followed by a second,

final reversal ("Rev2"), after reaching the learning criterion for Acq2. The two reversals Rev1 and Rev2 consisted in rewarding the previously unrewarded stimulus (during Acq1 and Acq2, respectively) and vice-versa, while Acq2 consisted in presenting circles to horses who received crosses during Acq1 and Rev1, and vice-versa, with or without changing the colour compared to Rev1 (Table 1). The rewarding stimuli for the different phases were assigned randomly to each horse before the study. For three of the six crib-biters and four of seven control horses, there was a colour change between Rev1 and Acq2 (Table 1).

In previous studies, it was shown that extended sessions of concentrated training could lead to a lack of motivation or to inappropriate and inefficient learning behaviour (McCall 1990; Rubin et al. 1980). To ensure that horses stay motivated, we decided, based on some preliminary tests ($N = 7$ non-stereotypic horses, not used in this study) to perform two sessions of 20 trials per day, and also to reward horses in cases when they chose the incorrect flap three times in a row (and thus did not obtain any reward).

Response measures

Stereotypy level

We scored the number of crib-biting events over time from the video recordings collected over 48 hours before the start of the experiment (see in *Experimental protocol* above), in order to assess the stereotypy level of the crib-biters at the time of the study. This score was converted into a frequency of crib-biting events per hour per horse. Based on these frequencies, we made three groups of crib-biters for the analysis: "S", strong crib-biters (58.37–65.76 times per hour, N= 3 horses); "M", medium crib-biter (25.03 times per hour, $N = 1$ horse); and "L" low crib-biters (1.09–9.06 times per hour; $N = 2$ horses).

Behavioural measures

All the learning tests were video recorded using a GoPro HERO3 to control for eventual errors of scoring for all the following behaviours, which were directly scored during the

301 tests by Experimenter 2: the choice of stimulus (correct or incorrect), the side chosen

302 (right or left), and crib-biting events. Crib-biting was defined as instances when the horse

303 grasped the top of the apparatus with its incisors, pulled back, contracted the neck

304 muscles and drew air into its oesophagus, emitting an audible grunt (McGreevy et al.

305 1995). Three stereotypic horses did crib-bite during at least one trial while others never

306 did so (Table 2).

307

308 *Physiological measures*

309 Before bringing a horse into the testing arena, it was equipped with a wireless heart-rate

310 monitor (MLE120X Bioharness Telemetry System, Zephyr) fixed on a specific girth.

311 During the acclimation and pre-training phases, the horses did also wear the girth for

312 habituation, but without recording any data. During the tests, we collected the ECG trace

313 continuously, allowing us to obtain the heart rate (HR) and the root mean square of

314 successive inter-beat interval differences (RMSSD) as indicator of the physiological stress

315 level of the subjects (von Borell et al. 2007). ECG gel was applied on the electrodes

316 before each use. The data were transmitted in real time to a laptop using AcqKnowledge

317 software v.7.2 (Biopac), and stored for later analyses. This allowed Experimenter 2 to

318 add live comments during the visual discrimination task indicating when each session and

319 each trial started. This enabled us to measure the physiological parameters precisely for

320 each phase and each trial. We extracted HR and RMSSD from good-quality sections with

321 clearly visible heartbeats on the ECG trace. We divided each session in five parts of equal

322 duration, and analysed, when possible, three segments of 10 s each per part (at the

323 beginning, middle, and end). We checked visually that the software was tracking the

324 heartbeats properly, and extracted HR and the inter-heartbeat (RR) intervals (ms). RR

325 intervals were then used to calculate RMSSD (ms). We then calculated an average value

326 per phase ("Acq1", "Rev1", "Acq2", "Rev2") for HR and for RMSSD. The total duration

327 over which we were able to extract HR and RMSSD for the analyses was comparable

328 between crib-biters and controls ("CB" group; means ± SD = 41.20 ± 18.59 s and "C"

329 group = 48.28 ± 10.56 s).

330

Statistical analysis

All our data were analyzed using generalised linear mixed models (GLMMs) or linear

mixed-effect models (LMMs), in R 3.0.2, as described below. The list of the fixed, control

factors and interactions terms included in each model are given in Tables 3 and 4.

Because many studies have shown that stereotypy levels (frequency of stereotypy

performance over time) can be strong predictors of learning abilities (Garner and Mason

2002; Garner et al. 2003; Vickery and Mason 2005), we first ran the models described

below to investigate differences between crib-biters ("CB") and control horses ("C"),

"GroupCB-C", and then reran the same models to investigate, this time, differences

between the four groups defined by the frequency of crib-biting events per hour over 48h

(see *Stereotypy level* above) as follows: "S" strong crib-biters, "M" medium crib-biter, "L"

low-frequency crib-biters and "C" control; "GroupSMLC" (Tables 3 and 4).

We first tested for group differences (GroupCB-C, or GroupSMLC, Tables 3 and 4) in the

number of training sessions the horses needed for the acclimation and pre-training

phases ("Session"). The fixed, control factors and interactions terms included in this

GLMM are described in Tables 3 and 4. The place where the horses were housed was

added as a random factor. Two-by-two comparisons between the different groups of the

factor GroupSMLC ("S", "M", "L", and "C") were then carried out using Tukey post-hoc

tests (function glht, package multcomp in R, multiple comparisons of means).

To investigate learning performances, we then tested for group differences (GroupCB-C,

or GroupSMLC, Tables 3 and 4) in the frequency of correct choices (i.e. number of correct

choices divided by the total number of trials; "Correct") during the last session of

acquisition for the first ("Acq1") and second ("Acq2") sets of learning tasks, and during

the first session of reversal for the first ("Rev1") and second ("Rev2") sets of learning

tasks. The fixed, control factors and interaction terms included in these LMMs are

described in Tables 3 and 4. The horse identity nested within the place where the horses

13

359 were housed was included as a random factor. Two-by-two comparisons between the

360 different learning phases (between "Acq1" and "Rev1", and between "Acq2" and "Rev2")

361 were then carried out using Tukey post-hoc tests (function glht, package multcomp in R,

362 multiple comparisons of means).

363

364 Finally, to compare the learning abilities (i.e. number of trials needed to reach the learning

365 criterion for each learning phase, "Trial") and the physiological stress level of GroupCB-C

366 and of GroupSMLC, we ran two separate sets of LMMs with Trial, HR or RMSSD as response

367 variables. The first set of LMMs was aimed at testing the learning abilities and stress levels

368 of crib-biters and controls during each learning phase (Acq1, Rev1, Acq2 and Rev2). In

369 this set, we investigated group differences (CB-C or SMLC) in Trial, HR or RMSSD during

370 the four learning phases. The fixed, control factors and interaction terms included in these

371 LMMs are described in Tables 3 and 4. Because of the small sample size, we additionally

372 carried out a power analysis for the effect of Group CB-C and GroupSLMC on Trial, HR and

373 RMSSD in order to calculate if the power of our analysis was large enough (pwr.f2 function,

374 pwr library; in R 3.0.2). Two-by-two comparisons between the different learning phases

375 were then carried out using Tukey post-hoc tests (function glht, package multcomp in R,

376 multiple comparisons of means). The second set of LMMs was aimed at testing the effect

377 of the change in the colour of the signal that some horses experienced between Rev1 and

378 Acq2 on Trial, HR or RMSSD (Table 1). Indeed, this change in colour (hereafter "Colour",

379 change in colour "Y", no change of colour "N") could have also been perceived as a reversal

380 by the horses. Since the factors Phase and Colour are correlated, we tested their effects

381 on the response variables in different sets of models. In the set used to test the effect of

382 the change in colour, the same fixed and control factors were included as in the set used

383 to test the effect of the phase (Table 3 and 4), except for the fixed factor Phase, which was

384 replaced by Colour. In this second set, only the data for Rev1 and Acq2 were included,

385 because we were interested specifically in the colour change or not between Rev1 and

386 Acq2. Because of the small sample size, we again carried out a power analysis for the effect

387 of GroupCB-C and GroupSLMC on Trial, HR and RMSSD in order to calculate if the power

388   of our analysis was large enough (pwr.f2 function, pwr library; in R 3.0.2). For all these

389   models (first and second sets), the horse identity nested within the place where the horses

390   were housed was included as a random factor.

391

392   For all models described above, the residuals were checked graphically for normal

393   distribution and homoscedasticity. To satisfy model assumptions, we used a square-root

394   transformation for Trial, and a cube-root transformation for RMSSD. All the resulting

395   parameters satisfying model assumptions were then entered into linear mixed-effects

396   models fit with Gaussian family distribution and identity link function (lme function, nlme

397   library, in R 3.0.2). Session did not meet the statistical assumptions despite

398   transformation. It was thus transformed to binomial data as follows; value equal or

399   higher than median = 1 or value lower than median = 0. This parameter transformed to

400   binomial data was input into a generalized linear mixed model fit with binomial family

401   distribution and logit link function (glmer function, lmerTest library, in R 3.0.2). For all

402   models, we used a standard model simplification procedure by removing each non-

403   significant term, until the deletion caused a reduction in goodness of fit (at which point,

404   the term was left in the model). We assessed the statistical significance of each factor by

405   comparing the model with and without the factor included using likelihood-ratio tests

406   (LRT). The significance level of the factors was set at α = 0.05.

407

408   **RESULTS**

409

410   *Acclimation and pre-training*

411   Crib-biters required significantly more sessions ("Session") to fulfil the learning criterion

412   before starting the discrimination learning task than controls (CB: 6.83 ± 2.99 sessions;

413   C: 4.57 ± 0.79 sessions; GLMM: Effect of GroupCB-C on Session; χ^2 = 7.29, df = 1, p =

414   0.007). There was also a significant effect of the four groups defined by the frequency of

415   crib-biting events per hour over 48h ("GroupSLMC": "S" strong crib-biters, "M" medium

416   crib-biter, "L" low-frequency crib-biters and "C" control) on sessions (GLMM: Effect of

417   GroupSMLC on Session, $\chi^2 = 11.94$, df = 4, p = 0.008). However, further two-by-two

418   comparisons did not show any significant difference between the four groups "S", "M",

419   "L" and "C" in the number of sessions needed to fulfil the learning criterion (p ≥ 0.98 for

420   all).

421

422   *Learning performance*

423   All 13 animals completed the four learning phases ("Phase"; "Acq1", "Rev1", "Acq2",

424   "Rev2"). There was a significant effect of Phase on the frequency of correct choices

425   ("Correct") in the last session of acquisition for the first and second sets of learning

426   tasks, and the first session of reversal for the first and second sets of learning tasks

427   (LMM: effect of Phase on Correct, $F_{3,36} = 15.51$, p < 0.0001; Fig 2). Further post-hoc

428   analyses showed a significant drop in the frequency of correct choices per session

429   between the last session of Acq1 (mean ± SD = 0.72 ± 0.18) and the first session of

430   Rev1 (Rev1 = 0.40 ± 0.17; Multiple comparisons of means; effect of Phase Acq1 versus

431   Rev1 on Correct; Z = –4.80, N = 13, p < 0.0001; Fig 2). There was also a significant

432   drop in the frequency of correct choices per session between the last session of Acq2

433   (0.76 ± 0.08) and the first session of Rev2 (0.44 ± 0.22; Multiple comparisons of means:

434   effect of Phase Acq2 versus Rev2 on Correct; Z = –4.80, N = 13, p < 0.0001, Fig 2). On

435   the other hand, there was no effect of group CB versus C on the frequency of correct

436   choice (LMM: effect of GroupCB-C on Correct; $F_{1,7} = 1.77$, p = 0.31) nor of GroupSMLC

437   (LMM: effect of GroupSMLC on Correct; $F_{3,5} = 0.59$, p = 0.65).

438

439   *Learning capacities*

440   There was neither effect of the two groups CB-C (LMM: effect of GroupCB-C on Trial; $F_{1,7}$

441   = 1.77, p = 0.23; Fig 3), nor of GroupSMLC (LMM: effect of GroupSMLC on Trial; $F_{3,5}$ =

442   0.962, p = 0.48) on the number of trials per phase needed to reach the learning criterion

443   ("Trial"). However, there was a significant difference between phases (Acq1; Rev1; Acq2;

444   Rev2) in Trial for all horses (LMM: effect of Phase on Trial; $F_{3,36} = 5.05$, p = 0.005; Fig 3).

445   The number of trials needed until the learning criterion was reached are shown in Fig 3.

446   Post-hoc comparisons showed that all horses needed significantly more trials for Rev1

447   than for Acq1 (Multiple comparisons of means: $Z = 3.64$, $N = 13$, $p = 0.002$; Fig 3), and

448   more trials for Rev1 than for Acq2 (Multiple comparisons of means: $Z = 2.92$, $N = 13$, p

449    $= 0.018$; Fig 3) and for Rev2 (Multiple comparisons of means: $Z = -2.65$, $N = 13$, $p =$

450   0.041 Fig 3). The other two-by-two comparisons were not significant ($p \geq 0.76$ for all).

451   In addition, considering only the first reversal and second acquisition, horses needed

452   more trials when there was a colour change ("Colour") (mean ± SD = 169.86 ± 95.12

453   trials) than when there was no change in colour (100.67 ± 87.13 trials) between Rev1

454   and Acq2 (LMM: effect of Colour on Trials; $F_{1,12} = 6.603$, $p = 0.025$).

455

456   For all the LMMs carried out on Trial, neither the interaction between Phase and

457   GroupCB-C or GroupSMLC, nor the sex, the age, the person leading the horse (the two

458   different persons), the type of signal (cross or circle), or the colour of the signal (black or

459   white) had a significant effect. These terms were thus removed during model selection.

460   Power analyses conducted on non-significant models revealed that the various LMMs

461   tested the effect of GroupCB-C and GroupSMLC on Trial had a power ≥ 0.94, suggesting

462   that a larger sample size would not have led to a significant result.

463

464   *Physiological parameters*

465   *Heart rate (HR)*

466   There was no difference in HR between the different learning phases (LMM: effect of

467   Phase on HR; $F_{3,35} = 1.03$, $p = 0.39$), nor between the two groups CB-C (LMM: effect of

468   GroupCB-C on HR; $F_{1,7} = 2.09$, $p = 0.20$), and no effect of GroupSMLC (LMM: effect of

469   GroupSMLC on HR; $F_{3,5} = 1.88$, $p = 0.25$). In addition, considering only the first reversal

470   and second acquisition, Colour had no effect on HR (LMM: effect of Colour on HR; $F_{1,12} =$

471   0.30, $p = 0.59$).

472

473   *Root mean square of successive inter-beat interval differences (RMSSD)*

474   There was no difference in RMSSD between the different learning phases (LMM: effect of

475   Phase on RMSSD; $F_{3,32} = 0.97$, $p = 0.41$), nor between the two groups CB-C (LMM: effect

476   of GroupCB-C on RMSSD; $F_{1,7} = 0.05$, $p = 0.84$), and no effect of GroupSMLC on RMSSD

477   (LMM: effect of GroupSMLC on RMSSD; $F_{3,5} = 0.91$, $p = 0.50$). In addition, Colour had no

478   effect on RMSSD (LMM: effect of Colour on RMSSD; $F_{1,12} = 1.35$, $p = 0.27$).

479

480   For all the LMMs carried out on HR or RMSSD, neither the interaction between Phase and

481   Group, nor the interaction between Colour and Group, nor the sex, the age, the person

482   leading the horse (two different persons), the signal (cross or circle), or the colour of the

483   signal had a significant effect. These terms were thus removed during model selection.

484   Power analyses conducted on non-significant models revealed that the various LMMs

485   tested the effect of GroupCB-C and GroupSLMC on HR or RMSSD had a power ≥ 0.87,

486   suggesting that a larger sample size would not have led to a significant result.

487

DISCUSSION

In this study, we used a reversal learning task, which has been used as a diagnostic for basal ganglia dysfunction, to compare the learning performances of crib-biting and control horses. According to our results, there is no indication that crib-biters suffer from such a dysfunction. Except for the acclimation phases, which took longer for crib-biters compared to the controls to achieve, we did not find any differences between crib-biters and control horses in the number of trials necessary to reach the learning criterion in any phase of the experiment. In fact, all horses reached the learning criterion and performed the two reversals. Interestingly, they also performed the second reversal in fewer trials compared to the first one, suggesting that they learned to learn. Unlike in other studies that found that crib-biting horses have altered learning abilities compared to other horses (Hemmings et al. 2007; Parker et al. 2009; Parker et al. 2008; Roberts et al. 2015), our subjects had the opportunity to crib-bite during the experiment, and hence potentially to reduce their stress level (Briefer Freymond et al. 2015), as shown by the absence of differences between the two groups in the physiological parameters that we measured. We could therefore suggest that previous research on learning performance could be the result of differences in stress levels experienced by crib-biters and control horses, although other studies did not collect physiological measures of stress.

Acclimation and pre-training

In this study, crib-biters needed a longer time than the controls to be acclimated to the learning apparatus, and to attain the conditions required to start the discrimination procedure (i.e. pushing the flaps alone five times on both sides of the apparatus without any intervention of the experimenter). This could be explained by the fact that crib-biters seem to be more stress sensitive (Briefer Freymond et al. 2015), and might thus need more time to be acclimated to a new situation. However, we did not collect any physiological indicators of stress during the pre-training phase, because in this phase, the movement of the horses was not standardised, unlike during the learning phase.

19

517   Indeed, it is recommended that only measures made during times of similar behavioural

518   pattern should be compared (von Borell et al. 2007). Future studies that additionally

519   measure stress parameters during habituation could inform us on the stress levels of

520   crib-biters in such situations.

521

522   *Learning capacities*

523   All the horses (crib-biters and controls) in our study performed the two reversal tasks,

524   and needed significantly more trials to reach the learning criterion for the first reversal

525   ("Rev1") than for the other phases (first acquisition, "Acq1", second acquisition, "Acq2",

526   and second reversal, "Rev2"). In addition, the significant drop that we observed in the

527   number of correct responses between the acquisition phases and their following

528   respective reversals suggests that learning had taken place (McBride and Parker 2015).

529   Although few studies suggest that horses possess the ability to perform reversal learning

530   based only on visual cues (Sappington et al. 1997; Voith 1975), this task seems to be

531   more challenging, and in some cases not achievable, compared to reversal learning

532   based on spatial cues (Hothersall et al. 2010; Martin et al. 2006). The fact that control

533   and even stereotypic horses learned the reversal task let us suggest that reversal

534   learning based only on visual cue is possible under certain conditions. Moreover, the fact

535   that horses performed the second reversal in less trials than the first one, confirmed that

536   horses learned to learn as it was also demonstrated in other studies testing horses in

537   visual or spatial discrimination tasks (Fiske and Potter 1979; Martin et al. 2006; Voith

538   1975; Warren and Warren 1962). As underlined by Brubaker and Udell (2016) the study

539   protocol and nature of the visual stimuli appear to affect a horse's ability to perform at

540   any given cognitive task. In our study, we adapted the experimental protocol, based on

541   preliminary tests and previous studies (Flannery 1997; Gabor and Gerken 2010; Hall et

542   al. 2003) in order to keep the horses motivated, as follows. Firstly, we chose to oppose

543   two signals that differed only in whether they were black or white, because such colours

544   seem to be easy to differentiate by horses. Indeed, horse ability to discriminate between

545   different colours seems to be limited due to their dichromatic vision (Blackmore et al.

20

546   2008). Secondly, we ensured that the stimuli were presented at the ground level.

547   Indeed, former studies demonstrated that horse performance was improved when stimuli

548   were presented at the ground level, compared at the height of 90 cm from the ground

549   (Hall et al. 2003). Thirdly, we adapted the number of trials per sessions and rewarded

550   the horses during the study if they had been choosing the wrong stimulus three times in

551   a row, in order to maintain the attention span and motivation of the horses (Flannery

552   1997; Rubin et al. 1980; Sappington et al. 1997). Finally, we waited until all the horses

553   reached the learning criterion before stopping the study, even when a relatively high

554   number of trials was required ($N = 537$). We suggest that similar precautions might help

555   improve motivation of horses in future cognitive studies.

556

557   *Learning performance of crib-biters compared to control horses*

558   We did not find any difference in the number of trials needed to reach the learning

559   criterion between crib-biter and control horses, contrary to previous studies on the same

560   topic (Hemmings et al. 2007; Parker et al. 2009; Parker et al. 2008; Roberts et al.

561   2015). Indeed, previous studies found that crib-biters might be more prone to habit

562   learning than to response-outcome learning (Hemmings et al. 2007; Parker et al. 2009;

563   Parker et al. 2008; Roberts et al. 2015). Therefore, we expected that they would reach

564   the learning criteria during Acq1 or Acq2 faster compared to the controls. In addition,

565   since previous studies also demonstrated that crib-biting horses need more operant

566   responses compared to the other horses before the extinction of a previously learnt

567   action (Hemmings et al. 2007; Roberts et al. 2015) and were unable to maintain

568   response–outcome learning in a continuously applied learning paradigm (Parker et al.

569   2008), we expected that the crib-biters would need more trials in Rev 1 and Rev 2 to

570   reach the criteria compared to controls. In contrast to these predictions, in our study, all

571   the crib-biters were able to achieve the different phases (Acq1, Rev1, Acq2, Rev2) in a

572   similar number of trials compared to the controls. An explanation for these discrepancies

573   between our studies and previous ones could be that in our experiment, the crib-biters

574   had the opportunity to crib-bite on the learning apparatus. It is not always clear whether

21

575   stereotypic horses had the opportunity to crib-bite, and did so, during previous studies

576   (Hemmings et al. 2007; Parker et al. 2009; Parker et al. 2008). However, in Roberts et

577   al. (2015), crib-biting straps were removed prior to the tests, although no information

578   about crib-biting events is specified. If crib-biting is indeed a coping strategy (Briefer

579   Freymond et al. 2015), reducing stress levels could, as a result, improve their learning

580   capacities, allowing them to achieve the same performances as non-stereotypic horses

581   (Schwabe and Wolf 2010; Valenchon et al. 2013). Even if other studies did not measures

582   stress parameters during learning tasks (Hemmings et al. 2007; Parker et al. 2009;

583   Parker et al. 2008; Roberts et al. 2015), this suggests that allowing crib-biting horses to

584   perform their stereotypic behaviour during learning could improve their learning abilities.

585

586   Since stereotypies in animals are often likened to human developmental, neurological or

587   severe psychiatric disorders (e.g., autism, obsessive compulsive disorder (OCD) or

588   schizophrenia) (McBride and Parker 2015), comparisons between our results and human

589   disorders can be made. Even if OCD patients usually report that they get a sort of relief

590   by performing their rituals, and that preventing performance increases their anxiety

591   (Boyer and Lienard 2006), the literature on autism in humans suggests that, on the

592   contrary, children exhibiting high levels of stereotypy fail to learn while engaged in

593   stereotypy (Cunningham and Schreibman 2008). On the other hand, our findings could

594   be related to results found in another human psychological disorder, named attention-

595   deficit/hyperactivity disorder (ADHD). Similarly to animal stereotypies, ADHD patients

596   show non-goal oriented motor movements. These movements are however, unlike

597   stereotypies, not executed as repetitive invariant patterns. A recent paper showed that

598   performing such movement is associated with an improvement in cognitive performance

599   (Sarver et al. 2015). Such findings are in accordance with our results in crib-biting

600   horses. As suggested in Hausberger et al. (2007), stereotypic horses differ from other

601   horses in their behaviour and may require specific training. Letting these horses the

602   possibility to perform their stereotypy might be, as our study suggests, one specific

603   feature to incorporate in learning protocols, which might then allow them to perform

604 successfully. Future studies could compare the learning capacities of crib-biter horses

605 prevented or not to crib-bite and of a corresponding number of control horses subjected

606 to the same treatment. To summarize, the results of our study do not support the

607 hypothesis that crib-biters display alterations in learning abilities, which could result from

608 impaired dopaminergic system. In addition, our findings suggest that, in the same way

609 as the performance of non-goal oriented motor movements improves cognitive

610 performances of ADHD human patients (Sarver et al. 2015), the performance of

611 stereotypic behaviour might improve crib-biting horse learning abilities.

612

613 During the acclimation and pre-training phase, however, crib-biters needed more

614 sessions than controls to attain the learning criterion, even if they also had the possibility

615 to crib-bite. Since we did not collect measures of stress indicators (e.g. HR, RMSSD)

616 during this period, we cannot make inferences about the stress level of stereotypic

617 horses compared to controls during this phase. Nevertheless, we could suggest that crib-

618 biting horses were less focussed on the task during the acclimation and pre-training

619 compared to the test phase, because horses had to manage too many other external

620 stimuli. A general difficulty of maintaining a task focus (i.e. attention) in stereotypic

621 compared to control horses has been suggested by Hausberger et al. (2007). These

622 authors proposed that the time invested in performing stereotypic behaviour throughout

623 the day and at night is likely to affect sleep quality and quantity in stereotypic horses.

624 This could lead to a general lower attention span in these horses than in non-stereotypic

625 horses. Attention state and motivation being primordial for learning (Cowan 1998;

626 Rochais et al. 2014), we could hypothesise that crib-biters might needed a longer time

627 than control horses to habituate to novel situations and be able to ignore and exclude

628 external stimuli (i.e. new area, apparatus). However, after a longer habituation than for

629 the controls, they might be able to focus on the cognitive task (i.e. test phase). During

630 the test phase itself, as a result of our protocol that was designed to maximise attention

631 span and motivation (e.g. short training sessions of around 20 min, "correction trials"

632 (Flannery 1997)), all horses, including crib-biters, seemed very attentive and motivated,

633 which might have boosted their performances. Attention deficits have also been

634 demonstrated in ADHD disorder, which has the particularity to induce difficulties in

635 maintaining task focus (Sarver et al. 2015). In order to test the hypothesis that crib-

636 biters are generally less attentive than other horses, future studies could evaluate the

637 distractibility (i.e. how much can an individual be distracted by external stimuli) of these

638 horses compared to non-stereotypic individuals (e.g. using a "distractibility test";

639 (Rochais et al. 2017)). Another indicator of attention that could be used to test such

640 hypothesis is spontaneous blink rate (SBR; (Roberts et al. 2015)). SBR is a basic

641 measure of dopamine transmission utilised to determine striatal functioning in stereotypy

642 performing humans and could also be applied to animals (Roberts et al. 2015). Using this

643 indicator, Roberts et al. (2015) demonstrated that crib-biters display lower SBR than

644 other horses (Roberts et al. 2015). Interestingly, SBR has been recently shown to

645 constitute an index of dopaminergic component of sustained attention and fatigue in

646 humans (Maffei and Angrilli 2018). By comparing the distractibility and attention of

647 stereotypic and control horses when performing cognitive tests, future studies might be

648 able to further highlight the need to adapt the design of training sessions to horses

649 suffering from stereotypies, in order to maximum their learning abilities and improve

650 their welfare (e.g. by avoiding frustration).

651

652 A last explanation for the discrepancies between our studies and other studies

653 investigating learning alteration in crib-biting horse could be that reversal learning tests

654 are perhaps not appropriate for assessing dopaminergic alterations that might be present

655 in crib-biters. Such alterations, including a higher number of dopamine receptor subtypes

656 in the ventral striatum or nucleus accumbens (Nac) and a lower number of such

657 receptors in the dorsomedial striatum (DMS) or caudate have been shown using post-

658 mortem analysis in crib-biting horses (McBride and Hemmings 2005). However,

659 behavioural flexibility, or the ability to adjust responses according to a change in the

660 environment, is mediated by a large neural network, including prefrontal-basal ganglia

661 circuits in addition to the dorsal and ventral striatum. As mentioned previously, the DMS

662   has been identified as an important structure for flexible responding (Castane et al.

663   2010). Indeed, DMS lesions, due to the role of this brain structure in learning, might

664   result in a switch from goal-directed to habit formation and thus in the impairment of the

665   development of habits (Yin et al. 2008). However the role of the ventral part of the DMS

666   (Nac) in instrumental performance remains nowadays controversial (Yin et al. 2008). For

667   example, some studies found that lesions in the Nac do not impair spatial, visual or

668   motor reversal in monkeys, Macaca fascicularis (Stern and Passingham 1995), while

669   other studies found that such lesions impaired both an initial discrimination and its

670   reversal in Lister hooded rats (Annett et al. 1989). A reason for these discrepancies

671   between studies could be that most studies on DMS or dorsolateral (DLS) lesions have

672   used rats, despite the fact that it is difficult to compare the physical location of dorsal or

673   ventral striatum in rat and other species such as primates for example (Yin et al. 2008).

674   To conclude, reversal learning paradigms are among the most widely used tests for

675   cognitive flexibility and there is accumulating evidence that DMS is involved in this type

676   of learning (Castane et al. 2010; Izquierdo et al. 2017; Ragozzino et al. 2003). However

677   the role of Nac, which has been suggested to be impaired in crib-biters (McBride and

678   Hemmings 2005), in reversal learning tasks is controversial (Yin et al. 2008). Therefore,

679   testing crib-biters with other cognitive tests than reversal learning might be perhaps

680   more valuable in order to investigate the suggested impairment in the Nac (McBride and

681   Hemmings 2005).

682

683   *Effect of the change in colour*

684   In our study, the colour of the signal always changed from white to black or vice-versa,

685   between the acquisition and the corresponding reversal. However, for half of the horses

686   (half of the crib-biters and four of the seven controls), a change in colour occurred also in

687   the middle of the learning procedure, between Rev1 and Acq2 (Table1, Fig 1). This

688   change in colour in the middle might have been experienced as an additional reversal

689   (based on colour only). In accordance with this hypothesis, the results showed a

690   significant effect of colour change between Rev1 and Acq2, with horses submitted to the

691   colour change needing more trials than the other horses. We suggest that further studies

692   including several acquisition phases with different visual stimuli should be aware that

693   changes in colours between phases might be perceived by the animals as reversals.

694

695   *Stereotypy level and performance*

696   The crib-biters in this study differed with regards to the strength of their stereotypy

697   (Table 1). Many studies have reported more cognitive difficulties in animals displaying a

698   higher frequency of stereotypic behaviour compared to less stereotypic ones (Garner and

699   Mason 2002; Garner et al. 2003; Vickery and Mason 2005). Indeed, stereotypic levels

700   have been shown to correlate with an increase in the persistence of inappropriate

701   responses in an extinction learning test in bears, *Ursus thibetanus* (Vickery and Mason

702   2005). However, our results did not show such a trend. Our three groups based on the

703   frequency of crib-biting of the horses assessed over 48h before the experiment started

704   (GroupSMLC, "S", strong crib-biters, "M", medium crib-biters and "L", low crib-biters),

705   did not differ in their learning performance. Therefore, we did not find any evidence

706   showing that the frequency of crib-biting is a factor that influences the cognitive abilities

707   of horses. This absence of group difference could also suggest that the stereotypic level

708   is not a good indicator of dopaminergic system alterations. Similar results have been

709   found in rhesus macaques, *Macaca mulatta* (Pomerantz et al. 2012). Interestingly, in this

710   study the authors found that some type of stereotypies did correlate with perseveration

711   while some did not. Future studies could investigate potential links between learning

712   performance and the time since a horse started crib-biting instead of the stereotypic

713   level.

714

715   *Physiological parameters*

716   We did not find any evidence for group differences in the sympathomedullary (SAM)

717   axis parameters measured in this study (HR and RMSSD) during the phases of

718   acquisitions and their respective reversals. Because one potential cause of

719   stereotypies is a previous exposure to a chronic stress situation that could induce

720   higher sensitivity to stress (Bhatnagar and Vining 2003), we would have expected

721   crib-biters to be more stressed than controls. However, neither the results of the

722   present study, nor those of our previous study revealed any difference in SAM

723   parameters between crib-biters and controls (Briefer Freymond et al. 2015). It also

724   suggests that the crib-biters were experiencing similar stress levels as controls

725   during the experiment, possibly as a result of crib-biting on the apparatus, which

726   might have reduced their stress levels (Briefer Freymond et al. 2015). We could also

727   have expected horses to be more physiologically stressed during the first reversal

728   compared to the other learning phase as this learning phase might be more

729   challenging for them, as displayed by the increased number of trials required to

730   achieve this task. However, we did not find any effect of the learning phase

731   (acquisition or reversal) on HR or RMSSD. Finally, it is possible that other

732   parameters than HR and RMSSD might be more adequate to measure stress during

733   a learning task involving locomotor behaviour. Indeed, HR and RMSSD are also

734   influenced by physical activity (von Borell et al. 2007). For this reason, only

735   measures made during times of similar behavioural pattern should be compared

736   (von Borell et al. 2007). Our assumption is that this is the case in our study,

737   because all horses (crib-biters and controls) had to perform the same trajectory,

738   and the same number of trials per session. However, further studies could aim at

739   designing tasks involving less movement and take also additional measures of

740   stress, such as behavioural measures (e.g. Equine Facial Action Coding Systems

741   (FACS) (Wathan et al. 2015), behaviour scores (Young et al. 2012)) during learning

742   tasks as well as during habituation.

743

744   *Animal welfare*

745   Animal welfare being of increasing public and scientific concern, it is important to

746   understand the link between stereotypic behaviour and animal welfare. It has been

747   proposed that stereotypic behaviour might indicate poor welfare only if a dopaminergic

748   dysfunction is present (Mason and Latham 2004). Indeed, even if stereotypies develop

749 under contexts of chronic stress state, their performance, once fully developed, might not

750 necessarily indicate poor welfare. For instance, in this case, the performance of

751 stereotypic behaviours might serve as coping mechanisms, helping individuals to reduce

752 their stress level (e.g *mantra effect"*, (Mason and Latham 2004)). Mason and Latham

753 (2004) hence proposed that stereotypies correlate with poor welfare only when they have

754 become a habit and, only when behaviours have changed in control and have become

755 environmentally insensitive. At this developmental stage, stereotypies are performed in a

756 more diverse set of situations and are harder to interrupt. However, such stage is not

757 easy to assess because the performance of stereotypies may vary between individuals, in

758 terms of stereotypy level over time. Mason and Latham (2004) also argued that

759 perseverative responding, resulting from *basal ganglia* dysfunction, also indicates poor

760 welfare. In fact, with perseveration, individuals may produce unnecessary and

761 inappropriate responses to environmental cues. In humans, as mentioned earlier,

762 perseveration is also correlated with human disorders, like schizophrenia, autism and

763 other brain injuries. In conclusion, if neurobiological changes are linked to stereotypies,

764 resulting in alterations in the learning profile of animals, stereotypic behaviour should

765 indicate poor welfare. However, in this study, we could not conclude to the existence

766 neurobiological alteration in crib-biters, since these horses did not need more trials to

767 perform the reversal learning tasks compared to control horses. Further studies on

768 cognitive abilities of crib-biter horses are thus required to determine the impact of this

769 stereotypy on horse welfare.

770

771 **CONCLUSION**

772 Our study did not reveal any difference in cognitive abilities between crib-biters and

773 controls and therefore we cannot conclude that stereotypic horses suffer from a

774 dopaminergic dysfunction. Indeed our results show that all horses, including stereotypic

775 horses and controls, were able to perform reversal discrimination tasks based on visual

776 cues, and that they even learned to learn (i.e. improve their performance from one

777 reversal to the next). An explanation for the discrepancies between our study and the

778   previous ones could be that, in our study, the crib-biters had the opportunity to crib-bite

779   on the learning apparatus, which might have enabled them to reduce their stress level,

780   as suggested by the lack of group difference in physiological stress parameters. Further

781   studies could test the learning capacities of crib-biters that are prevented or not to

782   perform the stereotypic behaviour against a group of non-stereotypic horses subjected to

783   the same treatment. Finally, our results point towards several parallels between horse

784   crib-biting behaviour and human developmental, neurological or psychiatric disorders,

785   such as ADHD disorders, suggesting that the study of horse crib-biting behaviour could

786   serve as a good animal model to better understand such disorders in human (Brace et al.

787   2015).

788

793

794   **ETHICS**

795   The experimental procedure for the horses was approved by the Federal Veterinary Office

796   (approval number VD 26777 bis; Switzerland).

797

798   **CONFLICTS OF INTEREST**

799   The authors declare that they have no competing interest.

800

801

802   **REFERENCES**

803

804   Annett LE, McGregor A, Robbins TW (1989) The effects of ibotenic acid lesions of the

805   nucleus accumbens on spatial learning and extinction in the rat. Behav Brain Res

806   31:231-242

807   Bhatnagar S, Vining C (2003) Facilitation of hypothalamic pituitary adrenal responses to

808   novel stress following repeated social stress using the resident/intruder paradigm.

809   Hormon & Behav 43:158-165

810   Blackmore TL, Foster TM, Sumpter CE, Temple W (2008) An investigation of colour

811   discrimination with horses (Equus caballus). Behav Proc 78:387-396

812   Boyd BA, McBee M, Holtzclaw T, Baranek GT, Bodfish JW (2009) Relationships among

813   Repetitive Behaviors, Sensory Features, and Executive Functions in High

814   Functioning Autism. Res Autism Spectr Disord 3:959-966

815   doi:10.1016/j.rasd.2009.05.003

816   Boyer P, Lienard P (2006) Why ritualized behavior? Precaution Systems and action

817   parsing in developmental, pathological and cultural rituals. Behav Brain Sci

818   29:595-613; discussion 613-550

819   Brace LR, Kraev I, Rostron CL, Stewart MG, Overton PG, Dommett EJ (2015) Auditory

820   responses in a rodent model of Attention Deficit Hyperactivity Disorder. Brain Res

821   1629:10-25

822   Briefer Freymond S, Bardou D, Briefer EF, Bruckmaier R, Fouché N, Fleury J, Maigrot AL,

823   Ramseyer A, Zuberbühler K, Bachmann I (2015) The physiological consequences of crib-

824   biting in horses in response to an ACTH challenge test. Physiol & Behav 151:121-

825   128 doi:https://doi.org/10.1016/j.physbeh.2015.07.015

826   Brubaker L, Udell MAR (2016) Cognition and learning in horses (Equus caballus): What

827   we know and why we should ask more. Behav Proc 126:121-131

828   doi:https://doi.org/10.1016/j.beproc.2016.03.017

829   Cabib S, Giardino L, Calzà L, Zanni M, Mele A, Puglisi-Allegra S (1998) Stress promotes

830   major changes in dopamine receptor densities within the mesoaccumbens and

831   nigrostriatal systems. Neuroscience 84:193-200

832   doi:https://doi.org/10.1016/S0306-4522(97)00468-5

833 Castane A, Theobald DE, Robbins TW (2010) Selective lesions of the dorsomedial

834   striatum impair serial spatial reversal learning in rats. Behav Brain Res 210:74-83

835   doi:10.1016/j.bbr.2010.02.017

836 Cowan N (1998) Attention and memory: An integrated framework. Oxford University

837   Press,

838 Cunningham AB, Schreibman L (2008) Stereotypy in Autism: The Importance of

839   Function. Res Autism Spectr Disord 2:469-479 doi:10.1016/j.rasd.2007.09.006

840 Dickinson A (1985) Actions and habits: the development of behavioural autonomy.

841   Philosophical Transactions of the Royal Society of London B, Biol Sci 308:67-78

842   doi:10.1098/rstb.1985.0010

843 Fiske JC, Potter GD (1979) Discrimination Reversal Learning in Yearling Horses. J of Anim

844   Sci 49:583-588 doi:10.2527/jas1979.492583x

845 Flannery B (1997) Relational discrimination learning in horses. Appl Anim Behav Sci

846   54:267-280 doi:10.1016/S0168-1591(97)00006-3

847 Gabor V, Gerken M (2010) Horses use procedural learning rather than conceptual

848   learning to solve matching to sample. Appl Anim Behav Sci 126:119-124

849   doi:10.1016/j.applanim.2010.06.006

850 Garner JP, Mason GJ (2002) Evidence for a relationship between cage stereotypies and

851   behavioural disinhibition in laboratory rodents. Behav Brain Res 136:83-92

852   doi:https://doi.org/10.1016/S0166-4328(02)00111-0

853 Garner JP, Meehan CL, Mench JA (2003) Stereotypies in caged parrots, schizophrenia and

854   autism: evidence for a common mechanism. Behav Brain Res 145:125-134

855 Hall CA, Cassaday HJ, Derrington AM (2003) The effect of stimulus height on visual

856   discrimination in horses. J. of Anim Sci 81:1715-1720

857   doi:10.2527/2003.8171715x

858 Hausberger M, Gautier E, Müller C, Jego P (2007) Lower learning abilities in stereotypic

859   horses. Appl Anim Behav Sci 107:299-306

860   doi:https://doi.org/10.1016/j.applanim.2006.10.003

861 Hemmings A, McBride SD, Hale CE (2007) Perseverative responding and the aetiology of

862   equine oral stereotypy. Appl Anim Behav Sci 104:143-150

863   doi:https://doi.org/10.1016/j.applanim.2006.04.031

864 Hothersall B, Gale E, Harris P, Nicol C (2010) Cue use by foals (Equus caballus) in a

865   discrimination learning task. Anim Cogn 13:63-74

866 Houpt KAM, S.M. (1993) Equine stereotypies Compend Contin Educ:1265-1271

867 Hughes BO, Duncan IJH (1988) The notion of ethological 'need', models of motivation

868   and animal welfare. Anim Behav 36:1696-1707

869   doi:https://doi.org/10.1016/S0003-3472(88)80110-6

870 Izquierdo A, Brigman JL, Radke AK, Rudebeck PH, Holmes A (2017) The neural basis of

871   reversal learning: An updated perspective. Neuroscience 345:12-26

872   doi:10.1016/j.neuroscience.2016.03.021

873 Lopez BR, Lincoln AJ, Ozonoff S, Lai Z (2005) Examining the relationship between

874   executive functions and restricted, repetitive symptoms of Autistic Disorder. J

875   Autism Dev Disord 35:445-460 doi:10.1007/s10803-005-5035-x

876 Luescher UA, McKeown DB, Halip J (1991) A cross-sectional study on compulsive

877   behavior (stable vices) in horses. Equine Vet J 27:14-18

878 Maffei A, Angrilli A (2018) Spontaneous eye blink rate: An index of dopaminergic

879   component of sustained attention and fatigue. Interna J. of Psychophysiol 123:58-

880   63 doi:https://doi.org/10.1016/j.ijpsycho.2017.11.009

881 Martin TI, Zentall TR, Lawrence L (2006) Simple discrimination reversals in the domestic

882   horse (Equus caballus): Effect of discriminative stimulus modality on learning to

883   learn. Appl Anim Behav Sci 101:328-338

884   doi:https://doi.org/10.1016/j.applanim.2006.02.011

885 Mason G, Latham N (2004) Can't stop, won't stop: is stereotypy a reliable animal welfare

886   indicator? Anim Welf 13:57-69

887   Mason G, Rushen J (2006) Stereotypic Animal Behaviour: Fundamentals and Applications
888   to Welfare. CAB Interantional, Wallingford OXON second ed.
889   McBride SD, Hemmings A (2005) Altered mesoaccumbens and nigro-striatal dopamine
890   physiology is associated with stereotypy development in a non-rodent species.
891   Behav Brain Res 159:113-118 doi:https://doi.org/10.1016/j.bbr.2004.10.014
892   McBride SD, Parker MO (2015) The disrupted basal ganglia and behavioural control: An
893   integrative cross-domain perspective of spontaneous stereotypy. Behav Brain Res
894   276:45-58 doi:https://doi.org/10.1016/j.bbr.2014.05.057
895   McBride SD, Parker MO, Roberts K, Hemmings A (2017) Applied neurophysiology of the
896   horse; implications for training, husbandry and welfare. Appl Anim Behav Sci
897   190:90-101 doi:https://doi.org/10.1016/j.applanim.2017.02.014
898   McBride SD, Perentos N, Morton AJ (2016) A mobile, high-throughput semi-automated
899   system for testing cognition in large non-primate animal models of Huntington
900   disease. J Neurosci Methods 265:25-33 doi:10.1016/j.jneumeth.2015.08.025
901   McCall CA (1990) A Review of Learning Behavior in Horses and its Application in Horse
902   Training. J Anim Sci 68:75-81 doi:10.1093/ansci/68.1.75
903   McGreevy P, McLean, A. (2010) Equitation Science. John Wiley & Sons, pp 328
904   McGreevy PD, Cripps PJ, French NP, Green LE, Nicol CJ (1995) Management factors
905   associated with stereotypic and redirected behaviour in the Thoroughbred horse.
906   Equine Vet J 27:86-91
907   Parker M, McBride SD, Redhead ES, Goodwin D (2009) Differential place and response
908   learning in horses displaying an oral stereotypy. Behav Brain Res 200:100-105
909   doi:https://doi.org/10.1016/j.bbr.2008.12.033
910   Parker M, Redhead ES, Goodwin D, McBride SD (2008) Impaired instrumental choice in
911   crib-biting horses (Equus caballus). Behav Brain Res 191:137-140
912   doi:10.1016/j.bbr.2008.03.009
913   Pomerantz O, Paukner A, Terkel J (2012) Some stereotypic behaviors in rhesus
914   macaques (Macaca mulatta) are correlated with both perseveration and the ability

915   to cope with acute stressors. Behav Brain Res 230:274-280

916   doi:10.1016/j.bbr.2012.02.019

917 Ragozzino ME, Kim J, Hassert D, Minniti N, Kiang C (2003) The contribution of the rat

918   prelimbic-infralimbic areas to different forms of task switching. Behav Neurosci

919   117:1054-1065 doi:10.1037/0735-7044.117.5.1054

920 Ringhofer M, Yamamoto S (2017) Domestic horses send signals to humans when they

921   face with an unsolvable task. Anim Cogn 20:397-405 doi:10.1007/s10071-016-

922   1056-4

923 Roberts K, Hemmings A, Moore-Colyer M, Hale CE (2015) Cognitive differences in horses

924   performing locomotor versus oral stereotypic behaviour. Appl Anim Behav Sci

925   168:37-44 doi:https://doi.org/10.1016/j.applanim.2015.04.015

926 Roberts K, Hemmings AJ, McBride SD, Parker MO (2017) Causal factors of oral versus

927   locomotor stereotypy in the horse. J of Vet Behav: Clinical Applications and

928   Research 20:37-43 doi:https://doi.org/10.1016/j.jveb.2017.05.003

929 Rochais C, Henry S, Hausberger M (2017) Spontaneous attention-capture by auditory

930   distractors as predictor of distractibility: a study of domestic horses (Equus

931   caballus). Scientific Report 7:15283

932 Rochais C, Henry S, Sankey C, Nassur F, Gorecka-Bruzda A, Hausberger M (2014) Visual

933   attention, an indicator of human-animal relationships? A study of domestic horses

934   (Equus caballus). Front in Psy 5:108

935 Rubin L, Oppegard C, Hindz HF (1980) The effect of varying the temporal distribution of

936   conditioning trials on equine learning behavior. J of Anim Sci 50:1184-1187

937 Sappington BKF, McCall CA, Coleman DA, Kuhlers DL, Lishak RS (1997) A preliminary

938   study of the relationship between discrimination reversal learning and

939   performance tasks in yearling and 2-year-old horses. Appl Anim Behav Sci

940   53:157-166 doi:10.1016/S0168-1591(96)01157-4

941 Sarver DE, Rapport MD, Kofler MJ, Raiker JS, Friedman LM (2015) Hyperactivity in

942   Attention-Deficit/Hyperactivity Disorder (ADHD): Impairing Deficit or

943   Compensatory Behavior? J of Abnorm Child Psychol 43:1219-1232

944   doi:10.1007/s10802-015-0011-1

945 Schwabe L, Wolf OT (2010) Learning under stress impairs memory formation. Neurobiol

946   Learn Mem 93:183-188 doi:https://doi.org/10.1016/j.nlm.2009.09.009

947 Schwabe L, Wolf OT (2011) Stress-induced modulation of instrumental behavior: From

948   goal-directed to habitual control of action. Behav Brain Res 219:321-328

949   doi:https://doi.org/10.1016/j.bbr.2010.12.038

950 Spruijt BM, van den Bos R, Pijlman FT (2001) A concept of welfare based on reward

951   evaluating mechanisms in the brain: anticipatory behaviour as an indicator for the

952   state of reward systems. Appl Anim Behav Sci 72:145-171

953 Stern CE, Passingham RE (1995) The nucleus accumbens in monkeys (Macaca

954   fascicularis). III. Reversal learning. Exp Brain Res 106:239-247

955 Toates F (2004) Cognition, motivation, emotion and action: a dynamic and vulnerable

956   interdependence. Appl Anim Behav Sci 86:173-204

957   doi:10.1016/j.applanim.2004.02.010

958 Valenchon M, Lévy F, Fortin M, Leterrier C, Lansade L (2013) Stress and temperament

959   affect working memory performance for disappearing food in horses, Equus

960   caballus. Anim Behav 86:1233-1240 doi:10.1016/j.anbehav.2013.09.026

961 Van der Harst JE, Baars AM, Spruijt BM (2003) Standard housed rats are more sensitive

962   to rewards than enriched housed rats as reflected by their anticipatory behaviour.

963   Behav Brain Res 142:151-156

964 Vickery SS, Mason GJ (2005) Stereotypy and perseverative responding in caged bears:

965   further data and analyses. Appl Anim Behav Sci 91:247-260

966 Voith VL (1975) Pattern discrimination, learning set formation, memory retention, spatial

967   and visual reversal learning by the horse. The Ohio State University

968 von Borell E et al. (2007) Heart rate variability as a measure of autonomic regulation of

969   cardiac activity for assessing stress and welfare in farm animals: A review. Paper

970   presented at the Physiology & Behavior, Stress and Welfare in Farm Animals,

971   10/22/2007

972   Warren J, Warren HB (1962) Reversal learning by horse and raccoon. J of Gen Psychol

973   100:215-220

974   Wathan J, Burrows AM, Waller BM, McComb K (2015) EquiFACS: the equine facial action

975   coding system. PLoS one 10:e0131738

976   Wickens CL, Heleski CR (2010) Crib-biting behavior in horses: A review. Appl Anim Behav

977   Sci 128:1-9

978   Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nature

979   reviews Neuroscience 7:464-476 doi:10.1038/nrn1919

980   Yin HH, Ostlund SB, Balleine BW (2008) Reward-guided learning beyond dopamine in the

981   nucleus accumbens: The integrative functions of cortico-basal ganglia networks.

982   Eur J Neurosci 28:1437-1448 doi:10.1111/j.1460-9568.2008.06422.x

983   Young T, Creighton E, Smith T, Hosie C (2012) A novel scale of behavioural indicators of

984   stress for use with domestic horses. Appl Anim Behav Sci 140:33-43

985

986

Table 1. Characteristics of the horses used in the experiment. Group (controls = "C"; crib-biters = "CB"; strong crib-biters = "S", medium crib-biters = "M", low frequency crib-biters = "L", controls = "C"), Sex (female = f; gelding = g, stallion = s), Breed (Franches-Montagnes = "FM"; English thoroughbred = "ET"; warmblood = "WB"; Camargue horse= "CA"; Hispano-Arabian = "HA"), year of birth, stimuli used in the first set of learning tasks (Acquisition 1 = "Acq1", Reversal 1 = "Rev1) and in the second set of learning tasks (Acquisition 2 = "Acq2", Reversal 2 = "Rev2"), and presence of a change in colour between Rev1 and Acq2 or not (change of colour = "Y", no change of colour "N").

Horses	Group	Breed	Sex	Birth	Acq1/Rev1	Acq2/Rev2	Colour
1	C	FM	f	1993			Y
2	C	FM	s	1996			Y
3	C	ET	g	1991			Y
4	C	FM	g	2001			Y
5	C	WB	g	2002			N
6	C	CA	f	2000			N
7	C	FM	s	2006			N
8	CB-S	FM	f	1997			Y
9	CB-M	FM	s	2002			Y
10	CB-L	FM	s	2005			Y
11	CB-M	FM	g	2004			N
12	CB-S	ET	g	2003			N
13	CB-S	HA	f	2004			N

37

996 **Table 3. Number of crib-biting events ("CBnb")** and number of trials ("Trial")

997 needed to attain the learning criterion for the corresponding phase ("Phase"; Acq1, Rev1,

998 Acq2, Rev2). "Horses" refer to the number attributed to each horses in Table 1.

999

Horses	CBnb	Trials	Phase
8	1	144	Acq1
8	0	94	Rev1
8	0	124	Acq2
8	0	105	Rev2
9	1	109	Acq1
9	4	247	Rev1
9	0	229	Acq2
9	0	9	Rev2
10	0	14	Acq1
10	0	149	Rev1
10	0	65	Acq2
10	0	1	Rev2
11	0	74	Acq1
11	0	420	Rev1
11	0	49	Acq2
11	0	220	Rev2
12	0	111	Acq1
12	0	478	Rev1
12	0	27	Acq2
12	0	157	Rev2
13	121	53	Acq1
13	745	298	Rev1
13	380	73	Acq2
13	2634	473	Rev2

1000

1001

Table 3. Abbreviations of the parameters used in the different analysis.

parameter abbreviation	definition
"Phase" (c)	Different phases of learning until the learning criterion (LC) is reached;
"Acq1"	First acquisition: phase during which the horses learned to choose the rewarded signal among a first set of two stimuli.
"Rev1"	First reversal: phase during which the horses learned to choose the signal that was unrewarded in Acq1.
"Acq2"	Second acquisition: phase during which the horses learned to choose the rewarded signal among a novel set of stimuli.
"Rev2"	Second reversal: phase during which the horses learned to choose the signal that was unrewarded in Acq2.
"Correct" (f)	Frequency of correct choices during the last session of Acq1 and Acq2, and during the first session of Rev1 and Rev2
"Sessions" (n)	Number of sessions of 10 min during the acclimation and pre-training (two sessions per day)
"Trial" (n)	Number of trials needed until the learning criterion is reached (6 correct trials in a row)
"Signal" (c)	Cross or circle
"Col" (c)	Colour of the signal (black or white)
"Person" (c)	Person leading the horse
GroupCB-C ("CB" or "C") (c)	Crib-biting or control group
GroupSMLC ("S", "M", "L" and "C") (c)	Groups of crib-biters based on the frequency of crib-biting events per hour over 48h (4 groups; strong crib-biters, medium crib-biters, low-frequency crib-biters and controls)

"Colour": "Y" or "N" (c)	Whether the colour changed between Rev1 and Acq2 (yes or no)
"HR" (m)	Heart rate (average value per horse per phase, in BPM)
"RMSSD" (m)	Root mean square of successive inter-beat interval differences (average value per horse per Phase, in ms)

c: category, f: frequency; m: mean; n: number

1003

1004

1005

1006

1007

1008 **Table 4. Response variables, fixed and control parameters used in the different**

1009 **models.** The abbreviations are described in Table 2. The crosses indicate which parameters

1010 and which response variable were used in the different models. The fixed parameters are

1011 the GroupCB-C or GroupSMLC, the Phase and the Colour depending on the model. The

1012 others parameters are control parameters. For the model with Correct as response

1013 variable, we selected only Rev1 and Acq2 among the other Phases. "[1]" indicates that we

1014 used either Phase or Colour as fixed parameters in the model.

1015

1016

Response variable	Acclimation	Learning performance	Learning capacities	physiology
Session	x			
Correct		x		
Trial			x	
HR				x
RMSSD				x
Fixed factors	**Acclimation**	**Learning performance**	**Learning capacities**	**physiology**
GroupCB-C	x	x	x	x
GroupSMLC	x	x	x	x
Phase		x	x^1	x^1
Colour			x^1	x^1
Phase x GroupCB-C		x	x^1	x^1
Phase x GroupSMLC		x	x^1	x^1
Colour x GroupCB-C			x^1	x^1
Colour x GroupSMLC			x^1	x^1
Control factors	**Acclimation**	**Learning performance**	**Learning capacities**	**physiology**
sex	x	x	x	x
age	x	x	x	x
Person	x	x	x	x
Col		x	x	x
Signal		x		

1017

1018

1019

1020 **Figure Legends**

1021

1022 **Figure 1. Test-apparatus for visual discrimination task.** The visual stimuli were

1023 inserted on the front side of the apparatus inside a plastic window (indicated in dash

1024 line). In case of a correct choice, the horses could reach the food through the

1025 corresponding flap. A vertical piece of wood was added in the middle of the apparatus.
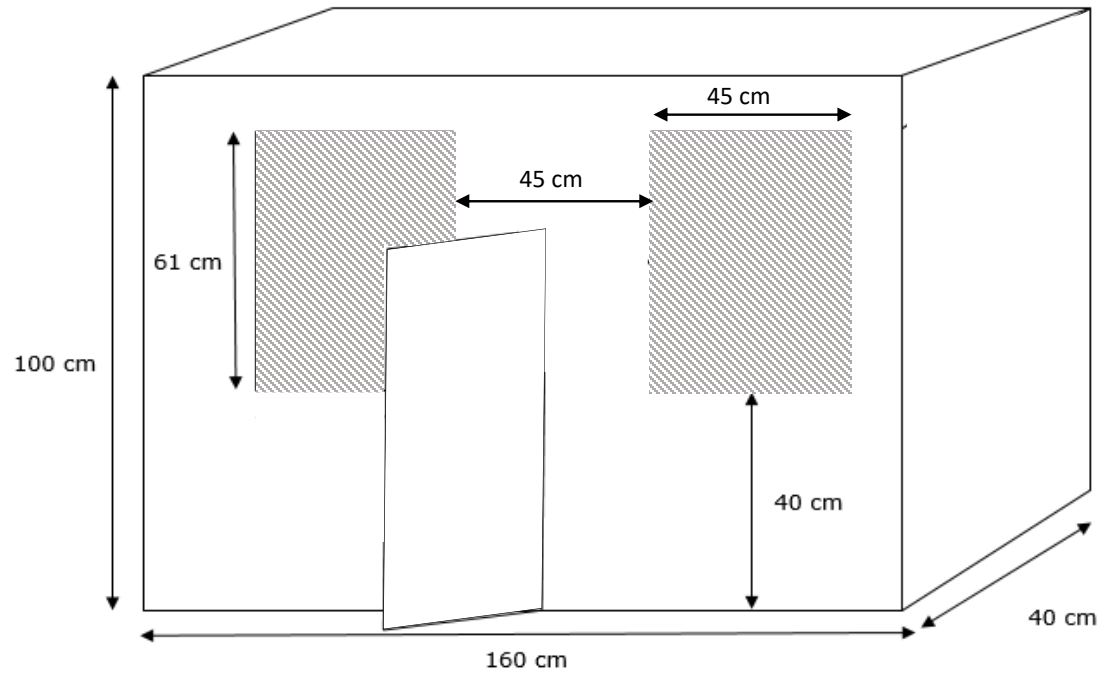
1026

1027 **Figure 2. Session-by-session summary of the performance of all horses**. Data are

1028 the frequency (±SD) of correct choices per session for all horses of a group (control

1029 horses in black and crib-biters in grey) combined. The dots indicate the different session.

1030 For each horse, once the learning criterion (6 correct trials in a row) was reached, it was

1031 assigned a score of 90% until all remaining animals reached the criterion within that

1032 acquisition or reversal phase. The different phases are the first acquisition (Acq1), the

1033 first reversal (Rev1), the second acquisition (Acq2) and the second reversal (Rev2).

1034 Significant differences between the last session of Acq1 and the first session of Rev1 and

1035 between the last session of Acq2 and the first session of Rev2 are indicated (*** $p <$
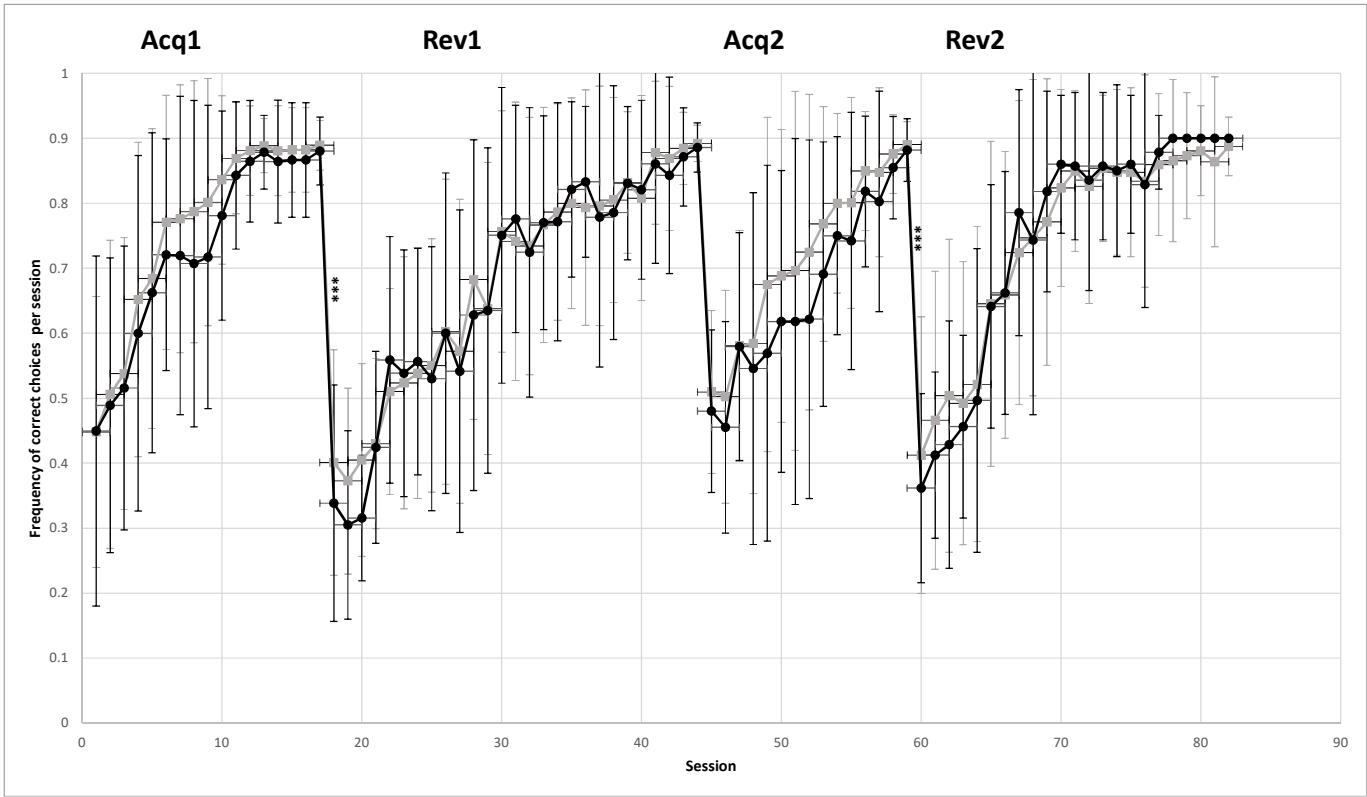
1036 0.001).

1037

1038

1039 **Figure 3. Learning capacities of crib-biters and control horses.** Number of trials

1040 until the learning criterion was attained for all crib-biters (group CB: $N = 6$ horses, in

1041 grey) and control horses (group C: $N = 7$ horses, in white), for each Phase (Acq1, Rev1,

1042 Acq2, Rev2). The different phases are the first acquisition (Acq1), the first reversal

1043 (Rev1), the second acquisition (Acq2) and the second reversal (Rev2). The learning

1044 criterion was fixed at six correct trials in a row. The black dots indicate the means.

1045 Significant differences between Phase are indicated as * $0.05 \leq p < 0.01$; ** $0.01 \leq p <$

1046 0.001.

1047

45 cm

45 cm

61 cm

100 cm

40 cm

160 cm

40 cm