

Author Identifying Information

How to stay perfect: the role of memory and behavioural traits in an experienced problem and a similar problem.

Pizza Ka Yee Chow*, Stephen E. G. Lea, Natalie Hempel de Ibarra, Théo Robert

Centre for Research in Animal Behaviour

Psychology Department

University of Exeter

Corresponding author: *kyc202@exeter.ac.uk

1 **ABSTRACT**

2 When animals encounter a task they have solved previously, or the same problem appears in
3 a different apparatus, how does memory, alongside behavioural traits such as persistence,
4 selectivity and flexibility, enhance problem-solving efficiency? We examined this question
5 by first presenting grey squirrels with a puzzle 22 months after their last experience of it (the
6 recall task). Squirrels were then given the same problem presented in a physically different
7 apparatus (the generalisation task) to test whether they would apply the previously learnt
8 tactics to solve the same problem but in a different apparatus. The mean latency to success in
9 the first trial of the recall task was significantly different from the first exposure but not
10 different from the last exposure of the original task, showing retention of the task. A
11 neophobia test in the generalisation task suggested squirrels perceived the different apparatus
12 as a different problem, but they quickly came to apply the same effective tactics as before to
13 solve the task. Greater selectivity (the proportion of effective behaviours) and flexibility (the
14 rate of switching between tactics) both enhanced efficiency in the recall task, but only
15 selectivity enhanced efficiency in the generalisation task. These results support the interaction
16 between memory and behavioural traits in problem solving, in particular memory of task
17 specific tactics that could enhance efficiency. Squirrels remembered and emitted task-
18 effective tactics more than ineffective tactics. As a result, they consistently changed from
19 ineffective to effective behaviours after failed attempts at problem-solving.

20 *Keywords:* problem solving, generalisation, positive transfer, behavioural traits, memory,
21 squirrels, problem-solving efficiency

22 INTRODUCTION

23 Problem solving ability, the ability to overcome obstacles and achieve a goal, has
24 been shown to bring advantages on various measures of fitness. For example, successful
25 problem solvers lay larger clutches of eggs and have increased mating success (review by
26 Boogert et al. 2010; Cauchard et al. 2013, Cole et al. 2012, Keagy et al. 2009, Preiszner et al.
27 2017, but also see Isden et al. 2013). Such impacts on fitness provide a justification for
28 extending investigation to mechanisms that are correlated with problem-solving, such as
29 behavioural traits. However, investigations in such area have only recently begun (Reader &
30 Laland 2003; review by Guez & Griffin 2016).

31

32 An increasing number of studies have now shown that certain behavioural traits are
33 important for problem solving. The key behavioural traits include persistence, motor diversity
34 ('behavioural variety' or 'exploration diversity'), selectivity (or 'behavioural selectivity') and
35 flexibility (e.g. Benson-Amram and Holekamp 2012; Benson-Amram et al. 2013; Biondi et
36 al. 2008; Cauchard et al. 2013; Chow et al. 2016; Diquelou et al. 2016; Griffin et al. 2014;
37 Griffin and Diquelou 2015; Overington et al. 2011; van Horik and Madden 2016; Thornton
38 and Samson 2012). Each of these traits has been shown to relate to problem-solving
39 performance in different ways. For example, increased selectivity enhanced problem-solving
40 efficiency, as measured by decreased latency to solve a problem, in black-throated monitor
41 lizards, *Varanus albigularis albigularis* (Manrod et al. 2008), in Atlantic cod, *Gadus morhua*
42 *L.* (Milot et al. 2014) and in grey squirrels, *Sciurus carolinensis* (Chow et al. 2016).
43 Increased motor diversity and persistence facilitated success rate in spotted hyenas, *Crocuta*
44 *crocuta* (Benson-Amram and Holekamp 2012; Benson-Amram et al. 2013) and Indian
45 mynas, *Sturnus tristis* (Griffin et al. 2014), and enhanced problem-solving efficiency in grey
46 squirrels (Chow et al. 2016). Increased flexibility, the rate of switching between tactics,

47 however, decreased solving efficiency in grey squirrels, as a result of decreased selectivity
48 (Chow et al. 2016).

49

50 The traits associated with success in a single novel complex task, as discussed above,
51 do not appear to have a fixed hierarchy of importance. A given trait may be salient in relation
52 to a particular task, its context, and perhaps the species involved. For example, selectivity
53 seems to be particularly important when animals return to a learned task after a delay, or
54 experience a new task that resembles one they had experienced previously. With regard to
55 returning to previously experienced task, selectivity appears to be an important factor in the
56 success of captive lions, *Panthera leo*, in solving a suspended puzzle box up to seven months
57 after experiencing it (Borrego and Downing 2016), in the success of goats, *Capra hircus*, in
58 solving a two-step food box challenge ten months after first experiencing it (Briefer et al.
59 2014), and in the success fat-tailed dunnarts, *Sminthopsis crassicaudata*, when they re-
60 experience a visual reversal learning task (Bonney and Wynne 2002). With regard to
61 situations where animals can apply previously learned tactics in a different (or novel)
62 apparatus through generalisation, selectivity has been shown to be important in stimulus
63 generalisation (e.g. Cuvo 2003), categorisation (e.g. Reichmuth Kastak and Schusterman
64 2002) and the generalisation of tool use (e.g. Macellini et al. 2012). Such success in
65 transferring previously learnt tactics to a different task depends on individuals being able to
66 recognise that it is the same (or a similar) task, and to recall the tactics that they learned in a
67 previous task.

68

69 The ability to recall and employ previously learned tactics to solve a given task after a
70 lapse of time, or to solve a similar task, highlights the interaction between selectivity and
71 cognitive mechanisms such as learning and memory in facilitating problem-solving

72 efficiency. The level of task information retained may affect the way that behavioural traits
73 vary across trials when individuals re-experience the same task. Hypothetical situations
74 include:

- 75 1) The ideal outcome, where individuals would immediately perform the effective
76 tactic. In this situation, additional motor diversity (use of alternative tactics),
77 flexibility (switch to another tactic), or persistence (attempts) in solving an
78 experienced problem would not be necessary.
- 79 2) The worst-case scenario, where individuals have completely forgotten the task and
80 are learning the task as if at their first experience. In this case, we would expect
81 individuals to increase selectivity (Chow et al. 2016; Manrod et al. 2008; Millot et
82 al. 2014), persistence (Biondi et al. 2008; Chow et al. 2016), and motor diversity
83 (Benson-Amram and Holekamp 2012; Griffin and Diquelou 2015) with increased
84 experience. Flexibility should not vary with increased experience (Chow et al.
85 2016).
- 86 3) The intermediate case, where individuals have retained some but not all
87 information relevant to a previously experienced task. In this case, the variation of
88 traits with trials would depend on how much information they have retained from
89 the past. In this situation, animals may show different types of retrieval strategies.
90 Two strategies have been described: an information-based and a guessing-based
91 strategy. In the information-based strategy, individuals retrieve effective tactics
92 based on the familiarity of the task and retained task information (Malmberg and
93 Xu 2007). Such a strategy implies that there will be switching between retained
94 tactics until asymptotic efficiency is again achieved. In a guessing-based strategy
95 there will inevitably be errors, but surprisingly it has been shown in humans that
96 these enhance retention, because guessing may lead to more elaborated

97 information processing of correct responses (Yan et al. 2014). Accordingly, if
98 either of these retrieval strategies is used, observed flexibility should increase
99 across trials and should enhance solving efficiency. However, an essential
100 difference between the two strategies lies in the way tactics change as the problem
101 is solved. In the information-based retrieval strategy, changes should not be
102 completely ‘random’ (i.e. behaviours in an individual’s repertoire should not all
103 have equal probability of being exhibited) whereas they should have in the
104 guessing-based strategy.

105

106 Here, we examined how memory, alongside behavioural traits, contributes to enhance
107 problem-solving efficiency by giving five grey squirrels, firstly a previously experienced task
108 22 months after they had last experienced it (hereafter, the “recall task”), and secondly a task
109 requiring a previously successful action to be performed in a physically different apparatus
110 (hereafter, the “generalisation task”). The squirrels had learned a specific solution for solving
111 a puzzle box involving food reward in the laboratory (hereafter, the “original task”) 22
112 months before the present experiments (Chow et al. 2016). We used Chow and colleagues’
113 methods to measure four behavioural traits, persistence (rate of attempts), selectivity
114 (proportion of effective behaviours), motor diversity (rate of emitting different types of
115 tactics) and flexibility (rate of switching between tactics after a failed attempt), on a trial-by-
116 trial basis. We chose grey squirrels for this study because they have demonstrated high
117 behavioural flexibility, in a number of situations, including a serial spatial reversal learning
118 task (Chow et al. 2015), a colour reversal learning task (Chow et al. 2017), and a problem
119 solving task (Chow et al. 2016). Grey squirrels are also known to have good long-term, at
120 least in the spatial domain: they are scatter-hoarders that cache thousands of nuts during the
121 autumn (Thompson and Thompson 1980), and they are able to re-locate their own caches

122 (Jacobs and Liman 1991) and artificial caches (Macdonald 1997) after long intervals of time.
123 While there is always a possibility that memory ability is domain specific, it is reasonable to
124 assume they would be able to remember the solutions to a problem over an extended period.
125 If this is the case, then squirrels would not only be able to solve the task when they re-
126 experience the original task, but they would also show significantly shorter latency to solve
127 the task compared with the first experience of the original task, as in the experiments on lions
128 (Borrego and Dowling 2016) and goats (Briefer et al. 2014) cited above.

129

130 We further explored what retrieval strategy squirrels were employing in these two
131 tasks by examining whether squirrels exhibited non-random changes in tactics or not. If
132 squirrels have completely retained the learned task tactics they used to solve the original task,
133 we predict that selectivity would remain at its highest (close to 1 as a proportion) whereas
134 motor diversity, flexibility and persistence would remain at their lowest, and none of these
135 traits would vary with increased experience (see situation 1 above). Such high selectivity
136 would be expected to be one key behavioural trait that enhances efficiency in both tasks.
137 However, as discussed above, if individuals have completely forgotten the task or only
138 retained some information about the original task, then we would observe characteristic
139 variations of these traits with increasing experience in the new situation (see situations 2 and
140 3 above).

141

142 **METHODS**

143 *Ethical notes*

144 This study was approved by the Ethical Review Group at the University of Exeter (no.
145 2012/ 533) and the experiment was carried out in accordance with the Association for the
146 Study of Animal Behaviour and Animal Behaviour Society guidelines and UK law.

147

148 *Subjects and Housing*

149 Five squirrels, living in the laboratory, participated in this study. They were named
150 Arnold, Leonard, Sarah, Simon and Suzy, and included two females and three males. Their
151 mean age was 6 years; see Supplementary Materials Table S1 for further information on each
152 squirrel. The temperature in their housing was controlled at a constant 19 °C, and lighting
153 was on a 12:12 day-night cycle, with all testing conducted during the light period. The
154 squirrels were housed in large cages that were constructed using metal mesh. In each cage,
155 there was a sliding metal door connected to an overhead tunnel. Only one squirrel was
156 allowed access to the test room at a time for this experiment. A metal mesh divided the test
157 room into two equally large cages (each 1.5 x 1.8 x 2.5 m). The front and ceiling of the cages
158 were metal mesh whereas the side and the back of the cages were solid concrete wall. One
159 cage had a touch screen panel, set 2 m above the floor as reported in Chow et al. (2017). A
160 camera (Panasonic SWHD-90) was set up in the adjacent cage to capture all behavioural
161 responses during the experiment. Further details of the housing and test room set-up are given
162 by Hopewell et al. (2010). All the squirrels had similar experimental histories in cognitive
163 tasks (see Table S1 in supplementary materials for details). Within the 22 months prior to the
164 present study, the squirrels did not interact with the puzzle box used by Chow et al. (2016) or
165 any similar problem-solving task, nor were they exposed to similar designs as enrichment;
166 they did participate in a serial spatial reversal learning task, as reported by Chow et al.
167 (2015). The squirrels were not food deprived, and water was provided *ad libitum*. We ensured
168 squirrels' motivation by using rewards (hazelnuts) that were different from their daily diet
169 (seeds, fresh fruit and vegetable). Doors allowing the squirrels to enter the test room by the
170 overhead tunnel from their home cages were opened during the times of day when they were

171 most active (0700-1100 and 1500-1800), and tests were carried out when a squirrel entered
172 the test room spontaneously. Data collection took place between May and July, 2015.

173

174 *Puzzle box for the recall task*

175 Fig. 1a shows the puzzle box that was presented to squirrels by Chow et al. (2016), 22
176 months before the present experiment; we used the same box for this experiment. The box
177 was a transparent Plexiglas cuboidal box (Length 25 x Width 25 x Height 25 cm) that had 10
178 holes on each side. 10 levers (each lever 29.8 x 1.5 x 0.5 cm thickness), five functional
179 (baited with hazelnuts) and five non-functional (without hazelnuts), were inserted across the
180 box through holes in opposite sides. The holes (2 cm x 0.9 cm W x H) on the box were
181 designed to be larger than the thickness of a lever (0.5 cm), so that squirrels could see and
182 smell the nuts but could not directly reach them after a lever was inserted. At one end of each
183 lever there was a three-sided container, and this was positioned just inside the box. Four
184 wooden legs were used to support the box, creating a 4.5 cm gap through which squirrels
185 could obtain the hazelnuts once they fell out of the containers. Although squirrels could use
186 many types of behaviours to solve the task, the apparatus was designed so specific behaviours
187 were effective (i.e. the most efficient way) for obtaining a nut and specific other behaviours
188 could not solve the task. The specific effective behaviours were pushing the ‘near-end’ of a
189 lever and pulling the ‘far-end’ (near- and far- end refer to proximity to the hazelnut bait)
190 while the specific ineffective behaviours were pulling the ‘near-end’ of a lever and pushing
191 its ‘far-end’.

192

193 *Puzzle box for the generalisation task*

194 Fig. 1b, c show the apparatus used in the generalisation task. It was a transparent
195 puzzle box in the shape of a four-sided triangular prism (triangle front: 35 x 19 x 18 cm;

196 Length x Width x Height, rectangular side: 25 x 20 cm) with five levers inserted. The puzzle
197 box had completely different physical characteristics and colour than the one used in the
198 recall task, but it still involved moving levers, so that we could examine whether squirrels
199 applied the learned effective and ineffective tactics to obtain the nuts. The length of the levers
200 was shorter than in the recall task and both ends of each lever were slightly curved (lever
201 dimensions: 23.5 x 2 x 0.2 cm L x W x H). The generalisation box had 5 holes (2 x 0.9 cm)
202 on each side, which were horizontally but not vertically aligned with holes on the opposite
203 side. Because squirrels showed a strong preference for choosing the functional levers (with
204 hazelnuts) both in the original (Chow et al. 2016) and the recall tasks (see results section), we
205 further increased the difference between the recall and generalisation task by including only
206 functional levers. As Fig. 1c shows, both lever ends protruded 1.5 cm out of the box. The box
207 was supported by four wooden legs, creating a 3.5 cm gap from its base. The base of the box
208 (32 x 10 x 3 cm) was a wooden sloped platform (in silver grey colour) which allowed a nut to
209 roll down once it had fallen. As in the recall task, squirrels could see and smell the rewards
210 but could not reach them directly. Squirrels were able to emit the same effective and
211 ineffective behaviours on each lever to obtain a nut: pulling the near-end or pushing the far-
212 end of a lever was ineffective, so they had to push the near-end or pull the far-end.

213 *Figure 1*

214 *Procedures*

215 Squirrels first participated in the recall task, so we could examine whether they
216 remembered the puzzle box they had experienced 22 months ago. The generalisation task was
217 presented six days later so as to examine whether squirrels could transfer the same effective
218 behaviours to a physically different box. We kept the same procedures as in Chow et al.
219 (2016) for both the recall and the generalisation tasks; squirrels were tested individually to
220 avoid confounding factors such as stimulus enhancement or social learning in the task. Each

221 squirrel participated in three blocks of four trials in each task (for a total of 12 trials), with a
222 one day break between each block (for a total of 14 testing days). In each trial, we placed the
223 box at the centre of the test room. A trial started when squirrels touched or manipulated any
224 part of the box. The trial ended when squirrels completed the task by obtaining all the
225 rewards, when they had not touched the apparatus for 15 minutes, or when 45 minutes had
226 elapsed, whichever came first. If a squirrel did not respond, we repeated the trial the next day.
227 This only happened with one squirrel, Suzy, in one trial in the recall task. After every trial,
228 we removed the odour left on the apparatus using disinfectant-impregnated cleaning wipes.
229 We also used wipes after baiting in order to minimise any human scents left on the apparatus.
230 For both tasks, the orientation of the apparatus and the direction the levers faced was pseudo-
231 randomised between trials. For the recall task, we additionally randomised whether a given
232 lever was functional or not. A single success at solving the problem was defined as a squirrel
233 causing a functional lever and/or a nut to drop. A trial therefore normally consisted of five
234 successes.

235

236 *Latency measurements*

237 *Contact latency.* For both the recall task and the generalisation task, we measured the latency
238 from when a squirrel entered the test room until it first used its nose or paws to touch the
239 apparatus. We measured the contact latency on the last trial of the recall task and on the first
240 trial of the generalisation task as neophobia. This allowed us to test whether the squirrels
241 perceived the pyramid-shaped apparatus as a novel stimulus in the generalisation task.

242

243 *Success latency.* We also measured the time taken to obtain each reward; this was used as a
244 measure of problem solving efficiency. Latency was timed from the moment when a squirrel
245 started to manipulate a functional lever until the nut it contained dropped. Not every

246 manipulation of a functional lever led to success, but the time spent in unsuccessful
247 manipulation on it was still included. For each trial, we summed all the latencies on
248 functional levers and then divided this total success latency by the number of functional
249 levers that a squirrel solved during that trial, to obtain the mean latency to each success.

250

251 *Measurement of behavioural traits*

252 The four behavioural traits, persistence, motor diversity, selectivity, and flexibility were
253 measured using methods standardised by Chow et al. (2016). The first author analysed all
254 behaviours from videos using the software Adobe Premiere Pro CS6; this allowed us to
255 analyse behavioural data on a frame-by-frame basis. The behavioural measures of each trait
256 co-vary with one another and it is therefore necessary to tease them apart analytically to
257 avoid multicollinearity. The measures also need to be normalised in some way, since the
258 longer a trial lasts, the more opportunity there is for a behaviour to be performed.

259 Accordingly, rates of occurrence of behaviours rather than raw counts were used, as in
260 previous experiments (e.g. Biondi et al. 2008; Chow et al. 2016; Griffin et al. 2014; Griffin
261 and Diquelou 2015; Papp et al. 2015). All measurements were taken trial-by-trial for each
262 task (12 trials). For the recall task, we recorded the measures on the functional levers only, to
263 allow direct comparison with the generalisation task in which only functional levers were
264 used.

265

266 *Selectivity.* Selectivity was measured as the proportion of effective behaviours. We counted
267 the number of effective behaviours (i.e. either pushing the near-end or pulling the far-end of a
268 functional lever) and the number of ineffective behaviours (i.e. either pushing the far-end or
269 pulling the near-end of a functional lever) in each trial. Then we divided the number of
270 effective behaviours by the total number of effective and ineffective behaviours for that trial.

271

272 *Persistence.* Persistence has been used to assess motivation (e.g. Biondi et al. 2008; Chow et
273 al. 2016; Griffin et al. 2014). We measured persistence as the rate of attempting to solve the
274 problem. An attempt was recorded whenever a squirrel used any of its body parts to
275 manipulate a functional lever, regardless of whether the manipulation was exhibited as
276 effective or ineffective behaviours directed at the box. A new attempt was counted when
277 squirrels switched to a different functional lever or when the squirrel returned to
278 manipulating the same lever after at least one second without having its body in contact with
279 the lever. We counted the total number of attempts in each trial on all functional levers and
280 then divided this number by the total success latency as defined above.

281

282 *Motor diversity.* Motor diversity was measured as the rate of using different tactics in solving
283 the problem. We used Chow et al. (2016)'s Table 1 to code the tactics that squirrels used
284 within solving a functional lever. Nine types of behaviour were coded: pull, push in, push up,
285 push down, tilt up, claw, lick, shake and any of two or more of these behaviours occurring
286 simultaneously (combined behaviours). We obtained the rate of motor diversity for each trial
287 by counting the number of types of behaviours that a squirrel exhibited during a trial (ranged
288 from 1-9) and then dividing this number by the total success latency for the trial, as defined
289 above.

290

291 *Flexibility.* Flexibility was measured as the rate of switching between tactics. A switch was
292 counted whenever a squirrel changed from any of the tactics listed in motor diversity to a
293 different one, regardless of whether either of the tactics involved was effective. We first
294 counted the number of switches between tactics and then divided this number by the total
295 success latency, as defined above, to obtain the rate of flexibility in each trial. To further

296 examine squirrels' retrieval strategies, we measured the mean number of 'non-productive'
297 switches (i.e. switches from effective to ineffective behaviours) across functional levers.

298

299 *Data analysis*

300 We used R v.3.3.2 (R Core Team, 2016) to analyse all behavioural data. All significance
301 levels reported are two-tailed and were considered as significant when $P < 0.05$.

302

303 For the recall task, we used exact binomial tests to examine whether each squirrel was
304 significantly more likely to direct attempts at functional levers (baited with hazelnuts) than at
305 non-functional levers (without hazelnuts). We then pooled the P -values using Fisher's
306 formula $\chi^2 = -2 \sum \ln(P)$ (Sokal and Rohlf 1995 p. 794). For the generalisation test, we used a
307 Wilcoxon signed-rank test to assess differences in contact latency from the recall test, and
308 Spearman's correlation to examine relationships between contact latency and mean success
309 latency on the first trial.

310

311 We used Generalized estimating equations (GEE) with exchangeable 'working' correlation
312 (Hardin and Hilbe 2003) to investigate 1) whether the mean latency to each success in the
313 first trial of the recall task differed from the mean latency to each success in the first trial and
314 the last trial of the original task; 2) whether the mean latency to each success in the last trial
315 of the recall task differed from the mean latency to each success on the first trial of the
316 generalisation task; 3) how the mean latency to each success varied across trials in each task;
317 4) how each behavioural trait (rate of attempts, rate of flexibility, rate of motor diversity and
318 proportion of effective behaviours) varied across trials; and 5) how the behavioural traits
319 contributed to increasing efficiency in the recall task and in the generalisation task,
320 separately. GEE is a quasiparametric statistical test for model estimates. Because small

321 sample size leads to underestimation of the variance of parameter estimates, we obtained the
322 *P*-values using the package ‘geesmv’ (Wang 2015), which adjusted the modified ‘sandwich’
323 variance estimator (Wang and Long 2011) for estimating the variance–covariance matrix of
324 the parameter estimates. This modified variance has been shown to be robust for experiments
325 that have very small sample size with each individual completing all trials, as in our case.

326

327 We used Pearson correlations to explore the relationships between covariates before model
328 testing. Attempt rate and motor diversity were highly correlated in the recall tasks ($r = 0.78$)
329 and in the generalisation task ($r = 0.86$). High correlation was also shown between attempt
330 rate and selectivity ($r = 0.53$) for the generalisation task. To avoid confusion in interpreting
331 the results due to multicollinearity, and, in line with the primary focus of this study on
332 memory for task effective behaviours, we selected variables for model estimations as follows.
333 We included attempt rate, selectivity, switch rate and trial number for the recall task, but
334 excluded attempt rate was excluded from the model estimation for the generalisation task,
335 because, given the high level of accuracy, it was confounded with the other traits.

336

337 **RESULTS**

338 *Performance in the puzzle box recall task*

339 Fig. 2a shows the median across squirrels of mean latency to success in the first trial (8
340 seconds) and the last trial (2 seconds) of the original task. Fig. 2b shows the median across
341 squirrels of mean success latency in the first trial of the recall task (3 seconds). Latency on
342 the first trial of the recall task as significantly different from its value on the first trial of the
343 original task (GEE $\chi^2_1 = 4.12$, $P = 0.032$), but not different from its value on the last trial of
344 the original task ($\chi^2_1 = 2.65$, $P = 0.104$). These results indicate some retention for the task 22

345 months after the last experience with this box. The latency to each success did not vary
 346 significantly across trials in the recall task ($\chi^2_1 = 0.30, P = 0.587$).

347 *Figure 2*

348 Fig. 3 shows that on the first exposure of the recall task after 22 months, squirrels made more
 349 attempts to solve functional levers (with hazelnuts) than non-functional levers (without
 350 hazelnuts). This preference for solving functional levers was significantly above chance
 351 (pooled binominal tests: $\chi^2_{10} = 49.25, P < 0.001$). Although squirrels could smell the nuts,
 352 their behaviours suggested that they often first used vision to approach a functional lever
 353 before using olfaction, presumably to assess the quality of a nut rather than locating the nut
 354 (See supplementary video VS1-S2).

355 *Figure 3*

356 Fig. 4a shows the variation of selectivity (i.e. the proportion of effective behaviours) across
 357 trials in the recall task. Selectivity did not vary significantly across trials ($\chi^2_1 < 0.001, P =$
 358 0.95). Fig. 4 b, c, d show the variations of persistence, motor diversity and flexibility,
 359 respectively in the recall task. With increased experience, squirrels significantly increased
 360 flexibility ($\chi^2_1 = 6.42, P = 0.011$), but not persistence ($\chi^2_1 = 0.05, P = 0.826$) or motor
 361 diversity ($\chi^2_1 = 0.67, P = 0.414$). The median across squirrels of mean number of non-
 362 productive switches (i.e. switches from effective to ineffective behaviours) was 0.6 and 0.4 in
 363 the first trial and the last trial of the recall task, respectively.

364 *Figure 4*

365 *Performance in the generalisation task*

366 We first verified that squirrels perceived the generalisation puzzle box as a different stimulus.
 367 We compared the latency to contact the apparatus in the last trial of the recall task with the
 368 first trial of the generalisation task. All squirrels took longer to approach the puzzle box in the
 369 first trial of the generalisation task (Median of mean latency = 23s) than in the last trial of the

370 recall task (Median of mean latency = 11s) and this difference was significant (Wilcoxon
371 signed-rank test: $W = 5$, $Z = -2.02$, $P = 0.043$). Neophobic responses toward the
372 generalisation apparatus was not correlated with the latency to each success in the first trial of
373 the generalisation task ($r_s = 0.30$, $P = 0.623$).

374

375 Fig. 2b shows the median of mean latency to each success across 12 trials in the
376 generalisation task. All squirrels took significantly less time to each success in the first trial
377 of the generalisation task (Median of mean latency = 2s) than in the first trial of the recall
378 task ($\chi^2_1 = 4.39$, $P = 0.036$). However, the latency to each success was not significantly
379 different between the last trial of the recall task (Median of mean latency = 1s) and the first
380 trial of the generalisation task ($\chi^2_1 = 0.67$, $P = 0.413$). Fig. 4a shows that squirrels consistently
381 showed a high proportion of effective behaviours (median proportion across squirrels = 0.93)
382 in the first trial of the generalisation task. This proportion did not vary significantly across
383 trials ($\chi^2_1 = 0.38$, $P = 0.536$), which suggested they quickly perceived the problem as the same
384 despite the changed appearance of the task. Fig. 2b shows that the latency to each success did
385 not significantly vary across trials ($\chi^2_1 = 1.61$, $P = 0.205$). Fig. 4b, c, and d show the variation
386 of persistence, motor diversity, and flexibility during the generalisation task. None of the
387 three behavioural traits varied significantly across trials (persistence: $\chi^2_1 = 1.12$, $P = 0.290$;
388 motor diversity: $\chi^2_1 = 0.54$, $P = 0.461$; flexibility: $\chi^2_1 = 0.06$, $P = 0.808$). The median across
389 squirrels of the mean number of non-productive switches (i.e. switches from effective to
390 ineffective behaviours) was 0.2 in both the first trial and the last trial of the generalisation
391 task.

392

393 *Factors associated with problem-solving efficiency in the recall task*

394 Table 1 (left panel) shows the results of the correlational analysis for the recall task. Model 1
 395 shows that only selectivity, measured as the proportion of effective behaviours, and
 396 flexibility, measured as the rate of switching to other tactics after a failed attempt, were
 397 significantly associated with efficiency. Specifically, high efficiency was related to a high
 398 proportion of effective behaviours ($\chi^2_1 = 25.39, P < 0.001$) and a high switch rate ($\chi^2_1 = 6.17,$
 399 $P = 0.013$). As with our previous finding in Chow et al. (2016), the non-significant effects of
 400 persistence, measured as the rate of attempts, and experience, recorded as trial number,
 401 suggest that selectivity and flexibility mediated the effect of experience and persistence on
 402 efficiency. Therefore, we ran two separate analyses to examine two covariates, experience
 403 and persistence, in relation to response variable, selectivity, in one model (Model 2) and
 404 flexibility in another model (Model 3). Model 2 (left panel) shows that selectivity was related
 405 to persistence ($\chi^2_1 = 22.65, P < 0.001$), but not experience ($\chi^2_1 < 0.01, P = 0.989$); higher
 406 persistence was associated with higher selectivity. Model 3 shows that flexibility was
 407 significantly related to both persistence ($\chi^2_1 = 6.03, P = 0.014$) and experience ($\chi^2_1 = 4.85, P =$
 408 0.028); increased flexibility was positively correlated to higher persistence and increasing
 409 experience. We ran a final model (Model 4) to examine whether persistence increased across
 410 trials and results showed it did not ($\chi^2_1 = 0.05, P = 0.826$). These results imply that
 411 persistence, a trait that is not affected by experience, affects efficiency on the recall test, but
 412 does so indirectly through increasing selectivity and flexibility.

413 *Table 1*

414 *Factors associated with problem-solving efficiency in the generalisation task*

415 Table 1 (right panel) shows the results for the generalisation task. Model 1 shows that the
 416 latency to each success was significantly related to selectivity ($\chi^2_1 = 12.05, P < 0.001$); a
 417 higher proportion of effective behaviours was associated with greater efficiency. Model 2

418 (right panel) shows that in this task, selectivity was not related to experience ($\chi^2_1 = 0.37$, $P =$
419 0.544) or flexibility ($\chi^2_1 = 0.06$, $P = 0.812$). The final analysis (Model 3) shows that flexibility
420 was not significantly related to experience ($\chi^2_1 = 0.06$, $P = 0.808$).

421

422 *Total effect of behavioural traits on efficiency*

423 For each Model, we obtained the effect sizes for each trait (Path β). We then calculated the
424 total effect (Total β) for each trait in Table 1. For both tasks, selectivity showed the highest
425 effect on efficiency (in the recall task: $\beta = -0.74$ and in the generalisation task: $\beta = -0.88$).

426

427 DISCUSSION

428 In this study, we examined how memory and behavioural traits improved problem-
429 solving efficiency when squirrels re-experienced a task that reappeared after a substantial
430 time had elapsed (the recall task), and when squirrels encountered the same problem in a
431 different apparatus (the generalisation task). We showed that all squirrels retained some
432 information from previous experience by showing a high proportion of effective behaviours
433 (Fig. 4a), indicating that their retention of the task extended to the specific tactics that were
434 effective in solving it (i.e. pushing at the near-end or pulling at the far-end of a lever). Such
435 information about task tactics facilitated squirrels' solution of the problem. All squirrels also
436 successfully transferred the same tactics to solve the problem when it appeared in a
437 physically different apparatus. Aside from memory, selectivity and flexibility were important
438 factors in increasing efficiency in the recall task, whereas only selectivity affected efficiency
439 in the generalisation task.

440

441 As discussed in the introduction, the level of retained task information may affect how
442 traits vary in the recall and the generalisation tasks. In our case, other than flexibility in the

443 recall task, none of the traits varied with increased experience, suggesting that the squirrels
444 remembered the task almost perfectly (Fig.4 a-d). The fact that squirrels consistently showed
445 a high proportion of effective behaviours (Fig. 4a) is in line with our prediction that memory
446 and selectivity are tightly related to each other and their interaction is the key trait to
447 advanced problem-solving efficiency in both tasks (Table 1). In both tasks, memory of which
448 tactics were effective may reflect a series of associations that have been formed in the past,
449 for example, an association between the cues, the behaviours and the rewards formed by
450 operant conditioning would allow the squirrels to promptly locate functional levers (Fig. 3),
451 emit the effective behaviours, and obtain the hazelnut.

452

453 Another behavioural trait, flexibility, measured as the rate of switching between
454 tactics after a failed attempt, was also found to be an important trait for achieving efficiency
455 in the recall task, but not in the generalisation task (Model 1). Flexibility also varied with
456 increased experience in the recall task (Fig. 4d). In the introduction, we discussed two
457 possible retrieval strategies under a recall situation: individuals would either explore the
458 effective tactics based on retained information (information-based) or explore all possible
459 tactics (guessing-based strategy). In both tasks, squirrels showed more effective behaviours
460 than ineffective behaviours, suggesting they were using an ‘information-based’ strategy so
461 that if they did emit an ineffective behaviour, they quickly switched to an effective one.
462 These results show how flexibility is related to memory, because a productive change of
463 tactics would involve remembering the correct tactic, and may also lead to the reinforcement
464 of the effective behaviours during the recall task. It follows that productive changes of tactics
465 allowed the squirrels to achieve efficiency in the recall task and to apply the same tactics
466 from the first trial of the generalisation task (Fig. 4a).

467

468 The final trait of particular interest in this study is persistence, measured as the rate of
469 making attempts to solve the problem. The role of persistence in solving novel problems is
470 well established (Benson-Amram and Holekamp 2012; Biondi et al. 2008; Chow et al. 2016;
471 Griffin et al. 2014, Papp et al. 2015, Thornton and Samson 2012; van Horik and Madden
472 2016). Persistence may largely reflect the motivation of individuals (e.g. Biondi et al. 2008;
473 Chow et al. 2016). When squirrels re-experience the same task, such motivation embraces
474 various aspects of problem solving, including goal-orientation to food reward (Fig. 3),
475 changes to another tactic after a failed attempt (Model 2 and Model 3), and motivation to
476 emit the effective behaviours to increase problem-solving efficiency (Table 1 Model 2).
477 However, unlike others who have argued that persistence may not be involved in any
478 cognitive process that could lead to problem-solving success (e.g. van Horik and Madden
479 2016; Thornton and Samson 2012), we suggest that such persistence may well interact with
480 cognitive processes in several ways. For example, paying attention to the functional cues or
481 properties of the task has been demonstrated in kea and crows using tools to solve a problem
482 (e.g. Auersperg et al. 2011; St Clair and Rutz 2013; Werdenich and Huber 2006). In our case,
483 squirrels may pay attention to cues such as the levers that contain hazelnuts to locate which
484 lever to solve. But unlike what has been found in tool-use studies, the use of cues did not
485 develop with increased experience during problem solving. Squirrels showed an immediate
486 strong preference to contact functional levers rather than non-functional levers, both when
487 they first encountered this puzzle box 22 months prior to this study (Chow et al. 2016) and in
488 the first trial of the recall task (Fig. 3). These results imply that squirrels quickly focused their
489 attention on the reward and reward-related components of the apparatus (levers) from their
490 first encounter with the puzzle box. Such attention may be developed from previous handling
491 or knowledge about the objects or food (e.g. Bird and Emery 2009; Taylor et al. 2010). The
492 effect of persistence was positively mediated by flexibility and selectivity in the recall task

493 (Model 3); that is, persistence was indirectly related to the latency to solve the task through
494 its effect on selectivity and flexibility. Given that persistence was also highly correlated with
495 selectivity in the generalisation task, we suggest that higher motivation led the squirrels to
496 emit effective behaviours in this task. Taking these trends together, one could deduce that
497 persistence is also related to retrieval from memory, which may also explain why persistence
498 did not show a significant increase (or decrease) across trials in either task (Fig. 4b).

499

500 The sample size in the present study was limited, and hence, we had limited degrees
501 of freedom available to explore other interactions between traits on problem-solving
502 efficiency; and we do need to be cautious in generalising from five squirrels to the whole
503 species. Nevertheless, we have shown that learning tactics for a given task can improve future
504 problem-solving efficiency if individuals are able to recall these tactics when they revisit the
505 same task, or when it is possible to apply them in a different apparatus. In these situations,
506 learning plays a minimal role whereas long-term memories of the effective tactics, along with
507 other factors are important for increasing efficiency. In a broader context, these results
508 highlight the mechanisms, including cognitive capacity and behavioural traits, that are
509 correlated with problem-solving ability and enable animals to achieve better problem solving
510 performance. This provides information about why these mechanisms evolved together. In
511 turn, it should be possible to investigate which of these factors are general across a range of
512 tasks, thereby making it plausible to try to obtain a measure of general cognitive ability ('g')
513 for an individual. It should also be possible to study which of these factors might differ
514 between species - for example, between the Eastern grey squirrel and the Eurasian red
515 squirrel, which it has largely replaced in Britain and Ireland. Ultimately, we would hope to be
516 able to highlight the factors that explain the varied fitness consequences associated with

517 cognitive capacity, for example, the relationship between 'g' and the success of some species
518 as invaders of new environments.

519

520 Acknowledgements

521 We thank D. Taylor for building the puzzle box. C. Soper and L. Goss for taking care of the
522 squirrels. We also thank L. Leaver for developing the captive squirrel facility.

Compliance with Ethical Standards

This study was not supported by any funding organisation.

Conflict of Interest

All authors declare that there is no conflict of interest.

Ethical approval

This study was approved by the Ethical Review Group at the University of Exeter (no. 2012/533). Squirrels were treated in accordance with Association for the Study of Animal Behaviour guidelines on animal welfare and UK law.

References

- Auersperg AMI, von Bayern AMP, Gajdon GK, Huber L, Kacelnik A (2011) Flexibility in Problem Solving and Tool Use of Kea and New Caledonian Crows in a Multi Access Box Paradigm. PLoS ONE 6:e20231.
- Bonney KR, Wynne CD (2002) Visual discrimination learning and strategy behaviour in the fat-tailed dunnart (*Sminthopsis crassicaudata*). J Comp Psychol 116:55-62.
- Benson-Amram S, Holekamp KE (2012) Innovative problem solving by wild spotted hyenas. Proc R Soc Lond B 279:4087-4095.
- Benson-Amram S, Weldele ML, Holekamp KE (2013) A comparison of innovative problem-solving abilities between wild and captive spotted hyenas (*Crocuta crocuta*). Anim Behav 85:349-356.
- Biondi LM, Bó MS, Vassallo AI (2008) Experimental assessment of problem solving by *Milvago chimango* (Aves: Falconiformes). J Ethol 26:113-118.
- Bird CD, Emery NJ (2009) Rooks use stones to raise the water level to reach a floating worm. Curr Biol 19:1410-1414.
- Boogert NJ, Fawcett TW, Lefebvre L (2010) Mate choice for cognitive traits: a review of the evidence in nonhuman vertebrates. Behav Ecol 22:447-459.
- Borrego N, Dowling B (2016) Lions (*Panthera leo*) solve, learn, and remember a novel resource acquisition problem. Anim Cogn 19:1019-1025. DOI: 10.1007/s10071-016-1009-y
- Briefer EF, Haque S, Baciadonna L, McElligott AG (2014) Goats excel at learning and remembering a highly novel cognitive task. Front Zool 11:20.
<http://www.frontiersinzoology.com/content/11/1/20>
- Cauchard L, Boogert NJ, Lefebvre L, Dubois F, Doligez B (2013) Problem-solving performance is correlated with reproductive success in a wild bird population. Anim Behav 85:19-26.
- Chow PKY, Leaver LA, Wang M, Lea SEG (2015) Serial reversal learning in grey squirrels: learning efficiency as a function of learning and change of tactics. J Exp Psychol Anim Learn Cogn 41:343-353.

Chow PKY, Lea SEG, Leaver LA (2016) How practice makes perfect: the role of learning, flexibility, and persistence in problem solving efficiency. *Anim Behav* 112:273-283.

Chow PKY, Leaver LA, Wang M, Lea SEG (2017) Touch screen assays of behavioural flexibility and error characteristics in Eastern grey squirrels (*Sciurus carolinensis*). *Anim Cogn* 20:459-471.

Cole EF, Morand-Ferron J, Hinks AE, Quinn JL (2012) Cognitive ability influences reproductive life history variation in the wild. *Curr Biol* 22:1-5.

Cuvo AJ (2003) On stimulus generalization and stimulus classes. *J Behav Educ* 12:77. doi:10.1023/A:1022374406394

Diquelou MC, Griffin AS, Sol D (2016) The role of motor diversity in foraging innovations: a cross-species comparison in urban birds. *Behav Ecol* 27:584-591.

Griffin AS, Diquelou M, Perea M (2014) Innovative problem solving in birds: a key role of motor diversity. *Anim Behav* 92:221-227.

Griffin AS, Diquelou MC (2015) Innovative problem solving in birds: a cross-species comparison of two highly successful passerines. *Anim Behav* 100:84-94.

Guez D, Griffin AS (2016) Unraveling the key to innovative problem solving : a test of learning versus persistence. *Behav Ecol* 00:1-12. doi:10.1093/beheco/arw055

Hardin JW, Hilbe JM (2003) *Generalized estimating equations*. Chapman and Hall/CRC.

Hopewell LJ, Leaver LA, Lea SEG, Wills AJ (2010) Grey squirrels (*Sciurus carolinensis*) show a feature-negative effect specific to social learning. *Anim Cogn* 13:219-227.

Isden J, Panayi C, Dingle C, Madden J (2013) Performance in cognitive and problem-solving tasks in male spotted bowerbirds does not correlate with mating success. *Anim Behav* 86:829-838.

Jacobs LF, Liman ER (1991) Grey squirrels remember the locations of buried nuts. *Anim Behav* 41:103-110.

Keagy J, Savard J-F, Borgia G (2009) Male satin bowerbird problem solving ability predicts mating success. *Anim Behav* 78:809-817.

Macdonald IMV (1997) Field experiments on duration and precision of grey and red squirrel spatial memory. *Anim Behav* 54:879–891. doi:10.1006/anbe.1996.0528

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

Malmberg KJ, Xu J (2007) On the flexibility and the fallibility of associative memory. *Mem Cognit* 35:545-556.

Manrod JD, Hartdegen R, Burghardt GM (2008) Rapid solving a problem apparatus by juvenile black-throated monitor lizards (*Varanus albigularis albigularis*). *Anim Cogn* 11:267-273.

Millot S, Nilsson J, Fosseidengen JE, Bégout M-L, Fernö A, Braithwaite VA (2014) Innovative behaviour in fish: Atlantic cod can learn to use an external tag to manipulate a self-feeder. *Anim Cogn* 17:779-785.

Overington SE, Cauchard L, Côté K-A, Lefebvre L (2011) Innovative foraging behaviour in birds: What characterizes an innovator? *Behav Processes* 87:274-285.

Page EB (1963) Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Amer Statist Assoc* 58:216-230.

Papp S, Vincze E, Preiszner B, Liker A, Bókony V (2015) A comparison of problem-solving success between urban and rural house sparrows. *Behav Ecol Sociobiol* 69:471-480.

Preiszner B, Papp S, Pipoly I, Seress G, Vincze E, Liker A, Bókony V (2017) Problem-solving performance and reproductive success of great tits in urban and forest habitats. *Anim Cogn* 20:53-63.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Reader SM, Laland KN (2003) *Animal innovation*. Oxford University Press.

Reichmuth Kastak C, Schusterman RJ (2002) Long-term memory for concepts in a California sea lion (*Zalophus californianus*). *Anim Cogn* 5:225-232.

- Sokal RR, Rohlf FJ (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*. 3rd edn. New York: W.H. Freeman.
- St Clair JJH, Rutz C (2013) New Caledonian crows attend to multiple functional properties of complex tools. *Phil Trans R Soc B* 368:20120415.
- Taylor AH, Elliffe D, Hunt GR, Gray RD (2010) Complex cognition and behavioural innovation in New Caledonian crows. *Proc R Soc Lond B* 277:2637-2643.
- Thompson DC, Thompson PS (1980) Food habits and caching behavior of urban grey squirrels. *Can J Zool* 58:701–710.
- Thornton A, Samson J (2012) Innovative problem solving in wild meerkats. *Anim Behav* 83:1459-1468.
- van Horik JO, Madden JR (2016) A problem with problem solving: motivational traits, but not cognition, predict success on novel operant foraging tasks. *Anim Behav* 114:189-198.
- Wang M, Long Q (2011) Modified robust variance estimator for generalized estimating equations with improved small-sample performance. *Stat Med* 30:1278-1291.
- Wang M (2015) crank: Modified Variance Estimators for Generalized Estimating Equations. R package version 1.0.
- Werdenich D, Huber L (2006) A case of quick problem solving in birds: string pulling in keas, *Nestor notabilis*. *Anim Behav* 71:855-863.
- Yan VX, Yu Y, Garcia MA, Bjork RA (2014) Why does guessing incorrectly enhance, rather than impair, retention? *Mem Cognit* 42:1373-1383.

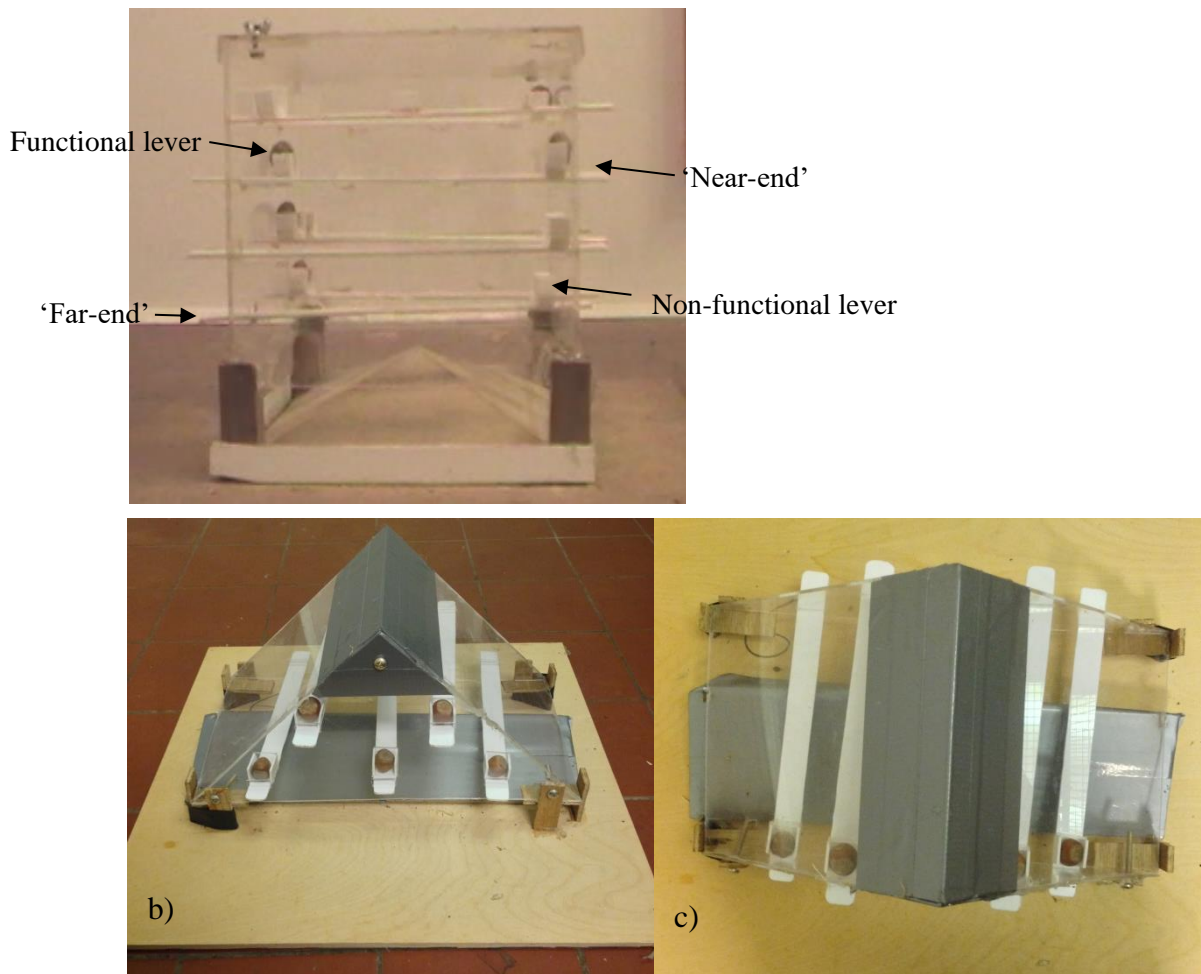


Figure 1. a) Recall task used the puzzle box that we had presented to the same squirrels 22 months before, as reported in Chow et al. 2016. Puzzle box for generalisation task b) front view and c) top view of the problem apparatus for the generalisation task. This problem is designed to keep the same solution as in the original task, but appear as a novel task for squirrels. A functional lever contains a nut whereas a non-functional lever is empty. A lever has two ends, the 'near-end' refers to the end close to the nut container whereas 'far-end' refers to the end far from the container.

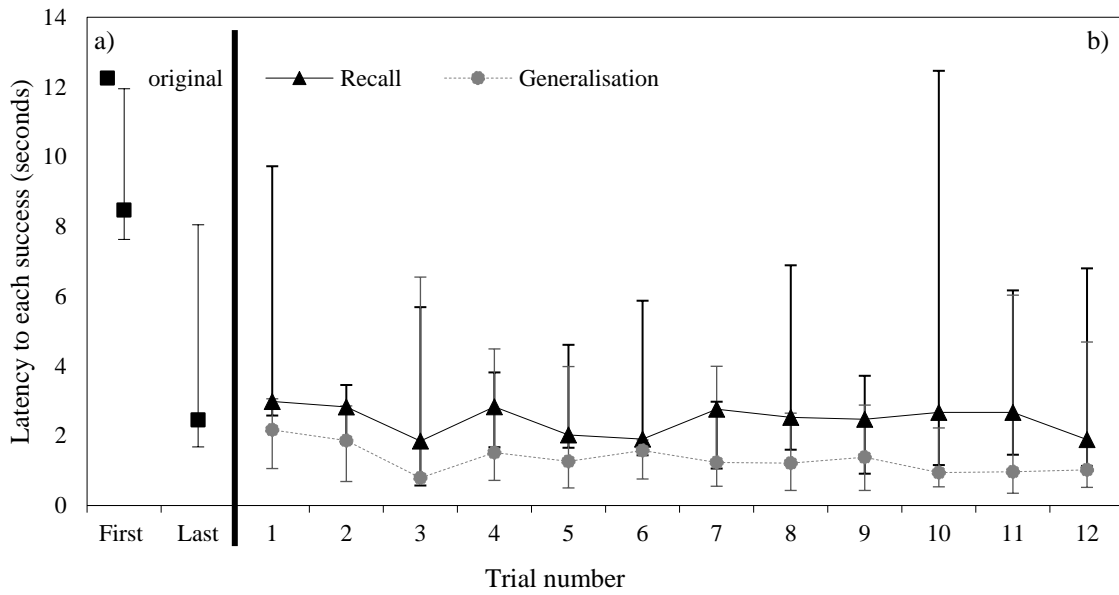


Figure 2. Median, maximum and minimum of mean latency to each success in a) the first and last trial of puzzle box task. b) Across 12 trials in the recall tasks, and c) across 12 trials in the generalisation task. N=5.

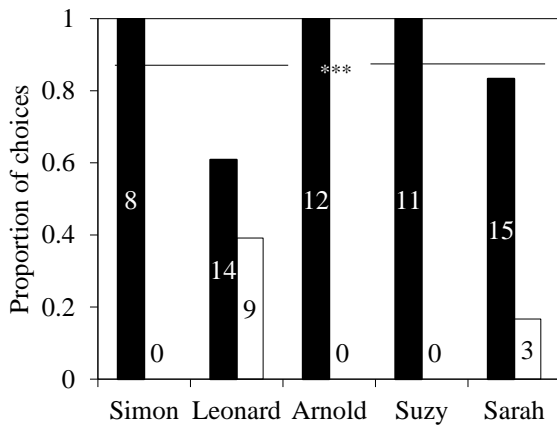


Figure 3. Proportion of attempts between functional levers (■) and non-functional levers (□) in the first trial of recall task. Number inside each bar indicates the actual number of attempts. N=5, ***P<0.001

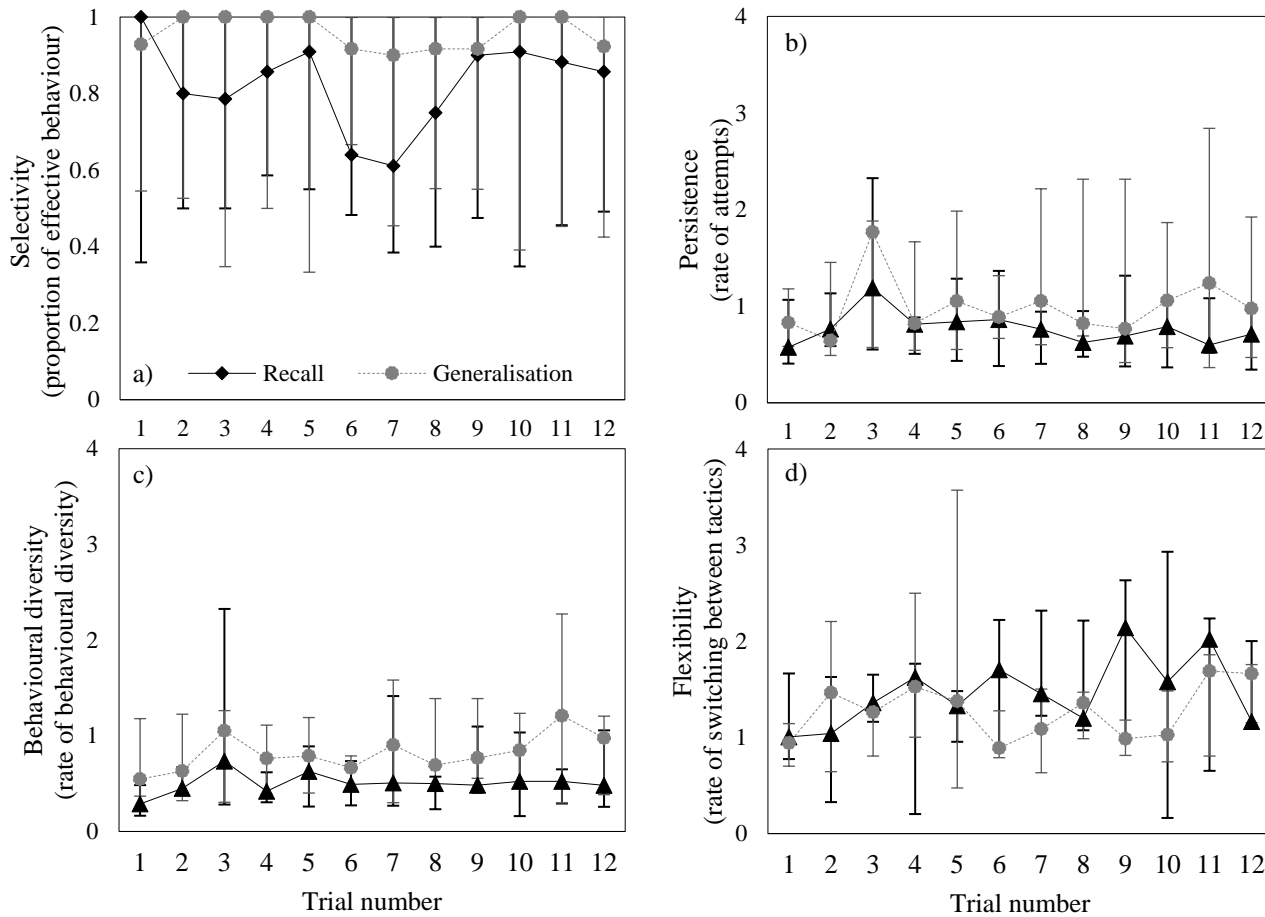


Figure 4. Medians (\pm minimum and maximum) of each behavioural trait across trials in the recall task and the generalisation task. a) Behavioural selectivity, the proportion of effective behaviours, across trials. b) Persistence, the rate of attempts, across trials. c) Behavioural diversity, the rate of behavioural diversity, across trials. d) Flexibility, the rate of switching between tactics, across trials. N=5.

Table 1. Summary for GEE models including estimates, chi-square values (χ^2), P values, effect size of each path (Path β), and total effect size (Total β) of each covariate.

DV	Covariates	Estimates	Recall test				Generalisation task				
			χ^2	P	Path β	Total β	Estimates	χ^2	P	Path β	Total β
Model 1											
Latency to each success	Experience	<0.01	<0.01	0.969	0.01	-0.10	-0.02	0.80	0.371	-0.06	-0.11
	Selectivity	-7.69	25.39	<0.001	-0.74	-0.74	-5.37	12.05	<0.001	-0.88	-0.88
	Persistence	-0.65	0.86	0.353	-0.17	-0.47	---	---	---	---	---
	Flexibility	-1.04	6.17	0.013	-0.38	-0.38	-0.06	0.13	0.713	-0.02	-0.03
Model 2											
Selectivity	Experience	<0.01	<0.01	0.989	<0.01		<0.01	0.37	0.544	0.06	
	Persistence	0.10	22.65	<0.001	0.26		---	---	---	---	
	Flexibility	---	---	---	---		0.01	0.06	0.812	0.01	
Model 3											
Flexibility	Experience	0.05	4.85	0.028	0.22		0.01	0.06	0.808	0.05	
	Persistence	0.41	6.03	0.014	0.29		---	---	---	---	
Model 4											
Persistence	Experience	<0.01	0.05	0.826	0.03						

Note that all measures are taken trial-by-trial (12 trials in total). Experience was recorded as trial number (total 12 trials); Persistence was measured as the rate of attempts; behavioural selectivity was recorded as the proportion of effective behaviours (i.e. either push the ‘near-end’ or pull the ‘far-end’); flexibility was measured as the rate of switching to another type of tactic after a failed attempt.