

FRONT MATTER

Title

- Full title: Human Ageing is Associated with More Rigid Concept Spaces
- Short title: Concept Change in Older Adults

Authors

Sean Devine^{1, 2*}, Cassandra Neumann¹, David Levari³, Robert C. Wilson^{4, 5, 6}, Ben Eppinger^{1, 7, 8}

* seandamiandevine@gmail.com

Affiliations

¹ Department of Psychology, Concordia University, Montreal, Canada

² Department of Psychology, McGill University, Montreal, Canada

³ Harvard Business School, Boston, MA, USA

⁴ Department of Psychology, University of Arizona, Tucson, USA

⁵ Cognitive Science Program, University of Arizona, Tucson, USA

⁶ Evelyn F. McKnight Brain Research Foundation, University of Arizona, Tucson, USA

⁷ PERFORM Center, Concordia University, Montreal, Canada

⁸ Lifespan Developmental Neuroscience Department, Technische Universität Dresden, Germany

Abstract

Prevalence-induced concept change describes a cognitive mechanism by which someone's definition of a concept shifts as the prevalence of instances of that concept changes. The phenomenon has real-world implications because this sensitivity to environmental characteristics may lead to substantial biases in judgements. While prevalence-induced concept change has been established in young adults, it is unclear how it changes as a function of human ageing. In this cross-sectional study, we explore how prevalence-induced concept change affects older adults' lower-level, perceptual, and higher-order, ethical, judgements. We find that older adults are less sensitive to prevalence-induced concept change than younger adults across domains. Using a combination of computational and experimental approaches, we demonstrate that these changes in judgements are sensitive to the pace with which the stimuli occur in the environment and are affected by the effort that subjects invest in order to make accurate decisions. Based on findings from three experiments we argue that older adults' concept spaces are more rigid than those of younger adults. However, what appear as an age-related cognitive "deficit" may turn out to be beneficial because it makes older adults less susceptible to biases in judgments.

40 “*The more things change, the more they stay the same.*”

41 — Jean-Baptiste Alphonse Karr and, later, Jon Bon Jovi

42 By 2068, almost 30% of the North-American population will be 65 years or older
43
44 (Statistics Canada, 2019; U.S. Census Bureau, 2018). As adults age, their judgments and
45 decisions will affect society more than ever before and will largely influence our collective future.
46 As such, it is critical to understand how the cognitive and motivational processes underlying
47 judgement and decision-making change with age and how these changes may affect real-world
48 decisions.

49 In this study, we explore how changes in one’s environment affect concept formation and
50 judgements in younger and older adults. Specifically, we consider how judgements about
51 perceptual and ethical concepts are affected by the prevalence of instances of them in the
52 environment. This phenomenon has been referred to as **prevalence-induced concept change**
53 (Levari et al., 2018). Prevalence-induced concept change describes the empirical observation that
54 as the numbers of instances of a given concept change in the environment, so do the boundaries
55 for that concept, such that they come to include instances that they would otherwise exclude. For
56 example, one task that measures prevalence-induced concept change requires participants to
57 serially judge whether individual dots that vary on a spectrum between blue and purple are in fact
58 blue or purple. When the relative frequency of objectively coloured dots in the environment is
59 consistent across the task (50% blue 50% purple dots), peoples’ judgements are relatively stable:
60 blue dots are judged as blue and purple dots as purple. However, if the number of blue dots
61 changes, dots initially judged as purple are later (after the prevalence changes) categorised as
62 blue. Interestingly these changes do not only occur on the perceptual level, but also arise in
63 higher-order social and ethical judgements. (Levari et al., 2018).

64 What these findings suggest is that, from judgement to judgement, people adjust their
65 concepts to environmental characteristics. In line with this view, recent work suggests that the
66 cognitive mechanisms underlying prevalence-induced concept change can be captured using

67 computational modeling (Wilson, 2018). In this model concept boundaries arise from competition
68 between the effect of the past stimulus and the effect of the past response on the judgement; the
69 former increasing prevalence-induced concept and the latter reducing it.

70 Why would we expect ageing-related changes in this regard then? Many cognitive
71 changes that come with healthy ageing offer good reason to suspect that older adults might be
72 differentially affected by prevalence-induced concept change compared to younger adults. Several
73 empirical findings suggest that older adults differ from young adults in terms of motivation,
74 postponement of gratification, and to what degree they value desired outcomes (Mather & Harley,
75 2016; Samanez-Larkin & Knutson, 2015; Eppinger et al., 2011). These differences in decision-
76 making processes can in part be explained by age-related differences in cognitive ability, such as
77 changes in executive function (Mayr, Spieler, & Kliegl, 2001), memory (Nyberg et al., 2012), and
78 processing speed (Salthouse, 1992; 1996). Specifically, two lines of research paint opposing
79 pictures of how older adults might make different concept judgements than younger adults when
80 the prevalence of instances of a concept in the environment changes.

81 On the one hand, previous work suggests that older adults have more difficulty learning
82 from uncertain outcomes compared to younger adults (Nassar et al., 2016). This difficulty
83 manifests as perseverative behaviour, whereby older adults have a tendency to repeat previous
84 responses despite changes in the environment (Bruckner et al., 2020; Eppinger, Walter, Heekeren,
85 & Li, 2013). With respect to the current study, these findings may suggest a decreased sensitivity
86 to prevalence-induced concept change in older adults, because the repetition of past choices
87 makes it less likely that a rarer category will be chosen after a shift in prevalence. The same can
88 also be expressed computationally, where perseverative behaviour is reflected by a higher
89 influence of past choice, which has previously shown to reduce prevalence-induced concept
90 change (Wilson, 2018).

91 On the other hand, results from several recent studies suggest that older adults have
92 difficulty converging on an accurate representation of the current state, particularly if these states
93 are latent (not directly observable) and need to be inferred from experience (Hämmerer et al.,
94 2019; Hämmerer, Müller, & Li, 2014; Eppinger, Heekeren, & Li, 2015). To help compensate for
95 this difficulty in distinguishing task states, older adults may outsource control to the environment
96 rather than relying on (sometimes inaccurate) internal representations (Mayr, Spieler, &
97 Hutcheon, 2015; Spieler, Mayr, & LaGrone, 2006; Lindenberger & Mayr, 2015). In the case of
98 prevalence-induced concept change, this outsourcing of control is likely to lead to increased
99 concept change. From a computational perspective, this tendency to outsource also represents an
100 increased sensitivity to the effect of previous stimulus, which has similarly been shown to
101 increase prevalence-induced concept change (Wilson, 2018).

102 Taken together, the results of the aforementioned studies point to two opposing
103 hypotheses: Hypothesis 1 suggests that older adults are **less** sensitive to prevalence-induced
104 concept change than younger adults, whereas Hypothesis 2 suggests that older adults are **more**
105 sensitive to prevalence-induced concept change than younger adults. To visualize our predictions
106 we used the computational model by Wilson (2018) to simulate data for each of the hypotheses
107 (see Figure 1).

108 To tease these hypotheses apart and gain a better understanding of the cognitive
109 mechanisms underlying age-related changes in prevalence-induced concept change we ran three
110 experiments. In the first experiment, we use an age-comparative study design and a computational
111 model to investigate age differences in prevalence-induced concept change in lower-level,
112 perceptual, and higher-level, moral, judgements. We show that, across domains, older adults are
113 less susceptible to prevalence-induced concept change than younger adults. In the second step, we
114 explore two potential explanations for these age-related changes in prevalence-induced concept
115 change by manipulating concept formation in younger adults. In Experiment 2, we vary the inter-

116 trial intervals (ITI) between occurrences of stimuli and show that a greater spacing of stimuli
117 reduces prevalence-induced concept change in younger adults (akin to the findings observed in
118 older adults). In Experiment 3, we provide incentives for consistent judgments and show that this
119 manipulation also shifts younger adults' judgments towards the behavior observed in older adults.
120 Across experiments, our findings suggest that age-related changes in prevalence-induced concept
121 change may reflect a combination of changes in the timing between stimuli and higher-level
122 changes in the motivation for accurate judgements.

123

124

Experiment 1

Method

Participants

127 We recruited 160 participants from the community and the university participation pool,
128 80 of which were older adults ($M_{\text{age}} = 70.10$; $s_{\text{age}} = 5.55$) and 80 of which were younger adults
129 ($M_{\text{age}} = 21.85$; $s_{\text{age}} = 2.27$). All participants were English-speaking, free of neurological or
130 psychiatric disorders, and free of any cognitive, motor, visual, or other condition(s) that would
131 impede their performance, including but not limited to a history of head trauma with loss of
132 consciousness, organic brain disorders, seizures, or neurosurgical intervention, to sensory deficits
133 (i.e. deafness, blindness, colour blindness, intellectual disability), or self-reported cognitive
134 impairment, and to a recent history of substance abuse. In each age group, 40 participants were
135 randomly assigned to either the decreasing prevalence condition or the stable prevalence
136 condition, in a counterbalanced order. In the former, they experienced a decreasing prevalence of
137 instances of the concept in both tasks detailed below. In the latter, the prevalence remained the
138 same throughout the entire experiment. All participants were compensated \$20 CAN or 2
139 participation pool credits for participating in the study. The study protocol was approved by the
140 Concordia Human Research Ethics Committee (certification number 30011191). In this

141 experiment and the two others detailed below, sample sizes were based on Levari et al. (2018)
142 data and the stopping rule for data collection was to collect until the end of the semester and
143 group sizes were even.

144 *Materials*

145 **The Dots Task.** In the Dots Task, participants had to judge the colour of an individual dot
146 presented on the screen. The task began with a series of instruction screens explaining the task to
147 the participant. These instructions were followed by a practice block consisting of 10 trials, in
148 which participants could familiarize themselves with the task. These trials were identical to trials
149 in the real task and consisted of 50% purple dots and 50% blue dots. Data from practice trials
150 were not analysed.

151 After the practice block, participants performed a total of 800 trials, divided into 16 blocks
152 of 50 trials each. In the decreasing prevalence condition, the number of blue dots in the
153 environment decreased as the number of blocks increased in a predetermined fashion.
154 Specifically, the proportion of blue relative to purple dots, was as follows for each of the 16
155 blocks: .50, .50, .50, .50, .40, .28, .14, .06, .06, .06, .06, .06, .06, .06, .06. In the stable
156 prevalence condition, the proportion of blue dots in the environment remained the same (.50)
157 across the experiment. In both cases, blue dots were defined as any dot who's RGB value was
158 between [0, 0, 254] and [49, 0, 205]. Purple dots were defined as any dot who's RGB value was
159 between [50, 0, 204] and [99, 0, 155]. Dot colours were randomly chosen for each trial based on
160 the number of trials per block (50) and the frequency with which blue and purple dots should
161 appear in a given block.

162 In each trial, participants judged the color of the dots as being either blue or purple by
163 pressing the 'A' or 'L' key on the keyboard. All stimuli were presented against a dark grey
164 background. Each trial started with a dot presented on the screen for 500 ms, followed by a
165 question mark that appeared on the screen until participants made a choice, and finally a blank

166 screen appeared for 500 ms as an ITI. Thus, all timing was fixed across participants, except that
167 which would arise from differences in response times. After each block text appeared that
168 indicated that the block was finished, which block the participant was now at, and offering them a
169 short break should they choose to take one.

170 **The Ethics Task.** In the Ethics Task, participants had to take on the role of a member of
171 an Ethics Review Board and judge whether fictitious research proposals were ethical or not
172 (phrased as whether they would allow these research studies to be conducted or not). All research
173 proposals were norm tested by Levari et al. (2018, see Supporting Online Material) to produce
174 scores depicting how ethical people found the 273 proposals. These scores were used to bin
175 proposals as unethical (80 proposals), ethical (113 proposals), or ambiguous (80 proposals). These
176 bins were used to calculate the proportion of proposals that appeared in each block (including the
177 practice trial). Just as in the Dots Task, participants were first presented with instruction screens
178 explaining the task to them. Following the instructions, participants completed a practice trial in
179 which they judged a research proposal using the keyboard keys. In this task, they pressed ‘A’
180 when they would not allow a study to be conducted and ‘L’ when they would.

181 Following the practice trial, participants began the test trials. All proposals in the
182 experiment were presented in black text against a dark grey background. The task consisted of
183 240 trials broken into 10 blocks. In the decreasing prevalence condition, the proportion of
184 unethical, ethical, and ambiguous proposals varied across blocks. Specifically, for the 10 blocks
185 of the study, the proportion of unethical proposals relative to ethical and ambiguous proposals
186 were as follows: .33, .33, .33, .33, .25, .17, .08, .04, .04, .04 (rounded to the nearest 2nd decimal).
187 In the stable prevalence condition, the proportion between the three types of proposals was the
188 same throughout the task: .33.

189 Each trial, participants read a proposal and pressed ‘A’ or ‘L’ on the keyboard indicating
190 whether they thought that the research should be allowed to be conducted on people or not. There

191 was no time limit on this choice. Following the choice, a fixation cross appeared on the screen for
192 500 ms, followed by the next proposal. Between each block, text appeared that indicated that the
193 block was finished, which block the participant was now at, and offering them a short break
194 should they choose to take one.

195 Both the Dots and Ethics Tasks described above were taken from Levari et al. (2018).
196 Both tasks were programmed in Python using the PsychoPy libraries.

197 ***Procedure***

198 Participants were recruited from the community through online or paper advertisements or
199 from Concordia’s participation pool. Participants were contacted by telephone or email and were
200 asked basic demographic information to determine initial eligibility. If eligible at this stage, they
201 were invited for a single two-hour session in the lab.

202 Once at the lab, participants were asked to fill out a consent form and complete the
203 Richmond HRR pseudoisochromatic test for colour vision (Cole, Lian, & Lakkis, 2016; see
204 Supplement for more details). Participants were then asked to complete the Dots Task and Ethics
205 Task, back-to-back. The order of these tasks was counterbalanced across participants. They were
206 told that they would be free to take short breaks during the tasks (between blocks) and a longer
207 break between the tasks, should they choose to. After completing both tasks, participants were
208 debriefed and paid \$20 for participating or were given their participation credits.

209 ***Computational Model***

210 We used a sequential decision-making model based upon logistic regression to explore
211 prevalence-induced concept change on a trial-by-trial basis (Wilson, 2018):

$$p_t = 1 - \frac{1}{1 + \exp(\beta_0 + \beta_f f_t + \beta_F F_t + \beta_c C_t)} \tag{Eq. 1}$$

212
213 In this model, p_t is the probability of classifying the current stimulus as blue or unethical, β_0
214 captures the overall bias for classifying the stimulus as blue or unethical, β_f captures the effect of

215 the current stimulus, β_F captures the effect of the past stimulus, and β_C captures the effect of past
216 response. F_t and C_t represent the exponentially weighted sum of past stimuli and past response
217 respectively. These parameters are controlled by two other parameters, λ_F and λ_C , which dictate
218 the rate of decay of the exponential weighting with larger values corresponding to slower decay.
219 They follow the following update rules:

$$F_t = \lambda_F F_t + f_t \quad (\text{Eq. 2})$$

$$C_t = \lambda_C C_t + c_t \quad (\text{Eq. 3})$$

220
221 This leaves six free parameters to be estimated ($\beta_0, \beta_f, \beta_F, \beta_C, \lambda_F$ and λ_C). These parameters
222 can be estimated using a standard maximum likelihood approach, where the parameters that
223 produce the maximum sum of log scores is taken (in practice, the minimum negative logged sum;
224 see Daw, 2011). Of these parameters, β_F and β_C are of the most theoretical interest. As β_F
225 decreases, the effect of previous stimuli becomes stronger and biases choice behaviour to the
226 opposite of the previous stimulus (e.g., if the previous dot was purple, the subsequent choice
227 would be blue). On the other hand, as β_C increases, responses become biased to match the
228 previous response (e.g., if one responded that the previous dot was blue, they would do so again
229 on their subsequent choice). Thus, according to Wilson (2018), the degree to which one is
230 sensitive to prevalence-induced concept change arises as a function of the relative strengths of
231 these opponent parameter weights.

232 ***Simulation of Predicted Results***

233 Based on our hypotheses, we simulated data using the computational model detailed
234 above. Specifically, we simulated data for three scenarios. First, H0 predicted that older and
235 younger adults would not differ in their sensitivity to prevalence-induced concept change. To
236 simulate this scenario, we imputed similar parameters as found in healthy young adults into the
237 computational model (parameters from Wilson, 2018). Second, H1 predicted that older adults
238 would be less sensitive to prevalence-induced concept change than young adults.

239 Computationally, we simulated this scenario by inputting greater β_C values into the model (greater
240 effect of past response). Finally, H2 predicted that older adults would be more sensitive to
241 prevalence-induced concept change than young adults, which we simulated by inputting lower β_F
242 values into the model. The results of these simulations are summarised in Figure 1.

243 Qualitatively, these simulations demonstrate the pattern of results we expect in accordance
244 with each hypothesis, as well as a rough estimate of the parameter values we think would underlie
245 participants' behaviour in each scenario.

246 *Statistical Analysis*

247 All data were analysed in R (version 3.6.1). For the Ethics task normed scores were
248 reversed, to make the plots in the same direction as the Dots Task, such that lower normed scores
249 now represented more ethical scenarios. The main analysis consisted of six general binomial
250 mixed-effects models that were implemented and fit using the *lme4* package (Bates, Maechler,
251 Bolker, & Walker, 2015). The main models in each task predicted response using age group
252 (young adult or older adult), condition (stable prevalence or decreasing prevalence), trial number,
253 and stimulus strength (colour in the Dots Task and normed ethicality scores in the Ethics Task) as
254 fixed effects, a random slope of trial, and a random intercept for each participant. All main effects
255 and interactions were explored. Two follow-up models were conducted in both tasks using the
256 same predictors, split by age group. In all statistical models, trial and stimulus strength were put
257 on a scale between zero and one. Model weights were estimated using the *nlminb* compared
258 across age groups and conditions using a between-groups 2x2 ANOVA.

259 Finally, we analysed response times using two between-groups 2x2 (age group \times
260 condition) ANOVA, one for each task. In line with best practices regarding HARKing
261 (Hollenback & Wright, 2017), we wish to disclose that these analyses were exploratory and not
262 based on original hypotheses. To supplement these exploratory analyses, we conduct two follow-
263 up studies, presented in more detail below.

264 **Results**

265 ***Choice Data***

266 Statistically-speaking, prevalence-induced concept change is reflected in a three-way
267 interaction between condition, trial, and stimulus strength, predicting responses. The effect size of
268 this interaction reflects the degree to which a participant's choice to categorize a given exemplar
269 (dot or research proposal) as one concept or another is influenced by a combined effect of three
270 factors: (a) the prevalence of instances in the environment (i.e., the effect of condition), (b) the
271 amount of time that has past (i.e., the effect of trial), and (c) the strength of the stimulus (i.e.,
272 blueness or ethicality). Thus, if younger and older adults differ in their sensitivity to prevalence-
273 induced concept change, we would expect to see a four-way interaction between these three terms
274 above and age group, as well as different effect sizes for the three-way effect within each of the
275 age groups.

276 Indeed, this is exactly what we observe. Results from mixed-effects regressions are
277 represented in Figure 2. In both tasks, there was a four-way interaction between age group,
278 condition, trial, and stimulus strength (In the Dots Task: $\beta = 4.32$, $SE = 0.70$, $p < .0001$, 95% CI =
279 [3.22, 5.42]; In the Ethics Task: $\beta = 1.13$, $SE = 0.21$, $p < .0001$, 95% CI = [0.72, 1.54]).
280 Furthermore, separate regression analyses for the two age groups revealed that the effect of
281 prevalence-induced concept change was stronger in younger adults ($\beta = 22.54$, $SE = 0.42$, $p <$
282 $.0001$, 95% CI = [21.73, 23.36]) than older adults ($\beta = 18.62$, $SE = 0.28$, $p < .0001$, 95% CI =
283 [18.06, 19.18]). In the Ethics Task the three-way interaction was significant in younger adults
284 ($\beta_{\text{Young Adults}} = 1.34$, $SE = 0.17$, $p < .0001$, 95% CI = [1.00, 1.69]), but not in older adults (β_{Older}
285 $\text{Adults} = 0.13$, $SE = 0.12$, $p = .2754$, 95% CI = [-0.11, 0.37]). Overall, these results suggest that
286 older adults are less sensitive to prevalence-induced concept change than younger adults.

288 ***Response Times***

289 Response time data across age groups are presented in Figure 3. Two 2x2 ANOVA (age
290 group \times condition) were conducted on each subjects' mean response time data. These analyses
291 revealed a significant main effect of age group on response time in both tasks (Dots Task: $F(1,$
292 $116) = 87.33, p < .0001, 95\% \text{ CI} = [0.23, 0.40], \text{ difference}_{\text{Older} - \text{Young}} = 0.28 \text{ seconds}$; Ethics Task:
293 $F(1, 116) = 28.19, p < .0001, 95\% \text{ CI} = [0.86, 3.30], \text{ difference}_{\text{Older} - \text{Young}} = 2.32 \text{ seconds}$), but no
294 statistically significant main effect of condition (Dots Task: $F(1, 116) = 0.12, p = .7529, 95\% \text{ CI}$
295 $= [-0.06, 0.11]$; Ethics Task: $F(1, 116) = 0.37, p = .5432, 95\% \text{ CI} = [-1.72, 0.71]$) or interaction
296 between age group or condition (Dots Task: $F(1, 116) = 1.20, p = .2768, 95\% \text{ CI} = [-0.19, 0.05]$;
297 Ethics Task: $F(1, 116) = 0.30, p = .5837, 95\% \text{ CI} = [-1.24, 2.21]$). These findings suggest that
298 older adults made slower responses in both tasks, but that neither group differed with regards to
299 response speed between conditions.

300 ***Computational Modeling Results***

301 We conducted a 2x2 ANOVA (age group \times condition) on each free parameter ($\beta_0, \beta_f, \beta_F,$
302 β_C, λ_F and λ_C). We found no statistically significant effect of age group on any of the estimated
303 parameters. However, both decay parameters, λ_F and λ_C , in the Dots Task showed a significant
304 interaction between age group and condition ($\lambda_F: F(1, 116) = 9.94, p = .0019, \hat{\eta}^2_G = 0.06, 95\% \text{ CI}$
305 $= [-0.42, -0.10]$; $\lambda_C: F(1, 116) = 11.92, p = .0007, \hat{\eta}^2_G = 0.07, 95\% \text{ CI} = [-0.55, -0.15]$). In both
306 cases, this interaction suggests that older adults have a slower decay parameter in the stable
307 condition, but quicker decay in the decreasing condition. No such interactions were found in the
308 Ethics Task.

309 Overall, these results did not substantiate our original prediction that older adults'
310 decreased sensitivity to prevalence-induced concept would be reflected by an increased influence
311 of the β_C parameter (greater influence of past choice; increased perseveration). We provide an
312 alternative explanation for the behavioural results we observe below.

313

314 **Discussion**

315 In this experiment, we demonstrate that older adults are less sensitive to prevalence-
316 induced concept change than young adults. The observed age differences are independent of the
317 complexity of the judgements (colour of dots versus ethical of research proposals) and may
318 indicate that concept spaces in older adults are less susceptible to environmental changes.

319 But, where do these age differences come from? We hypothesized that if age differences
320 existed, they could be explained by differences between young and older adults in estimated
321 parameter values using a computational model (Figure 1; Wilson, 2018). This hypothesis was not
322 substantiated, however. We did not find meaningful age differences in any of the estimated
323 parameter weights. While the model replicated Wilson’s (2018) overall pattern of results, neither
324 parameter accounted for age differences in concept change.

325 Apart from the reduction of prevalence-induced concept change in older adults, one major
326 difference in the behavior of younger and older adults is in their response times. As shown in
327 Figure 3, older adults responded much slower than younger adults. This age-related slowing is an
328 expected and well-documented feature of healthy cognitive ageing (Salthouse, 1992). However,
329 the consequence of the slower response times is that the interval between stimuli increases, which
330 may lead to a reduced impact of the previous stimulus on the current judgement. As noted by
331 Wilson (2018), increases in this ITI may lead to a reduction in prevalence-induced concept
332 change, given that the weakest effects observed in Levari’s and colleagues’ (2018) results were in
333 the task with the longest space between stimuli (the Ethics Task). To account for the potential
334 effects of prolonged responses on prevalence-induced concept change, we modified the original
335 computational model such that it accounts for response time differences (see Supplement).
336 Simulated data from this model (Figure 4) qualitatively resemble the empirical findings in Figure
337 2 and suggest a reduction of prevalence-induced concept change in older adults.

363

364 **Method**

365 *Participants*

366 We recruited 36 young adults ($M_{\text{age}} = 20.89$; $s_{\text{age}} = 2.09$) from the community and the
367 university participation pool. All participants met the inclusion criteria mentioned in the Method
368 section for Experiment 1. The participants were randomly assigned to either the decreasing
369 prevalence condition or the stable prevalence condition, in a counterbalanced order. All
370 participants were compensated \$20 CAN or 2 participation pool credits for participating in the
371 study.

372 *Materials and Procedure*

373 **ITI-Modified Dots Task.** This version of the Dots Task is the same as the original
374 version used in Experiment 1 in all respects except for one: the ITI between stimuli was increased
375 from 500ms to 2000ms. This time of 2000 ms was chosen as a way to test whether simple timing
376 differences between stimuli affected sensitivity to prevalence-induced concept change (a)
377 between the age groups (i.e., to make the case that longer responses on the part of older adults
378 may affect sensitivity to the phenomenon) and (b) between the tasks (i.e., if reduced prevalence-
379 induced concept change in the Ethics Task might be driven by longer response times across both
380 age groups). The procedure was the same as in Experiment 1.

381 *Analysis*

382 As before, the main analysis consisted of two general binomial mixed-effects models. The
383 first model compared young adults in both experiments, predicting response using study
384 (Experiment 1 or Experiment 2), condition (stable prevalence or decreasing prevalence), trial
385 number, and colour strength as fixed effects. All main effects and interactions were explored. The
386 follow-up model used the same predictors, but in Experiment 2 only. In all models, trial and
387 stimulus strength were converted to scores between zero and one.

388 **Results**

389 ***Choice Data***

390 Choice data for Experiment 2 are presented in Figure 5A. We found a significant four-way
391 interaction between study, condition, trial, and colour strength ($\beta = -27.00$, $SE = 1.13$, $p < .0001$,
392 95% $CI = [-29.24, -24.76]$), such that the effect of prevalence-induced concept change was
393 significantly smaller for young adults in Experiment 2 compared to those in Experiment 1. A
394 follow-up analysis revealed a still significant, but dramatically smaller, interaction between
395 condition, trial, and colour on response for participants in Experiment 2 ($\beta = -4.41$, $SE = 1.05$, p
396 $< .0001$, 95% $CI = [-6.93, -2.39]$).

397 **Discussion**

398 In this experiment, we explored whether the degree of prevalence-induced concept change
399 depends on the spacing between stimuli. Our results support this hypothesis and demonstrate a
400 significant reduction in prevalence-induced concept in younger adults from Experiment 1 to
401 Experiment 2. It is worth noting though that, even in the long-ITI condition in Experiment 2, the
402 prevalence-induced concept change effect was still present and statistically significant. These
403 results support the possibility that a reduced sensitivity to prevalence-induced concept change in
404 older adults could be due to general slowing and the consequential increased spacing between
405 stimuli. It is important to keep in mind, however, that even in the task with the longest ITI (the
406 ethics task in experiment 1) the effect of prevalence-induced concept change remained present
407 and statistically significant. Therefore, our evidence suggests that while prevalence-induced
408 concept change is sensitive to the spacing between stimuli, it is not dependent on it as a cognitive
409 phenomenon.

410
411
412
413
414

Experiment 3

In Experiment 3, we tested whether reduced concept change in older adults in Experiment 1 was due to a speed-accuracy trade-off, where they sacrificed faster responses for more accurate ones. If this was the case, the difference in response times we observed between young adults and older adults might not be due to a processing limitation (i.e., general slowing), but rather to a more deliberate and cautious approach to the task that prioritized accuracy (or perhaps consistency) over response speed.

To test whether speed-accuracy trade-offs of this kind could affect sensitivity to prevalence-induced concept change, we tested young adults on a modified version of the Dots Task, where we provided additional credit for correct responding and block-to-block performance feedback. We hypothesized that participants would slow down their responses to prioritize accuracy in order to avoid losing credit and that this strategy would result in reduced prevalence-induced concept change.

Method

Participants

We recruited 50 young adults ($M_{\text{age}} = 21.27$; $s_{\text{age}} = 1.87$) from the university participation pool. All participants met the inclusion criteria mentioned in the Method section for Experiment 1 and Experiment 2 and half of the participants were randomly assigned to either the decreasing prevalence condition or the stable prevalence condition, in a counterbalanced order.

Materials and Procedure

Rewarded Dots Task. This version of the Dots Task offered participants rewards in the form of class credits for correctly identifying dots as blue or purple. The task was the same as Experiment 1, with an ITI of 500ms. Because presenting performance feedback on each trial

440 would increase the space between stimuli, effectively increasing the ITI, feedback was provided
441 at the end of each block.

442 Participants began each session with 4 credits. With each incorrect response, participants
443 lost 0.0025 credits (conveyed to them by saying 20 mistakes would cost them 0.05 credits). The
444 minimum number of credits they could receive was 2 credits, to ensure that all participants were
445 fairly compensated for their time. Between each block, text appeared that indicated that the block
446 was finished, which block the participant was now at, the number of errors they made, and the
447 amount of credits they had remaining. The procedure for this Experiment was the same as in
448 Experiment 1 and 2, except that all participants were compensated with the participation pool
449 credits they earned during the Rewarded Dots Task for participating in the study.

450 *Analysis*

451 The main analysis consisted of two general binomial mixed-effects models. The first
452 model compared young adults in Experiment 1 and Experiment 3, predicting response using study
453 (Experiment 1 or Experiment 3), condition (stable prevalence or decreasing prevalence), trial
454 number, and colour strength as fixed effects, a random slope of trial, and a random intercept for
455 each participant. All main effects and interactions were explored. The follow-up model used the
456 same predictors, but in Experiment 3 only. In all models, trial and stimulus strength were put on a
457 scale between one and zero.

458 Two one-way ANOVA were also conducted comparing response times between young
459 and older adults in the Dots Task in Experiment 1 vs. the Rewarded Dots task in Experiment 3, as
460 a check to demonstrate that our experimental manipulation affected response strategies, using
461 response time as a proxy.

462

463 **Results**

464 ***Choice Data***

465 Choice data for Experiment 3 are presented in Figure 5B. We found a significant four-way
466 interaction between study, condition, trial, and colour strength ($\beta = -7.24$, $SE = 1.98$, $p = .0002$,
467 95% CI = [-11.10, -3.37]), such that the effect of prevalence-induced concept change was
468 significantly smaller for young adults in Experiment 3 compared to those in Experiment 1. A
469 follow-up analysis revealed a significant interaction between condition, trial, and colour on
470 response for participants in Experiment 3 ($\beta = 16.41$, $SE = 1.16$, $p < .0001$, 95% CI = [14.10,
471 18.70]). However, the size of this effect was smaller than the effect size observed in young adults
472 during Experiment 1 ($\beta = 22.54$ in Experiment 1 vs. $\beta = 16.41$ in Experiment 3).

473 ***Response Times***

474 We found a significant main effect of experiment on response times for both sets of young
475 adults ($F(1, 128) = 32.87$, $p < .0001$, 95% CI = [0.08, 0.16], Mean difference_{YoungExp3- YoungExp1} =
476 0.12 seconds), such that younger adults in Experiment 3 took more time responding than younger
477 adults in Experiment 1. Furthermore, we found a significant main effect of age group between
478 young adults in Experiment 3 and older adults in Experiment 1 ($F(1, 128) = 17.21$, $p < .0001$,
479 95% CI = [0.08, 0.23], Mean difference_{OldExp1- YoungExp3} = 0.16 seconds), such that older adults
480 still responded slowest overall across experiments.

481 **Discussion**

483 In this experiment, we aimed to demonstrate that participants who engaged in a speed-
484 accuracy trade-off would demonstrate reduced prevalence-induced concept change (as observed
485 in older adults in Experiment 1). Our results support this view. Young adults in Experiment 3
486 responded more slowly than young adults in Experiment 1 and, in turn, were less sensitive to
487 prevalence-induced concept change (i.e., they made more accurate judgements). Importantly,
488 these differences were not confounded by differences in the spacing of stimuli.

489 These findings suggest that when participants have a vested interest in maximizing
490 accurate responding, they can reduce—though not eliminate—prevalence-induced concept
491 change. The results line up with the older adults’ finding in Experiment 1 and indicate that
492 reduced sensitivity to prevalence-induced concept change in older adults may not only result from
493 general slowing, but may be due to a difference in how older adults approached the task; namely,
494 by trading speed for accuracy in their responses.

496 **General Discussion**

497 The purpose of this study was to investigate how prevalence-induced concept change
498 affects judgements of older adults across two conceptual domains: perception and ethics. Based
499 on previous findings (e.g., Nassar et al., 2016; Lindenberger & Mayr, 2015), we hypothesized that
500 older adults would either be less sensitive (H1) or more sensitive (H2) to prevalence-induced
501 concept change than younger adults. Our results support H1, demonstrating that older adults were
502 less sensitive to prevalence-induced concept change in their judgements about the colours of dots
503 and not significantly affected by the phenomenon in their ethical judgements about fictitious
504 research proposals. We furthermore predicted that this reduction would be explained by increased
505 perseveration on the part of older adults, reflected in estimated weights in a computational model
506 (Wilson, 2018). However, we failed to find meaningful age differences in model parameters.
507 Thus, we conducted exploratory analyses to test the possibility that the observed behavioural
508 differences in sensitivity to prevalence-induced concept change might be due to differences in the
509 spacing between stimuli between young and older adults. If this were the case, we hypothesized
510 that these differences would be due to general slowing (Verhaeghen & Cerella, 2002) and/or a
511 speed-accuracy trade-off (Starns & Ratcliffe, 2010; 2012; Salthouse, 1979) on the part of older
512 adults. To test these hypotheses we conducted two follow-up experiments: In Experiment 2, we
513 changed task and increased the time between stimuli without changing participants motivation; In

514 Experiment 3, we kept the spacing between stimuli constant but manipulated the decision process
515 by incentivizing accurate responses. In both cases, we found notable reductions in prevalence-
516 induced concept change in younger adults.

517 There are three important messages that can be taken from these experiments. The results
518 of the first experiment suggest that older adults' concepts seem to be more stable than those of
519 younger adults when faced with a changing task environment. This finding dovetails nicely with a
520 body of research demonstrating that older adults have greater difficulty than younger adults
521 updating behaviour despite changes in the environment (Eppinger, Hämmerer, & Li, 2011; Nassar
522 et al., 2016; Hämmerer et al., 2019). As Wilson (2018) has suggested, the types of serial
523 judgements where prevalence-induced concept change may affect judgments can be thought of as
524 a form of implicit learning, where the underlying state of a stimuli (e.g., the average blueness or
525 ethicality) is implicitly estimated based on recently seen instances of the concept (cf. Cleeremans,
526 Destrebecqz, & Maud Boyer, 1998; Nassar et al., 2012; McGuire, Nassar, Gold, & Kable, 2014;
527 Wilson & Niv, 2012). From this perspective, older adults may have more difficulty learning these
528 latent states of stimuli and default to their original responses (Nassar et al., 2016; Bruckner et al.,
529 2020). Notably, in most task environments these impairments in inferences about latent states are
530 associated with performance deficits. Even outside the lab, it is not difficult to see how a
531 difficulty in learning from uncertain environment may pose significant problems in day-to-day
532 life. In contrast, in the current task, older adults' reduced sensitivity to environmental statistics
533 was *protective* against some of the negative consequences that could be associated with
534 prevalence-induced concept change (e.g., claiming something is ethical, when you previously said
535 it was not). Thus, while cognitive ageing has negative consequences in many contexts (learning,
536 memory, etc.), it may have unexpected benefits for judgement and decision-making, such as in
537 the case of prevalence-induced concept change.

538 Second, prevalence-induced concept change is sensitive to the spacing between stimuli.

539 As demonstrated in Experiment 2, prevalence-induced concept change can be meaningfully
540 reduced by increasing the ITI from 500 ms (in Experiment 1) to 2000 ms. As such, it is possible
541 that the reduction in prevalence-induced concept change observed in older adults might be a
542 byproduct of general slowing: As people age, they respond more slowly and, in tasks where the
543 spacing of stimuli affects the prevalence-induced effect, they experience less prevalence-induced
544 concept change. On the one hand, this makes intuitive sense when we consider that in real life
545 instances of categories often change prevalence at a rate of days or months and not seconds. As
546 such, we would expect that the magnitude of the prevalence-induced effect would decrease in
547 accordance with the time it has to take effect (i.e., a smaller effect that carries out over days
548 nevertheless affects our judgements). In the same sense then, this finding also highlights the
549 automatic nature of concept change as a process whose effect is greatest when information is
550 processed rapidly. On the other, more work is needed to determine the degree to which the
551 spacing of stimuli affects prevalence-induced concept change in the lab and in the real world. In
552 the current study, despite reductions in prevalence-induced concept change, changes in ITI did not
553 affect the statistical significance of the effect. As such, it may prove beneficial to design an
554 experimental paradigm that is less sensitive to subtle differences in stimuli spacing and that more
555 readily mimics real-world instances of prevalence-induced concept change (cf. Yarkoni, 2019).

556 Third, prevalence-induced concept change can be reduced if participants are motivated to
557 respond accurately. This point fits within Kahneman's (2003; cf. 2011) dual-system framework,
558 where inconsistencies in our judgments can be tapered if we engage in what he calls "System 2
559 Thinking" (slower, more effortful, thinking). That is, when young adults in Experiment 3—a
560 population that is not affected by general slowing—responded more slowly relative to young
561 adults in the Dots Task in Experiment 1, they also experienced less prevalence-induced concept
562 change and were more consistent and accurate in their judgements overall. This finding provides

563 support for the view that older adults engaged in a speed-accuracy trade-off in Experiment 1 that
564 led to a reduction in their sensitivity to prevalence-induced concept change. If this is the case, it is
565 particularly interesting that older adults engaged in this trade-off without external incentive, while
566 young adults required substantial motivation to do so (i.e., to be offered double course-credits for
567 their participation).

568

569 **Future Directions**

570 The results of the current study answer our initial research question and go beyond it to
571 describe two general factors that may influence one's sensitivity to prevalence-induced concept
572 change. However, this work is but a starting point in describing the intricacies of how and why
573 concept change occurs.

574 Future work should focus on better understanding the relationship between deliberation
575 and prevalence-induced concept change. In this paper, we have referred to a certain style of
576 decision-making as a speed-accuracy trade-off: one where participants traded fast responses for
577 more accurate ones. However, the history of this term within the ageing literature refers more
578 narrowly to the observation that older adults show disproportionately longer response times and
579 favour accuracy when presented with feedback (see Salthouse, 1979). This process is likely
580 different than the one used by older adults in our sample, who responded more slowly and more
581 accurately, without feedback. The same can be said of young adults in Experiment 3, who
582 received unspecific feedback after long delayed fashion (at the end of each block). As such, it
583 remains unclear *how* participants use these slower response times to adapt behaviour: Do longer
584 response times simply extend the space between stimuli and thus reduce concept change, in line
585 with our results from Experiment 2, or do participants use this extra time to make more confident
586 choices that increase accuracy when motivated to do so?

587 Taken together, these points emphasize the need for future work that focuses on studying
588 prevalence-induced concept change in more ecologically valid settings, where the temporal
589 spacing of stimuli mimics that which we would expect outside of the lab (e.g., hours, days,
590 weeks) and people's motivations are taken into account.

591

592 **Conclusion**

593 The current study shows that as we age, our judgements about concepts become more
594 rigid as we face a changing world. While older adults are still generally susceptible to prevalence-
595 induced concept change in basic perceptual tasks, they seem to resist it entirely in higher-order
596 concept judgements about ethics. To shed light more on the cognitive and motivational
597 mechanisms underlying these age-related changes we performed two follow-up experiments in
598 younger adults. The results demonstrate that prevalence-induced concept change is automatic in
599 nature, occurring most prominently when information is processed quickly (Experiment 1).
600 Furthermore, we show that concept change is reduced when subjects are motivated to make slow
601 and deliberate judgements. Taken together, our results suggest that when making judgements
602 older adults engage more in this slow and effortful response mode. Doing so results in a
603 conceptual rigidity that can be beneficial in curtailing biases in judgement that result from
604 changes in the prevalence of events in the environment.

605 The current findings have real-world relevance when considering the degree to which
606 older adults' use of concepts will come to affect the future direction of our society. As we age, it
607 seems our concepts remain more stable, even if the world around us presents us with continued
608 reason to change them. It is in this sense that the quote at the beginning of this paper earns its
609 relevance: the more things (our age and our environment) change, the more they (our concepts)
610 stay the same.

611

Significance

The current research is significant for at least three reasons. First, we directly replicate the prevalence-induced concept change effect in young adults across two conceptual levels (Levari et al., 2018). Second, we demonstrate how older adults' concept space is more resilient to concept change than that of younger adults, a fact worth bearing in mind given that older adults' judgements risk to become increasingly important in coming years. Third, we show that prevalence-induced concept change is automatic in nature, occurring most prominently when information is processed quickly, whereas concept change is reduced when people are motivated to make slow and deliberate judgements. In line with this point, we demonstrate that older adults engage in such slow, deliberate, decision-making and, by doing so, demonstrate a conceptual rigidity that can be beneficial in curtailing biases in judgement that result from changes in the prevalence of events in the environment.

- 626 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects
627 Models Using lme4. *Journal of Statistical Software*, 67(1).
628 <https://doi.org/10.18637/jss.v067.i01>
- 629 Bruckner, R., Nassar, M., Li, S.C., & Eppinger, B. (2020). Default beliefs guide learning
630 under uncertainty in children and older adults. <https://psyarxiv.com/nh9bq/>
- 631 Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the
632 front. *Trends in Cognitive Sciences*, 2(10), 406–416. [https://doi.org/10.1016/S1364-6613\(98\)01232-7](https://doi.org/10.1016/S1364-6613(98)01232-7)
- 634 Cole, B. L., Lian, K. Y., & Lakkis, C. (2006). The new Richmond HRR pseudoisochromatic
635 test for colour vision is better than the Ishihara test. *Clinical and Experimental*
636 *Optometry*, 89(2), 73-80. <https://doi.org/10.1111/j.1444-0938.2006.00015.x>
- 637 Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision*
638 *Making, Affect, and Learning: Attention and Performance XXIII*, 1–26.
639 <https://doi.org/10.1093/acprof:oso/9780199600434.003.0001>
- 640 Eppinger, B., Hämmerer, D., & Li, S. C. (2011). Neuromodulation of reward-based learning
641 and decision making in human aging. *Annals of the New York Academy of Sciences*,
642 1235, 1. <https://doi.org/10.1111/j.1749-6632.2011.06230.x>
- 643 Eppinger, B., Heekeren, H. R., & Li, S. C. (2015). Age-related prefrontal impairments
644 implicate deficient prediction of future reward in older adults. *Neurobiology of Aging*,
645 36(8), 2380–2390. <https://doi.org/10.1016/j.neurobiolaging.2015.04.010>
- 646 Eppinger, B., Nystrom, L. E., & Cohen, J. D. (2012). Reduced sensitivity to immediate
647 reward during decision-making in older than younger adults. *PLoS ONE*, 7(5), 1–10.
648 <https://doi.org/10.1371/journal.pone.0036953>
- 649 Eppinger, B., Walter, M., Heekeren, H. R., & Li, S. C. (2013). Of goals and habits: Age-
650 related and individual differences in goal-directed decision-making. *Frontiers in*
651 *Neuroscience*, 7(7 DEC), 1–14. <https://doi.org/10.3389/fnins.2013.00253>
- 652 Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E. J., Derrfuss, J., Imperati, D., & Brown,
653 S. (2011). The speed-accuracy tradeoff in the elderly brain: a structural model-based
654 approach. *Journal of Neuroscience*, 31(47), 17242-17249.
655 <https://doi.org/10.1523/JNEUROSCI.0309-11.2011>
- 656 Hämmerer, D., Müller, V., & Li, S. C. (2014). Performance monitoring across the lifespan:
657 Still maturing post-conflict regulation in children and declining task-set monitoring in
658 older adults. *Neuroscience and Biobehavioral Reviews*, 46(P1), 105–123.
659 <https://doi.org/10.1016/j.neubiorev.2014.06.008>
- 660 Hämmerer, D., Schwartenbeck, P., Gallagher, M., FitzGerald, T. H. B., Düzel, E., & Dolan,
661 R. J. (2019). Older adults fail to form stable task representations during model-based
662 reversal inference. *Neurobiology of Aging*, 74, 90–100.
663 <https://doi.org/10.1016/j.neurobiolaging.2018.10.009>

- 664 Harrison XA, Donaldson L, Correa-Cano ME, Evans J, Fisher DN, Goodwin CED,
665 Robinson BS, Hodgson DJ, Inger R. (2018). A brief introduction to mixed effects
666 modelling and multi-model inference in ecology. *PeerJ* 6, e4794
667 <https://doi.org/10.7717/peerj.4794>
- 668 Hollenbeck, J. R., & Wright, P. M. (2017). Harking, Sharking, and Tharking: Making the
669 Case for Post Hoc Analysis of Scientific Data. *Journal of Management*, 43(1), 5–18.
670 <https://doi.org/10.1177/0149206316679487>
- 671 Kahneman, D. (2011). *Thinking Fast and Slow*. Farrar, Straus and Giroux.
- 672 Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics.
673 *American Economic Review*, 93(5): 1449-1475.
674 <https://doi.org/10.1257/000282803322655392>
- 675 Levari, D. E., Gilbert, D. T., Wilson, T. D., Sievers, B., Amodio, D. M., & Wheatley, T.
676 (2018). Prevalence-induced concept change in human judgment. *Science*, 360(6396),
677 1465–1467. <https://doi.org/10.1126/science.aap8731>
- 678 Li, S. C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: from neuromodulation
679 to representation. *Trends in cognitive sciences*, 5(11), 479-486.
680 [https://doi.org/10.1016/S1364-6613\(00\)01769-1](https://doi.org/10.1016/S1364-6613(00)01769-1)
- 681 Lindenberger, U., & Mayr, U. (2015). Cognitive Aging : Is There a Dark Side to
682 Environmental, 18(1), 7–15. <https://doi.org/10.1016/j.tics.2013.10.006>.Cognitive
- 683 Mayr, U., Spieler, D. H., & Hutcheon, T. G. (2015). When and why do old adults outsource
684 control to the environment? *Psychology and Aging*, 30(3), 624–633.
685 <https://doi.org/10.1037/a0039466>
- 686 Mayr, U., Spieler, D. H., & Kliegl, R. (2001). *Ageing and Executive Control*. Psychology
687 Press.
- 688 McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable
689 influences on learning rate in a dynamic environment. *Neuron*, 84(4), 870-881.
690 <https://doi.org/10.1016/j.neuron.2014.10.013>
- 691 Nassar, M. R., Bruckner, R., Gold, J. I., Li, S. C., Heekeren, H. R., & Eppinger, B. (2016).
692 Age differences in learning emerge from an insufficient representation of uncertainty in
693 older adults. *Nature Communications*, 7(May 2015).
694 <https://doi.org/10.1038/ncomms11609>
- 695 Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasley, B., & Gold, J. I. (2012).
696 Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature*
697 *neuroscience*, 15(7), 1040. <https://doi.org/10.1038/nn.3130>
- 698 Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory
699 aging and brain maintenance. *Trends in cognitive sciences*, 16(5), 292-305
700 <https://doi.org/10.1016/j.tics.2012.04.005>
- 701 Salthouse, T. A. (1979). Adult age and the speed-accuracy trade-off. *Ergonomics*, 22(7),
702 811–821. <https://doi.org/10.1080/00140137908924659>

- 703 Salthouse, T. A. (1992). Influence of processing speed on adult age differences in working
704 memory. *Acta psychologica*, 79(2), 155-170. [https://doi.org/10.1016/0001-](https://doi.org/10.1016/0001-6918(92)90030-H)
705 [6918\(92\)90030-H](https://doi.org/10.1016/0001-6918(92)90030-H)
- 706 Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition.
707 *Psychological review*, 103(3), 403. <https://doi.org/10.1037/0033-295X.103.3.403>
- 708 Samanez-Larkin, G. R., & Knutson, B. (2015). Decision making in the ageing brain: changes
709 in affective and motivational circuits. *Nature Reviews Neuroscience*, 16(5), 278-289.
710 <https://doi.org/10.1038/nrn3917>
- 711 Mather, M., & Harley, C. W. (2016). The locus coeruleus: Essential for maintaining
712 cognitive function and the aging brain. *Trends in cognitive sciences*, 20(3), 214-226.
713 <https://doi.org/10.1016/j.tics.2016.01.001>
- 714 Spieler, D. H., Mayr, U., & LaGrone, S. (2006). Outsourcing cognitive control to the
715 environment: Adult age differences in the use of task cues. *Psychonomic Bulletin and*
716 *Review*, 13(5), 787–793. <https://doi.org/10.3758/BF03193998>
- 717 Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise:
718 Boundary optimality in the diffusion model. *Psychology and Aging*, 25(2), 377–390.
719 <https://doi.org/10.1037/a0018022>
- 720 Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary
721 optimality with both trial-limited and time-limited tasks. *Psychonomic bulletin &*
722 *review*, 19(1), 139-145. <https://doi.org/10.3758/s13423-011-0189-3>
- 723 Statistics Canada. (2019). *What will the population of Canada look like in 2068?* Retrieved
724 from: <https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2019050-eng.htm>.
- 725 U.S. Census Bureau (2018). The 2017 National Population Projections. Retrieved from
726 <https://www.census.gov/programs-surveys/popproj.html>
- 727 Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: a review of
728 meta-analyses. *Neuroscience and Biobehavioral Reviews*, 26, 1–9.
- 729 Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion
730 model: An empirical validation. *Memory & cognition*, 32(7), 1206-1220.
731 <https://doi.org/10.3758/BF03196893>
- 732 Wilson, R. C., & Niv, Y. (2012). Inferring relevance in a changing world. *Frontiers in human*
733 *neuroscience*, 5, 189. <https://doi.org/10.3389/fnhum.2011.00189>

734 **Acknowledgments**

735
736 **Funding:** This work was supported by a grant from the Social Sciences and Humanities
737 Research Council of Canada (SSHRC) [CGS-M 2019] awarded to S.D. B.E. received
738 support from the German Research Foundation (DFG; SFB 940/2 B7), as well as the
739 Natural Sciences and Engineering Research Council of Canada (NSERC; N01882).
740 Furthermore, this research was undertaken, in part, thanks to funding from the Canada
741 Research Chairs program.

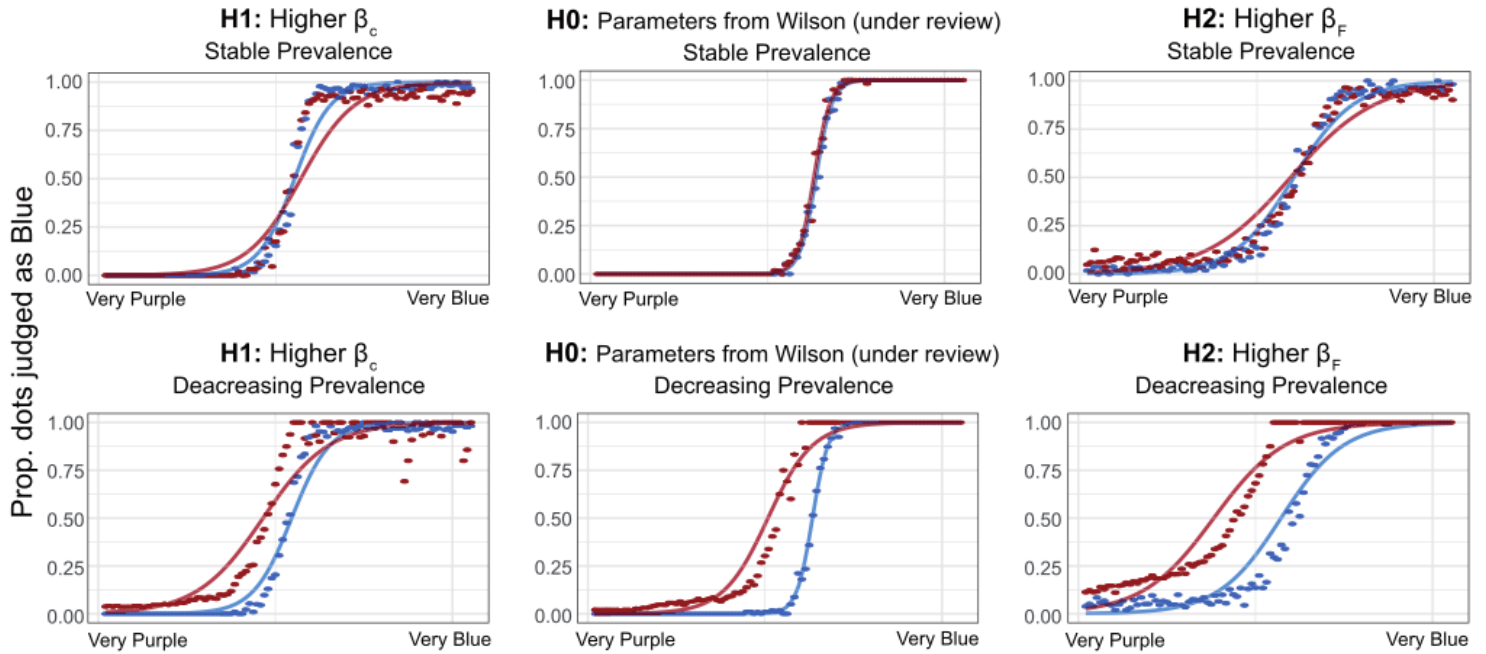
742
743 **Author contributions:** S.D. conceived of the three experiments, oversaw testing,
744 analysed the data, and composed the initial version of the manuscript. C.N. tested
745 participants and provided critical comments on the manuscript in its early stages. D.L.
746 provided guidance on the project and commented on the manuscript. R.W. helped
747 implement the computational models and provided comments on the manuscript. B.E.
748 supervised the project and offered crucial revisions.

749
750 **Competing interests:** The authors declare no conflict of interest.

751
752 **Data and materials availability:** All task code, raw data, and analysis scripts are
753 available at <https://osf.io/pbzv3/>.

754

756 **Figure 1 Simulations**

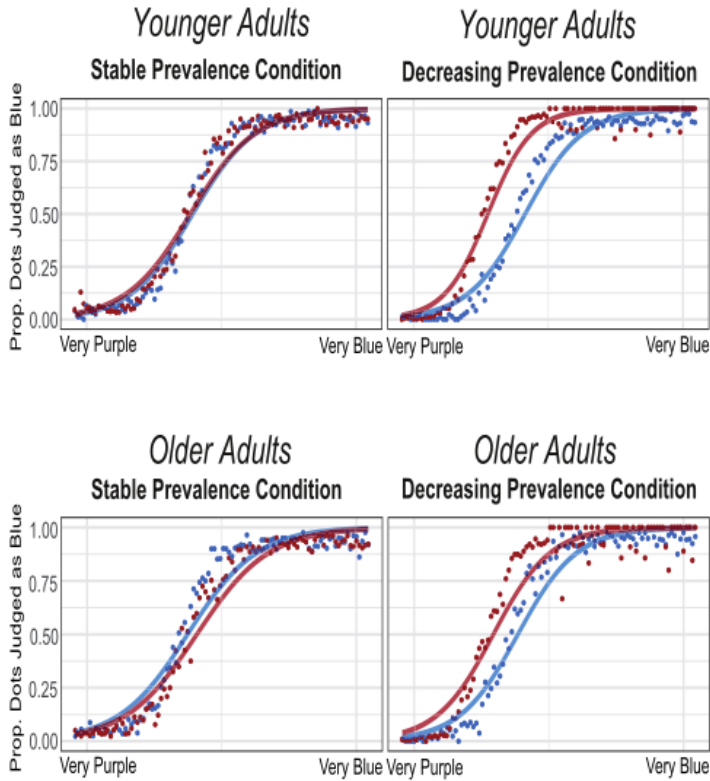


757

758 **Fig. 1. Simulated data representing three hypotheses.** H1 (far left) shows decreased
759 prevalence-induced concept change if older adults have a higher average β_C , representing a
760 greater effect of past response on current response (i.e., more perseverance). H2 (far right) shows
761 increased prevalence-induced concept change if older adults have higher absolute β_F values
762 (negative in reality), representing a greater effect of past stimuli on current response (i.e., more
763 outsourcing). Finally, H0 (center) represents a scenario where older and younger adults do not
764 differ in their sensitivity to prevalence-induced concept change.
765

Figure 2 Experiment 1

A) Dots task (colour judgements)



B) Ethics task (ethicality judgements)

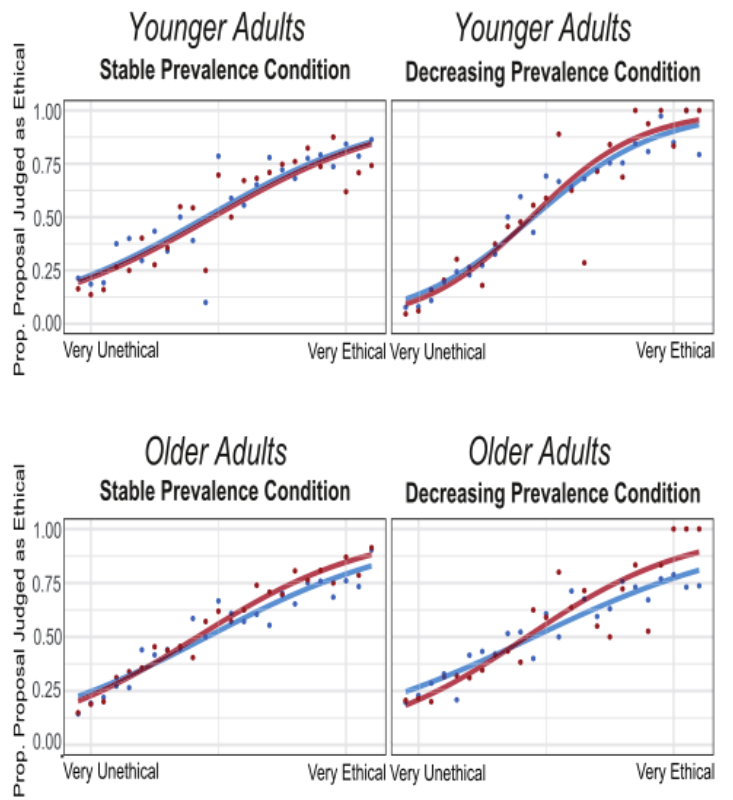
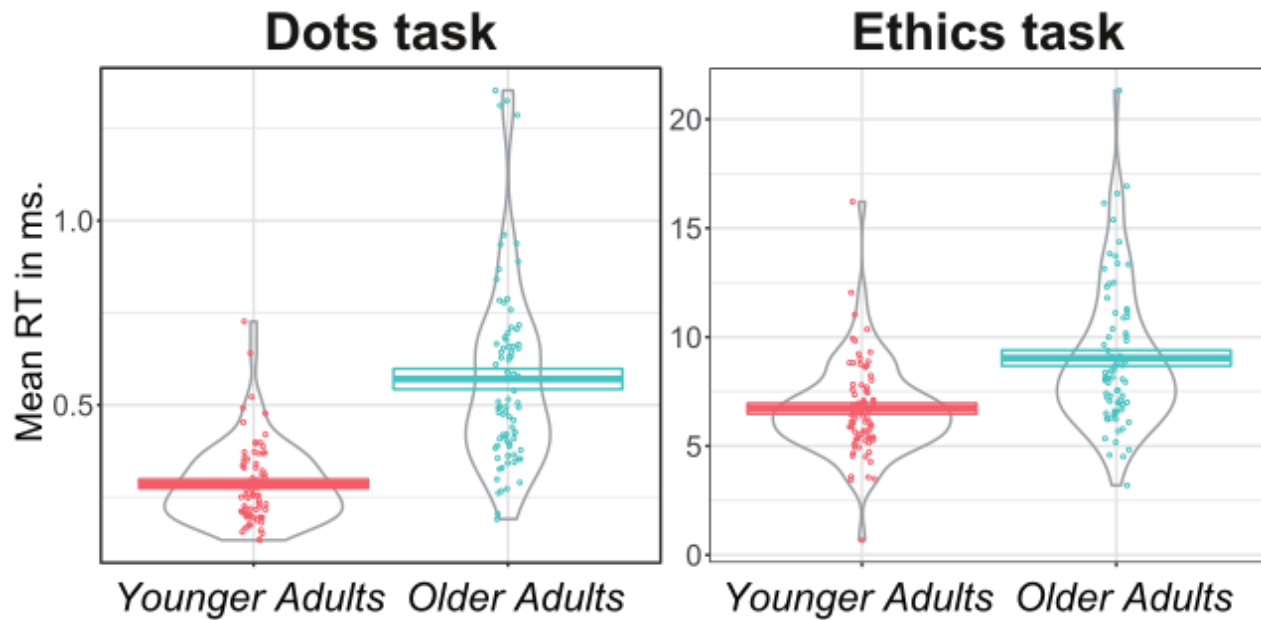


Fig 2. Concept judgements in (A) the Dots Task and (B) the Ethics Task. The x-axis represents stimulus strength: blueness in the Dots Task and ethicality in the Ethics task. The y-axis represents the percentage of dots/proposals judged as blue/ethical. Points represent the percent of choices for the corresponding stimulus strength, averaged across subjects within that cell. Curves represent fitted binomial regression curves.

Figure 3

Reaction times Experiment 1

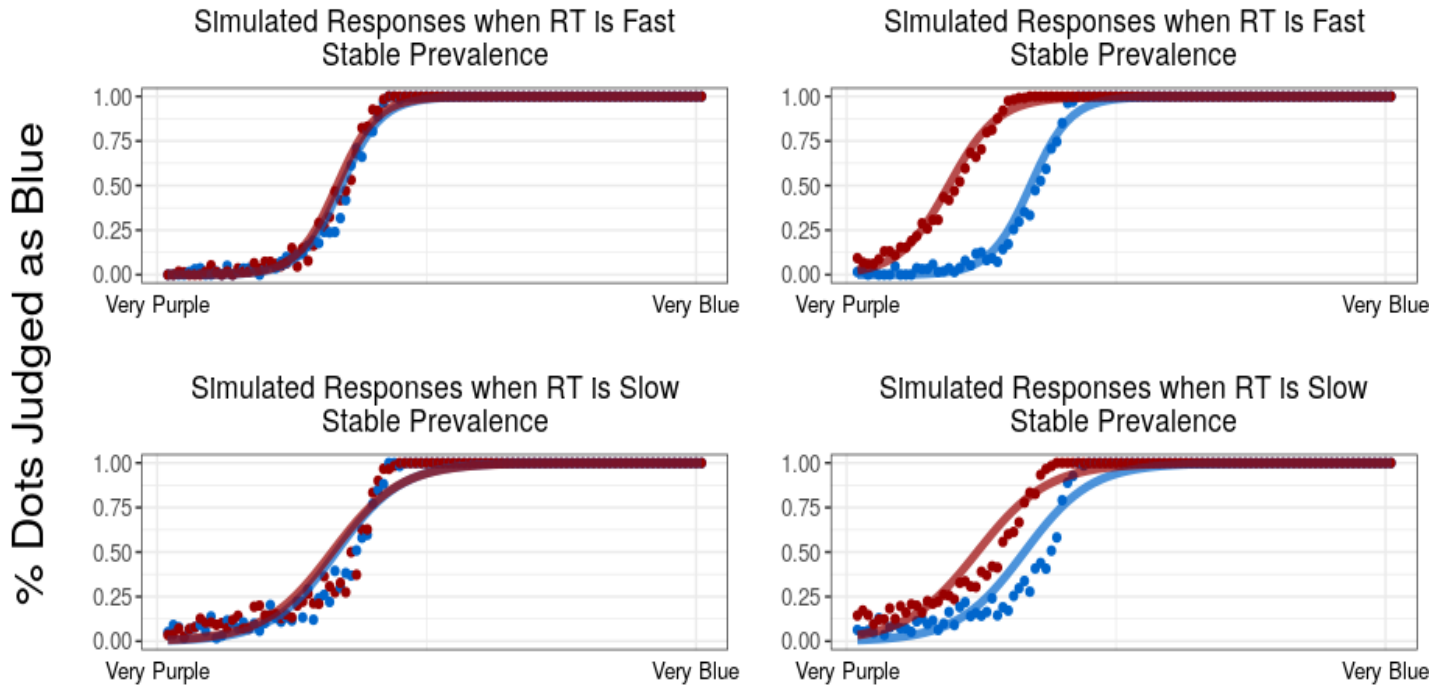


773

774 **Fig. 3. Pirate plots of mean response times in both age groups across both tasks.** X-axis is the
775 age group. Y-axis is the mean response time in seconds per participant (note the difference in
776 scale across tasks). Each point represents an individual participant's mean response time. Boxes
777 represent standard error and horizontal lines represent group means.

778

Figure 4

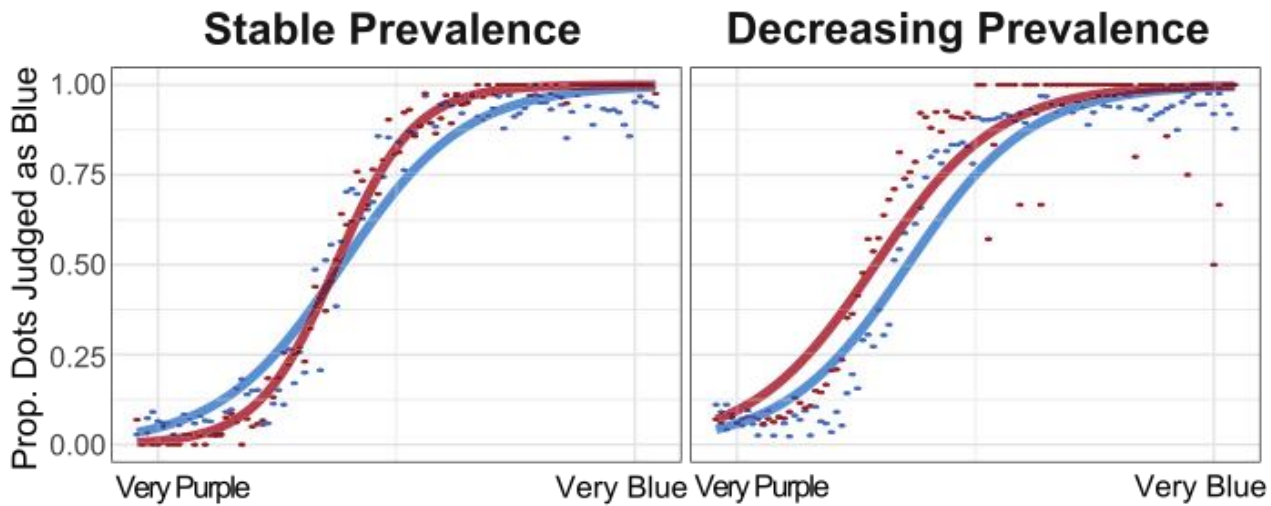


779

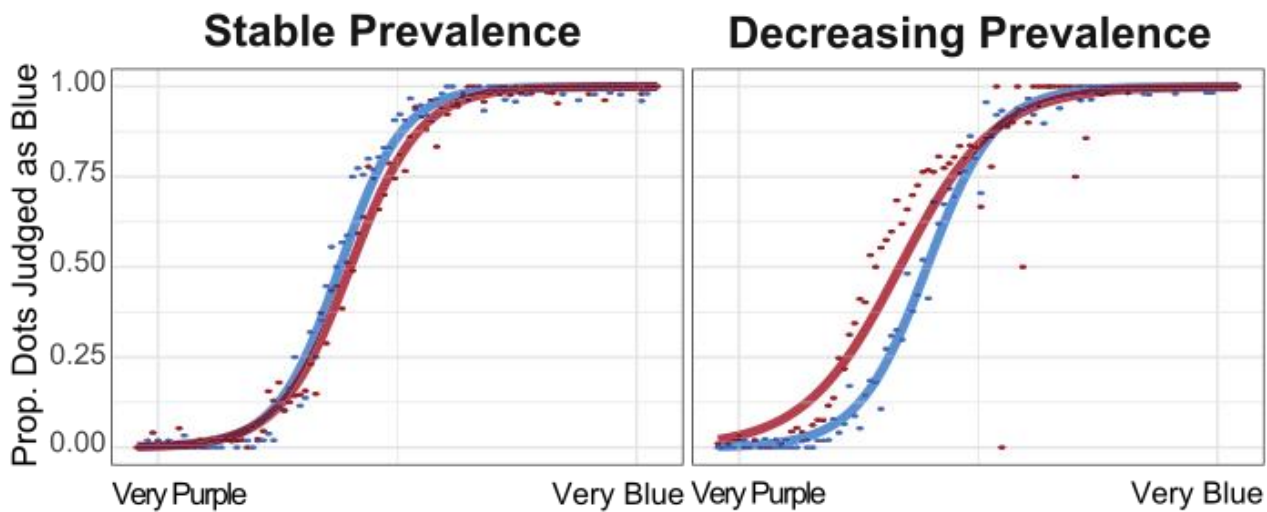
780 **Fig. 4. Exploratory data simulation using a modified model that incorporates response**
781 **times.** In both fast and slow response time groups, the same free parameters are used; only
782 response times vary. Note that distributions are shifted to the left because parameters are not
783 optimised for this modified model.
784

Figure 5

A) Experiment 2 (ITI manipulation)



B) Experiment 3 (Incentive manipulation)



785

786 **Fig. 5. Concept judgements in (A) ITI-modified Dots Task and (B) the Rewarded Dots**
787 **Task.** The x-axis represents stimulus strength (i.e., blueness). The y-axis represents the
788 percentage of dots judged as blue. Points represent the percent of choices for the corresponding
789 stimulus strength, averaged across subjects within that cell. Curves represent fitted binomial
790 regression curves. Blue points and lines represent the first 200 trials in the Dots Task and red ones
791 represent the final 200 trials in the Dots Task.

792

Supplementary Materials

HRR Colour Vision Test

The HRR colour vision test is a short screening test to ensure that participants' colour vision is adequate for the Dots Task. Specifically, this test was included to ensure that participants did not differ in how they experienced the stimuli in the Dots Task. Furthermore, given that many older adults experience sensory deficits that compromise their ability to discriminate colours, we also wanted to control for age-related impairments in colour discrimination.

The test contains 24 plates (pages), each displaying either one or two symbols, which can be a circle, a cross, or a triangle, four of which are demonstration plates to explain the task and six of which are screening plates used to classify participants based on their colour vision. The remaining plates are used to grade the severity of certain deficiencies. Only the first 10 plates were used in this study, as is standard in assessing basic colour discrimination (Cole, Lian, & Lakkis, 2006). The symbols on each plate are constructed of coloured dots that would be difficult or impossible to discern if someone were colourblind (Cole, Lian, & Lakkis, 2006).

An experimenter presented the plates to participants one at a time and asked them to identify how many symbols they saw, what the symbols were, and to outline those symbols with a brush. Participants in this study were graded as pass/fail, receiving a failing grade as soon as they either failed to identify one of the symbols or misidentified a symbol. A passing grade was only given if all plates were correctly identified.

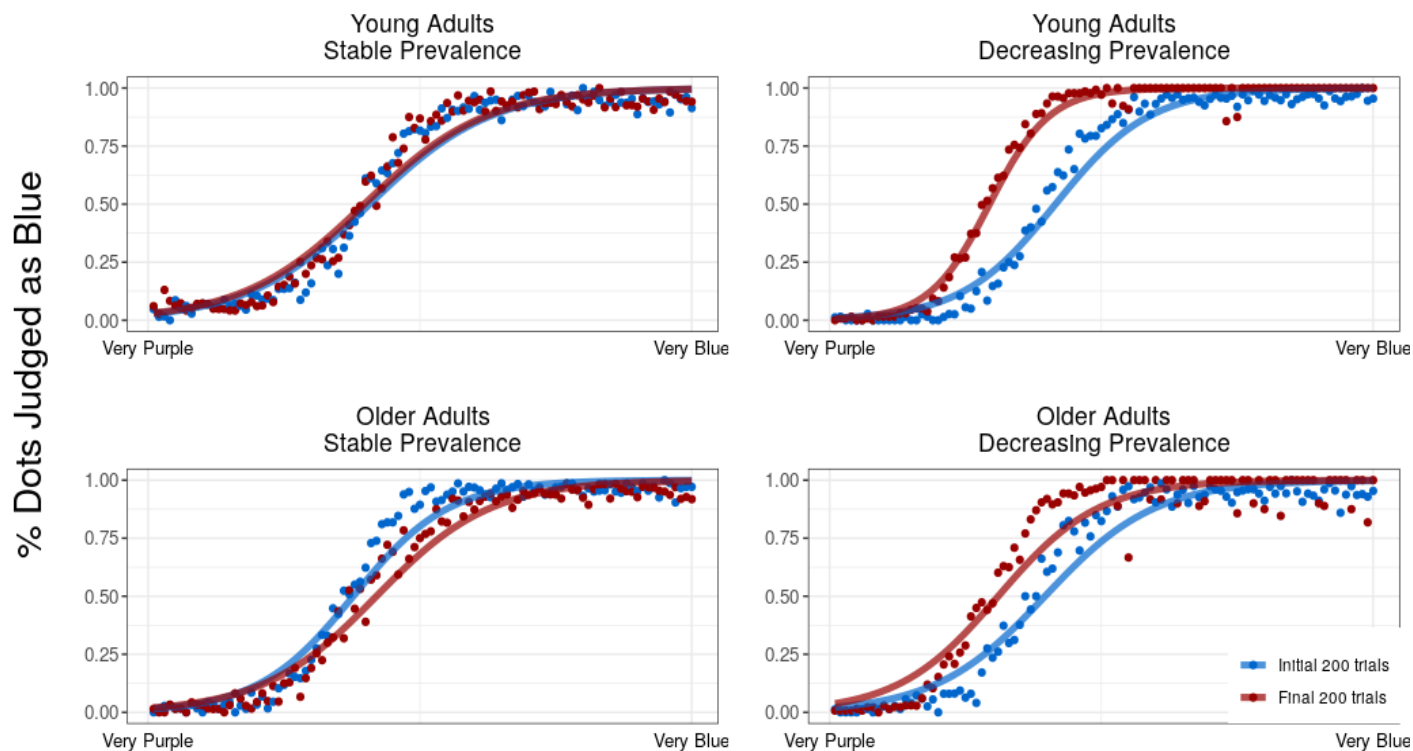
Unsurprisingly, more older adults failed the colour vision test than younger adults (14 older adults vs. 6 younger adults). To test whether these participants affected the overall pattern of results we observed in Experiment 1, we ran the same binomial mixed-effects models described in the main text with them excluded (see Experiment 1 *Statistical Analysis*). However, excluding these participants alone would result in unbalanced data in terms of age group (74 young adults

819 vs. 64 older adults), which is known to contribute to decreased power and imprecise estimates in
820 mixed-effects models (Harrison et al., 2018). As such, we randomly sampled 64 young adults
821 from the 74 who did not fail the colour vision test and ran the models.

822 The patterns of results from the overall binomial regression match those presented in the
823 main text (Experiment 1 Results). Namely, we found a significant four-way interaction between
824 age group, condition, trial, and colour strength ($\beta = 3.72$, $SE = 1.37$, $p = .0064$), such that older
825 adults were less sensitive to prevalence-induced concept change than younger adults (see
826 Supplemental Figure 1). Therefore, to avoid the issue of unbalanced data and to increase power,
827 all participants were included in the regressions presented in the main text.

828

Experiment 1 Choice Data Without Failed HRR Participants



830

831 **Fig. S1. Concept judgements in the Dots Task after participants who failed the HRR colour**
 832 **vision test are excluded and groups are balanced.** The x-axis represents stimulus strength:
 833 blueness in the Dots Task and ethicality in the Ethics task. The y-axis represents the percentage of
 834 dots/proposals judged as blue/ethical. Points represent the percent of choices for the
 835 corresponding stimulus strength, averaged across subjects within that cell. Curves represent fitted
 836 binomial regression curves.
 837

838 **Response Time Modified Computational Model**

839 As a first exploration into the role of response times in prevalence-induced concept
840 change, we modified the original computational model (Wilson, 2018; see Method section main
841 text). Specifically, we kept the majority of the model the same, but weighted the decay parameters
842 such that longer responses would incur slower decay on a trial-by-trial basis:

$$F_t = \lambda_F^{RT_t} F_t + f_t \quad (\text{Eq. 4})$$

$$C_t = \lambda_c^{RT_t} C_t + c_t \quad (\text{Eq. 5})$$

843 We used the same parameters in all cases, but manipulated the response time distributions.
844 For the “Low RT” group, we produced a random normal distribution using the same mean and
845 standard deviation we observed in the young adults data above. For the “High RT” group we did
846 the same, using the mean and standard deviation from the older adults. Qualitatively, we can see
847 these simulated data appear to resemble the empirical distributions shown in Figure 2. Simulated
848 data from this modified model is presented in Figure 4.

849 We used this simulated data as preliminary support that response times may be driving the
850 differences in sensitivity to prevalence-induced concept change we observed between older and
851 younger adults and not participants’ sequential decision-making patterns (see main text).

852
853