Prevalence-Induced Concept Change in Older Adults

Sean Devine

A Thesis
In the Department
of
Psychology

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Arts (Experimental Psychology) at
Concordia University
Montreal, Quebec, Canada

May 2020
© Sean Devine, 2020

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Sean Devine
Entitled: Prevalence-Induced Concept Change in Older Adults

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Experimental Psychology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Andreas Arvanitogiannis Chair
Chair's name

Roberto de Almeida Examiner
Examiner's name

Karen Li Examiner
Examiner's name

Ben Eppinger Supervisor
Supervisor's name

Approved by ______________________________
Chair of Department or Graduate Program Director
(Aaron Johnson)

__________ 2020 ______________________________
Dean of Faculty
(André G. Roy)

ABSTRACT

Prevalence-Induced Concept Change in Older Adults

Sean Devine
Concordia University, 2020

Prevalence-induced concept change describes a cognitive phenomenon whereby judgements about concepts shift as the prevalence of exemplars of that concept changes. For instance, in a task where people have to judge whether the colour of an ambiguously-coloured dot is blue or purple, if the frequency of objectively blue dots in the environment decreases, people judge more dots to be blue than they did initially. While this phenomenon has been explored in young adults, it is unclear how it affects older adults. Past work suggests that older adults simultaneously rely less on internal representations, but that they also tend to perseverate more on cognitive tasks. Thus, the question arises: Do older adults outsource control and their decisions become more susceptible to change or are they more rigid in their judgements than younger and resistant to prevalence-induced concept change?

In the current study, we explore how prevalence-induced concept change affects older adults' lower-level, perceptual, and higher-order, ethical, decision-making. We find that older adults are less sensitive to prevalence-induced concept change than younger adults across both domains. An exploratory analysis is conducted on response times to help elucidate the mechanism(s) underlying these differences. These analyses demonstrate that older adults respond more slowly than younger adults in both tasks. We offer two interpretations of this finding, both with implications for prevalence-induced concept change research more broadly: general slowing and/or a speed-accuracy trade-off. Overall, our results suggest that older adults' judgements about concepts may be less flexible than younger adults' when faced with a changing world.

Acknowledgements

This research was supported by a grant awarded by SSHRC (CGS - M2019) to me, Sean Devine. I would first like to thank Dr. Ben Eppinger who has supported me long before this thesis, has always been available when needed, and who has provided me with excellent research experience that I am sure will serve me far beyond what I could detail in this short section. This thesis would not have been possible without his continued guidance. Second, I would like to thank Cassandra Neumann and Alexa Ruel, who were indispensable in the entire process of conducting this study, from conception to data collection. If it was not for them, there is no doubt in my mind that this thesis would be much worse. Third, I would like to thank all of the participants who were gracious enough to volunteer their time to participate in this experiment. You are the lifeblood of psychological research and all of this would quite literally be impossible without you. Last, but not least, I'd like to thank my partner, Chelsea, who has been a key source of inspiration and encouragement at every step of my academic career and throughout the authorship of this thesis. Thank you.

Table of Contents

List of Figures

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

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

Similarly to most Western countries, by 2068, almost 30% of the Canadian population will be 65 years or older (Statistics Canada, 2019). This increase is accompanied by a parallel decrease of working-age people to near 60% of the population. Together, these demographic changes attest to the fact that older adults are slated to take on an important role in shaping Canadian society. As baby-boomers age, their judgements and decisions about the world will affect society more than ever before and largely influence the future direction of our country. As such, it is critical to explore the cognitive mechanisms that underlie these decisions, such that we might gain a better understanding of how older adults come to make such important decisions and, in turn, how we might improve overall decision-making as we age.

The focus of this thesis will be one such cognitive mechanism: the use of concepts. In order to understand the world, we must all make judgements about concepts. These judgements range from basic perceptual ones, such as if a banana is sufficiently ripe to eat, to more abstract judgements about the world, such as whether some actions are morally right or wrong. It has long been known that both high and low-level judgements like this play an important role in how we relate to the physical world. Even prior to the establishment of psychology as a science, philosopher Immanuel Kant stated that “intuition and concepts constitute [...] the elements of all our cognition” (Kant, 1998). Importantly, however, what we judge is not always fixed; our judgements have to be applied to a changing world. For instance, when deciding to buy a banana or not, we must judge the ripeness of many different batches of bananas that vary in quality when we visit the supermarket. When we do so, we are faced with a dilemma: apply a fixed criterion to changing exemplars, such that we decide not to buy any banana at all if none are sufficiently ripe, or adjust our criterion to what is available, buying a banana we did not before deem ripe. Similarly, when judging the morality of a new law, we must decide whether to apply a fixed judgement of “rightness” or to adapt to changing circumstances.

In a recent paper, Levari and colleagues (2018) explored how we make judgements in a changing world in such a way. In a phenomenon that they termed **prevalence-induced concept change**, they found that as the numbers of exemplars of a given concept increase in the environment (e.g., more ripe bananas), our judgements about that concept (“ripe” vs. “unripe”) change such as to include exemplars that they would otherwise exclude. For example, the authors used a task where participants (young adults) had to serially judge whether individual dots that vary on a spectrum between blue and purple were in fact blue or purple (The Dots Task; see Figure 1). When the relative frequency of objectively coloured dots in the environment were equal and consistent across this task (50% blue dots, 50% purple dots), peoples’ judgements were relatively stable: If they judged a dot to be blue in the first trials, they judged the same dot to be blue in the last trials. However, if the number of blue dots in the environment decreased over the task (50% blue dots in the first trials, but gradually shifted to 4% blue dots in the last trials), participants were more likely to judge purple dots as blue by the end of the experiment. Put simply, when the prevalence of blue dots in the environment changed, the boundary for what counted as “blue” expanded. Thus, as the authors claim, the concept itself changed; hence, prevalence-induced concept change.

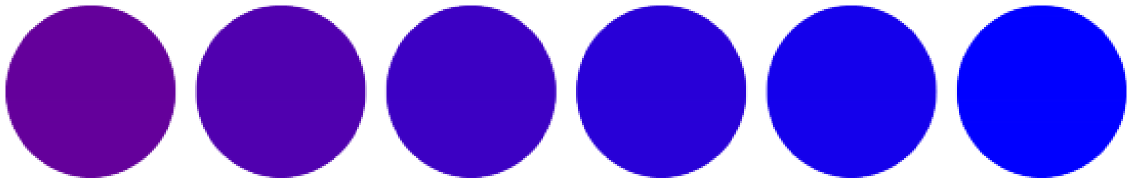

Figure 1. From Levari et al. (2018, Supplemental). Coloured dots in their experiments comprised 100 dots ranging from approximately RGB 100-0-155 (very purple) to RGB 0-0-255 (very blue). Each dot was presented one at a time and participants had to judge whether a given dot was blue or purple.

Critically, Levari et al. (2018) claimed that this change did not only occur for lower-level perceptual phenomena like colour perception, but that it indeed also arose in higher-order cognitive judgements. In Study 6 of their paper, they showed that the same principle applied to judgements about the threateningness of faces. As the prevalence of threatening faces in the environment decreased, people were more likely to judge non-threatening faces as threatening, compared to when the prevalence did not change. Similarly, in Study 7, they showed that prevalence-induced concept change also occurred in ethical judgements. Here, participants took on the role of a member of an internal review board and had to judge a series of research proposals to determine if they were ethical or not. Again, as the number of ethical proposals in the environment decreased, people became more likely to judge a research proposal as unethical than when the prevalence remained stable.

The purported implications of these findings may be substantial. Not only do Levari et al. (2018) suggest that concepts, once formed, are not stable, but that they are subject to continuous update based on information from one's environment. Furthermore, they claim that prevalence-induced concept change offers a cognitive explanation to broader social phenomena where this shift is suspected to take place, such as why people continue to view the world as a dangerous place, despite empirical evidence suggesting just the opposite (Levari et al., 2018).

While this research was the first to give a name to this phenomenon, it is not without its criticism (see Discussion for more detail). One limitation to Levari et al. (2018) work that the current thesis will focus on is its use of a homogenous group of participants, namely young adults. While young adults are a common population to sample from, there are both intuitive and theoretical reasons to be interested in how prevalence-induced concept change affects judgements in older adults. Firstly, as mentioned above, older adults are quickly becoming main decision-makers in Canadian society. If prevalence-induced concept change does generalise to broader social decision-making—as it does to facial and ethical judgements—it would be important to understand how it might differentially affect the judgements and decisions of a segment of the population that has a substantial amount of decision-making power in society and stands to gain even more. Second, prevalence-induced concept change is assumed to take place over long periods of time, during which individuals can observe changes in the prevalence of exemplars and adjust their judgements accordingly. Thus, in real-life, those who are most likely to experience the brunt of prevalence-induced concept change are older adults who have been alive long enough to see varying prevalences of exemplars in the environment and, as a result, adjust their judgements over time. In this sense, it is of obvious ecological interest to explore how older adults might be affected by this phenomenon.

From a theoretical perspective, there are also compelling reasons to think that older adults differ cognitively from younger adults in terms of how they make judgements and decisions. For instance, there is well-documented evidence that older adults differ in terms of their ability to process information and make decisions based on this processed information (Murman, 2015). These differences are in part due to cognitive changes in executive function (Mayr, Spieler, & Kliegl, 200), memory (Lezak, Howieson, Bigler, & Tranel, 2012), and processing speed (Kerchner et al., 2012) that occur naturally in healthy ageing. Older adults also differ in terms of higher-order decision-making strategies, such as motivation, postponement of gratification, and to what degree they value desired outcomes (Sparrow & Spaniol, 2016; Eppinger, Nystrom, & Cohen, 2012). Some of these decision-making trends are hypothesised to be linked to a broader decline in fluid intelligence associated with older age, which may negatively affect decision-making that depends on the learning and updating of new information (Samanez-Larkin & Knutson, 2015; Sparrow & Spaniol, 2016). In one case, these cognitive differences might be protective against prevalence-induced concept change, whereas, in the other, they may increase its effects. Thus, in this thesis, we put forward the following two, opposing, hypotheses:

H1: Older adults are **less** sensitive to prevalence-induced concept change than younger adults

H2: Older adults are **more** sensitive to prevalence-induced concept change than younger adults

In the case of H1, beyond the general deficits in learning mentioned above, previous work suggests that older adults have specific difficulty learning from uncertain outcomes compared to younger adults (Nassar et al., 2016). In computerised tasks, this difficulty manifests as perseverative behaviour, whereby older adults have a tendency to repeat previous responses despite changes in the environment (Bruckner et al., in prep; Eppinger, Walter, Heekeren, & Li, 2013). This perseveration is an indication that older adults are less likely than younger adults to update predictions about the environment, even when doing so would be advantageous (i.e., it would be more rewarding, as in the studies cited above). In terms of prevalence-induced concept change, perseverative behaviour is exactly the opposite of behaviour that would lead to a change in judgements over time. That is, repetition of past choices makes it less likely that a rarer category will be chosen after a shift in prevalence. For example, if the Dots Task mentioned above begins with an even split between blue and purple dots, and if the participant demonstrates perseverative behaviour, they are more likely to stick to their original choices than be swayed by the later change in the relative frequency of purple dots and blue dots. Indeed, this is exactly what Wilson (2018) found when he computationally modeled prevalence-induced concept change, using Levari's et al. (2018) data. In this paper, Wilson (2018) argued that prevalence-induced concept change was the outcome of a sequential decision-making process, whereby participants' choices were in large part governed by the set of past stimuli they observed and the past choices they made. Here, a higher influence of past choice (a greater weight on the past choice parameter in the model) on current behaviour drove prevalence-induced concept change down. Thus, older adults' tendency to perseverate—to be more consistent in their choices—may reduce the effects of prevalence-induced concept change on their judgements (H1).

In the case of H2, results from several recent studies suggest that older adults may be less able to converge on an accurate representation of the current state, particularly if these states are latent (not directly observable) and need to be inferred from experience (Hämmerer et al., 2019; Hämmerer, Müller, & Li, 2014; Eppinger, Heekeren, & Li, 2015). To help compensate for this difficulty in distinguishing task states, older adults preferentially outsource control to the environment rather than rely on (sometimes inaccurate) representations (Mayr, Spieler, & Hutcheon, 2015; Spieler, Mayr, & LaGrone, 2006). As Lindenberger and Mayr (2014) point out, such environmental outsourcing can also bias performance on some cognitive tasks. In the case of prevalence-induced concept change, older adults' tendency to outsource control should increase sensitivity to changes in the prevalence of events in the environment. In other words, this outsourcing of control is likely to lead to increased judgement change and, thus, technically incorrect responses (judging objectively purple dots as blue). As Wilson's (2018) model highlights, an opponent process to the effect of past response discussed in the previous paragraph is the effect of previous stimuli, such that people with a high value on this model weight are more likely to choose the opposite of the past stimulus and, thus, demonstrate more prevalence-induced concept change. In this sense, it is possible that older adults' tendency to outsource control to the environment—that is, to rely more on cues from task stimuli instead of their own representations of, say, the colour blue—increases their sensitivity to prevalence-induced concept change (H2).

Together, these hypotheses paint opposing pictures of older adults' sensitivity to prevalence-induced concept change. On the one hand, H1 predicts that older adults will respond more consistently throughout the whole task than younger adults and, as such, their judgements will remain more stable regardless of changes in the prevalence of exemplars in the environment. H2 on the other hand predicts that older adults' judgements will be more affected by changes in the prevalence of exemplars due to a difficulty in converging on an accurate representation of the concept and, hence, a tendency to outsource control to the environment.

To tease these hypotheses apart, the current study will utilize two experimental paradigms taken from Levari et al. (2018) to explore how prevalence-induced concept change differentially affects older adults' judgements compared to younger adults. In the first, the Dots Task, participants judge the colour of dots that vary between blue and purple. In the second, the Ethics Task, participants judge the ethicality of various (fictitious) research study proposals that range from very unethical to very ethical. Our results support H1 by demonstrating that older adults are less sensitive to prevalence-induced concept change than younger adults in both tasks. An exploratory analysis is conducted on response times to potentially elucidate the cognitive mechanisms underlying these differences in prevalence-induced concept changes between younger and older adults' judgements.

Method

Participants

We recruited 132 participants from the community and the university participation pool, 66 of which were older adults (60 years and older) and 66 of which were younger adults (between 18 to 35 years). Based on self-report, all participants were English-speaking, free of neurological or psychiatric disorders, and free of any cognitive, motor, visual, or other condition(s) that would impede their performance, including but not limited to a history of head trauma with loss of consciousness, organic brain disorders, seizures, or neurosurgical intervention, to sensory deficits (i.e. deafness, blindness; intellectual disability), or self-reported cognitive impairment, and to a recent history of substance abuse. Eleven of these participants (six older adults, five younger adults) were excluded due to failing the HRR pseudo-isochromatic plates colour vision test (Cole, Lian, & Lakkis, 2006), indicating abnormal colour-vision which would preclude them from completing the Dots Task (more details on the HRR colour test below). One participant was excluded for failing to complete both tasks in the study. The number of excluded participants happened to be evenly spread among young and older adults, leaving a final sample of 120 participants, with an even split of 60 young adults (51 women; M_{age} = 21.75; s_{age} = 2.28) and 60 older adults (47 women; M_{age} = 69.78; s_{age} = 5.21). The groups did not significantly differ by sex (β = -0.22, SE = 0.67, p = .7389). In each age group, participants were randomly assigned to either the decreasing prevalence condition (48 women; M_{age} = 45.48; s_{age} = 23.90) or the stable prevalence condition (49 women; M_{age} = 46.05; s_{age} = 24.80), in a counterbalanced order. Conditions did not significantly differ in terms of overall age($F(1, 116)$ = 0.59, p = 0.4450), nor did they differ between age within each age group (i.e., there was no significant interaction between age group and condition in predicting age; $F(1, 116) = 1.37$, $p = .2240$; Stable-YA: $M_{age} = 21.60$, Decrease-YA: $M_{age} = 21.90$, Stable-OA: $M_{age} = 70.50$, Decrease-OA: $M_{age} = 69.06$). Furthermore, the conditions did not significantly differ with regards to participants' sex, either across age groups ($\beta = 0.26$, SE = 0.73, $p = .7182$), nor within age groups ($\beta = -0.64$, SE = 0.95, $p = .5036$). In the decreasing prevalence condition, participants experienced a decreasing prevalence of exemplars in both tasks detailed below. In the stable prevalence condition, the prevalence of exemplars remained the same throughout the entire experiment (see task descriptions below for more details). Participants were either compensated $20 CAD or 2 participation pool credits for participating in the study, based on their preference and whether they were Concordia students. This study protocol was approved by the Concordia Human Research Ethics Committee (certification number 30011191).

Materials

HRR pseudo-isochromatic test for colour vision. 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.

Past work has demonstrated that certain colour preferences change as a result of healthy ageing (e.g., older adults prefer the colour blue less than young adults; Dittmar, 2001). Furthermore, these changes are thought to be attributable to age-related alterations in colour discrimination. Given that results from the Dots Task assume that the only source of bias in response should be prevalence-induced, it was important to ensure that all participants did not differ with regards to their baseline colour preferences. In practice, given that many older adults experience sensory deficits that compromise their ability to discriminate colours (cf. Fiorentini, Porciatti, Morrone, & Burr., 1996)—which in turn impact their colour preferences (Dittmar, 2001)—this meant controlling for age-related impairments in colour discrimination using the HRR colour vision test.

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 (see Appendix A). 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.

It is worth noting that, despite the control in colour discrimination impairments that the HRR colour test affords, there are still likely differences in baseline colour preferences between young and old adults. Research has suggested that various other elements than colour discrimination may affect age-related colour preferences. For instance, limitations in visual imagery abilities in old age have been cited as a potential contributor to colour preference differences between young and old adults (Dittmar, 2001). We attempted to control for some of these differences in our statistical analyses , by treating each participant as a random factor (see Analysis below). Nonetheless, it is likely that old adults and young adults in our sample may still have differed in their overall colour preference, which might have affected colour judgements in the Dots Task. As such, our results should be interpreted with this information in mind.

The Dots Task. In the Dots Task, participants had to judge the colour of an individual dot presented on the screen (see Figure 2A). The task began with a series of instruction screens explaining the task structure, duration, and response format to the participant (see instructions in Appendix B). These instructions were followed by a practice block consisting of 10 trials, in which participants became familiar with the task. These trials were identical to trials in the real task and consisted of 50% purple dots and 50% blue dots. Data from practice trials were not analysed.

After the practice block, participants performed 800 test trials that were divided into 16 blocks of 50 trials each. In the decreasing prevalence condition, the number of blue dots in the environment decreased as the number of blocks increased in a predetermined fashion (based on Levari et al., 2018; see Figure 2A). In the stable prevalence condition, the proportion of blue dots in the environment did not change; it was always .50. In both cases, blue dots were defined as any dot for which the RGB value was between [0, 0, 254] and [49, 0, 205]. Purple dots were defined as any dot for which the RGB value was between [50, 0, 204] and [99, 0, 155]. Dot colours were randomly drawn without replacement for each trial based on the number of trials per block (50) and the frequency with which blue and purple dots should appear on a given block (always .50 in the stable prevalence condition and varying in the decreasing prevalence condition).

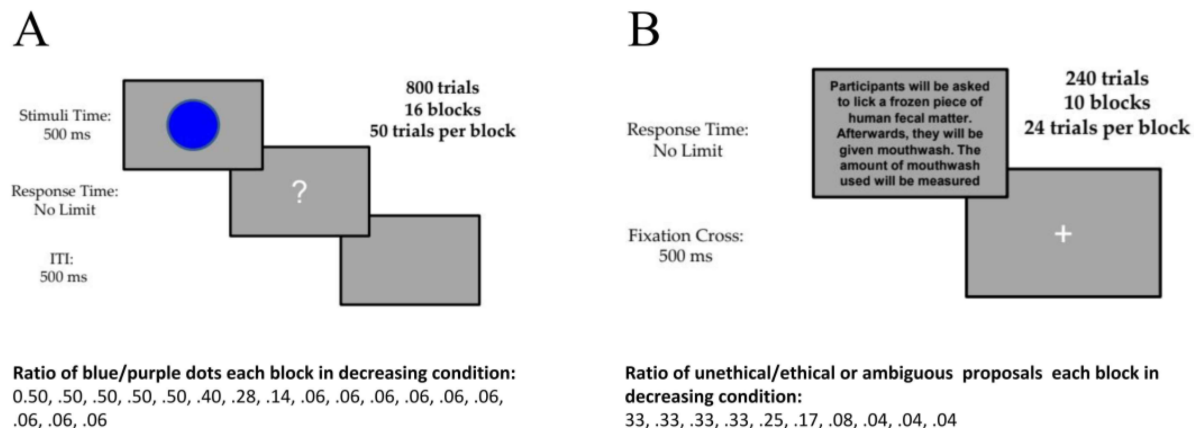

Figure 2. Task design for (A) the Dots Task and (B) the Ethics Task.

As can be seen in Figure 2A, on each trial, participants judged just one of these dots as being either blue or purple by pressing the ‘A’ or ‘L’ key on the keyboard. All stimuli were presented against a dark grey background. Each trial went as follows: A dot was presented on the screen for 500 ms, a question mark appeared on the screen until participants made a choice, and a blank screen appeared for 500 ms. Thus, the timing was fixed across participants, except that which would arise from differences in response times. Between each block, text appeared that indicated that the block was finished, which block the participant was now at, and offering them a short break should they choose to take one.

The Ethics Task. In the Ethics Task, participants had to take on the role of a member of an Ethics Review Board and judge whether fictitious research proposals were ethical or not (phrased as whether they would allow these research studies to be conducted or not; see Figure 2B). All research proposals were norm-tested by Levari et al. (2018, see Supporting Online Material) to produce scores depicting how ethical people found the 273 proposals.

These scores were used to bin proposals as unethical (80 proposals), ethical (113 proposals), or ambiguous (80 proposals). These bins were used to calculate the proportion of proposals that appeared in each block (including the practice trial).

As in the Dots Task, participants were first presented with instruction screens explaining the task to them (see Appendix C). Following the instructions, participants completed a practice trial in which they judged a research proposal using the keyboard keys. In this task, they pressed ‘A’ when they would not allow a study to be conducted and ‘L’ when they would.

All proposals in the experiment were presented in black text against a dark grey background. The test phase consisted of 240 trials broken into 10 blocks. In the decreasing prevalence condition, the proportion of unethical, ethical, and ambiguous proposals varied across blocks (see Figure 2B). In the stable prevalence condition, the proportion between the three types of proposals was the same throughout the task: .33.

Each trial, participants read a proposal and pressed ‘A’ or ‘L’ on the keyboard indicating whether they thought that the research should be allowed to take place or not. There was no time limit on this choice. Following the choice, a fixation cross appeared on the screen for 500 ms, followed by the next proposal. Between each block, text appeared that indicated that the block was finished, which block the participant was now at, and offering them a short break should they choose to take one.

Both the Dots and Ethics Tasks described above were taken from Levari et al. (2018). Both tasks were programmed in Python using the PsychoPy libraries. Task code is available upon request.

Procedure

Participants were recruited from the community or from Concordia's participation pool (see Appendix D for examples of recruitment materials). Participants were contacted by telephone or email and were asked basic demographic information to determine initial eligibility. If eligible at this stage, they were invited for a single two-hour session in the lab.

Once at the lab, participants asked to fill out a consent form (see Appendix E). After providing written consent, they completed the HRR colour vision test. If they failed the colour vision test, we still asked them to complete the Dots and Ethics tasks, but we excluded their data after the fact. After completing the colour vision test, they were asked to complete the Dots Task and Ethics Task, back-to-back. The order of these tasks was counterbalanced across participants. They were told that they would be free to take short breaks during the tasks (between blocks) and a longer break between the tasks, should they choose to. After completing both tasks, participants were debriefed and paid $20 for participating or were given their participation credits.

Analysis

All data were analysed in R (version 3.6.1). Analysis scripts are available upon request. For the Ethics task specifically, normed scores were reversed, to be able to plot them in the same direction as in the Dots Task, such that lower normed scores represent more ethical scenarios. Figures were created using the *ggplot2* library (Wickham, 2016). The main statistical analyses consisted of six general binomial mixed-effects models using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015). Models in each task predicted response using age group (between-subjects; young adult or older adult), condition (between-subjects; stable prevalence or decreasing prevalence), trial number (within-subjects), and stimulus strength (within-subjects; colour in the Dots Task and normed ethicality scores in the Ethics Task) as fixed effects, a random slope of trial, and a random intercept for each participant. We considered all main effects and interactions. We then ran two follow-up models in both tasks using the same predictors, split by age group. In all models, trial and stimulus strength were converted to a scale between 0 and 1.

Finally, we were interested in exploring how response times varied across age groups (young vs. old) and condition (stable vs. decreasing). In line with best practices regarding HARKing (Hollenback & Wright, 2017), we wish to disclose that these analyses were exploratory and were not based on our original hypotheses. Rather they were motivated by two factors: First, by a desire to provide a potential explanation of the observed age differences in sensitivity to prevalence-induced concept change; Second, in hopes of supplementing the prevalence-induced concept change literature by exploring the effect of condition on response times. Accordingly, we analysed response times using two between-groups 2x2 (age group × condition) ANOVA, one for each task. Further work will be needed to replicate and further interpret the observed response time differences (see the Discussion for suggestions of future work).

Results

Choice Data

In both tasks, prevalence-induced concept change is reflected as a three-way interaction between condition, trial, and stimulus strength, predicting responses. That is, the effect size of this interaction reflects the degree to which a participant's choice to categorise a given exemplar (dot or research proposal) as one concept or another is influenced by a combined effect of three factors. First, the changes in the relative frequency of a concept in the environment (i.e., the effect of condition). Second, the number of exemplars that the participant has been exposed to (i.e., the effect of trial). Third, the strength of the stimulus (i.e., the effect of blueness or ethicality). Thus, if older and younger adults differ in their sensitivity to prevalence-induced concept change, we would expect to see a four-way interaction between these three terms above and age group (dichotomized as young adult or older adult) and different effect sizes for this effect within each of the age groups.

Indeed, this is exactly what we observe. Results from the mixed-effects regressions are represented in Figure 3. In both tasks, there was a significant four-way interaction between age group, condition, trial, and stimulus strength (In the Dots Task: $\beta = 8.49$, SE = 0.37, $p < .0001$, 95% CI = [6.93, 10.00]; In the Ethics Task: $\beta = 0.90$, SE = 0.25, $p = .0004$, 95% CI = [0.40, 1.41]). We followed up on these regression analyses with two mixed-effects regressions separately for the two age groups, using the same predictors as above in both tasks (except for age group of course). This revealed that the effect of prevalence-induced concept change—again represented here as an interaction between condition, trial, and stimulus strength— was much stronger in younger adults ($\beta = 25.74$, SE = 0.90, $p < .0001$, 95% CI = [25.00, 26.40]) than older adults ($\beta = 17.44$, SE = 0.30, $p < .0001$, 95% CI = [16.90, 18.00]) in the Dots Task and was only statistically significant for younger adults in the Ethics Task ($\beta_{\text{Young Adults}} = 1.19$, SE = 0.22, $p < .0001$, 95% CI = [0.76, 1.63]; $\beta_{\text{Older Adults}} = 0.21$, SE = 0.14, $p = .1324$, 95% CI = [-0.06, 0.49]). Full regression tables are available in Appendix F.

Given the complexity of the interactions represented in Figure 3, interpreting the standardized regression weights as an effect size—that is, the logged odds ratio of a mixed-effects, three- or four-way, interaction—would prove difficult and potentially uninformative. Rather, to illustrate how prevalence induced-concept change affects judgements across age groups, here we will briefly describe a section of our data in detail. Take for instance judgements in the decreasing prevalence condition for a dot that was 33% blue. In the first 200 trials, 19% of young adults and 30% of older adults considered this dot to be blue. In the last 200 trials however, 73% of young adults now considered the dot to be blue, whereas only 58% of older adults considered the dot to be blue. Similar results were found for the Ethics Task. For a research proposal that had a normed rating of about 33% ethical, 33% of young adults and 44% of older adults stated that they would not allow this study to take place in the first 48 trials. In the last 48 trials, 50% of younger adults now would allow this study to take place. This is in contrast to only 42% of older adults in the last 48 trials who would allow the study to take place; a small decrease from the first 48 trials.

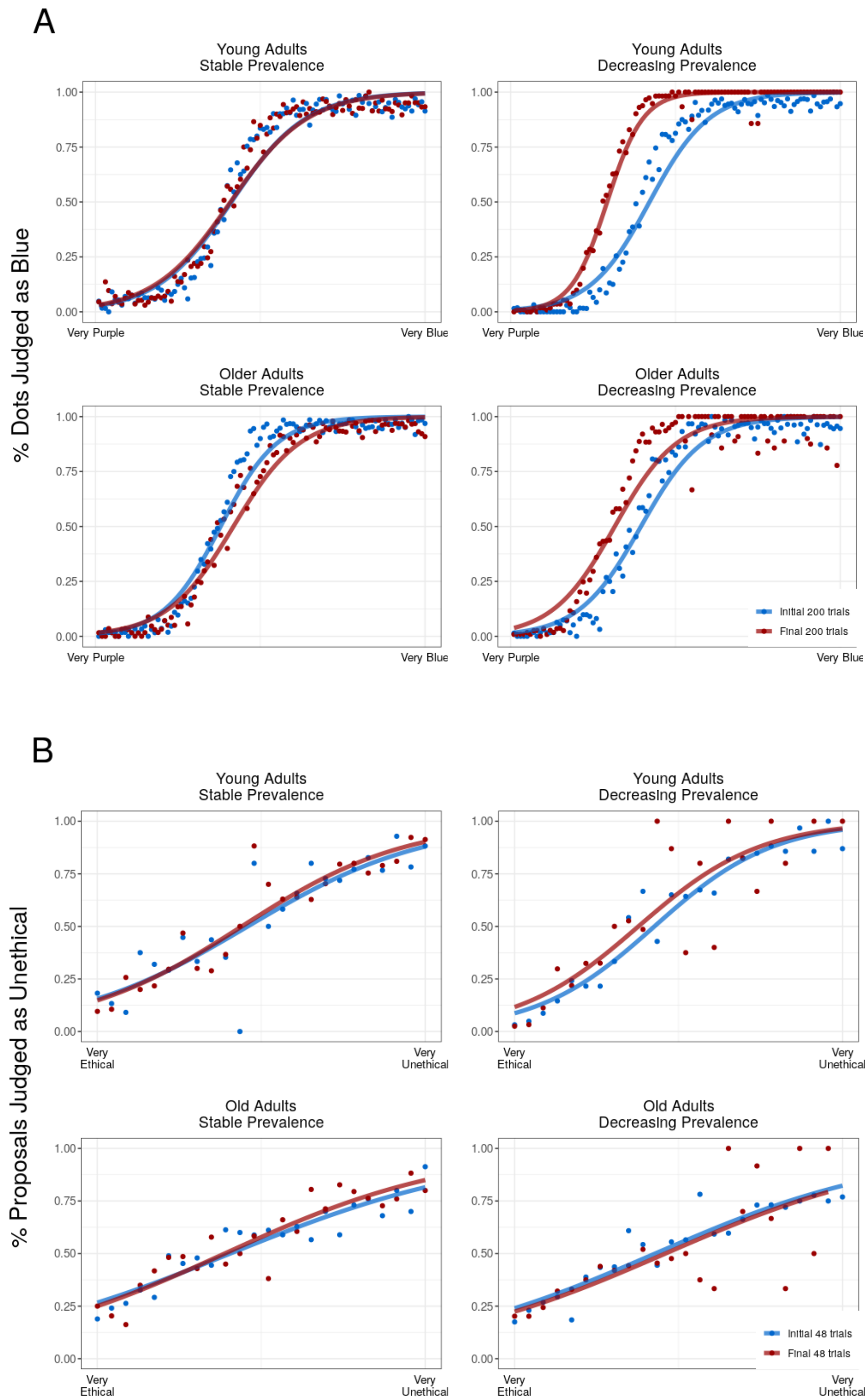

Figure 3. Concept judgements in (A) the Dots Task and (B) the Ethics Task. In the Dots Task, the y-axis represents the percent of dots judged as blue. In the Ethics Task, it represents the percent

of proposals judged as unethical. The x-axis represents stimulus strength: blueness in the Dots Task and ethicality in the Ethics task. Points represent the percent of choices for the corresponding stimulus strength, averaged across subjects within that cell (e.g., the percent of dots judged as blue by young adults in the stable prevalence condition when the dot present on the screen was very blue). Curves represent fitted binomial regression curves. Blue points and lines represent the first 200 trials in the Dots Task and first 48 trials in the Ethics Task. Red ones represent the final 200 trials in the Dots Task and final 48 trials in the Ethics Task. The first row in both groups of plots represents choice data from young adults and the second row represents choice data from older adults.

These, admittedly anecdotal, examples provide an expression of the size of the effects in our regression results. That is, they demonstrate that when the prevalence of exemplars in the environment decreases, both younger and older adults' concepts expanded to include exemplars they previously did not, but that this phenomenon occurred to differing degrees depending on the participants' age.

Response Times

Response time data across age groups is presented in Figure 2. Two 2x2 ANOVA (age group x condition) were conducted on each subject's mean response time data (see the Analysis section above for our rationale).

These analyses revealed a significant main effect of age group on response time in both tasks (Dots Task: $F(1, 116)$: 51.05, $p < .0001$, 95% CI = [0.17, 0.34], $\text{difference}_{\text{Older - Young}} = 0.21$ seconds; Ethics Task: $F(1, 116)$: 23.47, $p < .0001$, 95% CI = [1.23, 4.04], $\text{difference}_{\text{Older - Young}} =$

2.43 seconds), but they did not indicate a statistically significant main effect of condition or interaction between age group or condition (see Table 1). Thus, older adults responded more slowly than younger adults overall, irrespective of condition, in both tasks.

Figure 4. Pirate plot of response times in both age groups across both tasks. X-axis is the age group. Y-axis is the mean response time in seconds per participant. Each point represents an individual participant's mean response time. Boxes represent 95% confidence intervals and horizontal lines represent group means. Pink represents younger adults and blue represents older adults.

Table 1

ANOVA Table for response time In Dots and Ethics Task

Source	**SS**	***df***	***F***	***p***	***95% CI***
			Dots Task		
Age Group	1.38	1	51.05	< .0001	[0.17, 0.34]
Condition	0.01	1	0.34	.5633	[-0.06, 0.10]
Age Group * Condition	0.04	1	1.63	.204	[-0.20, 0.04]
			Ethics Task		
Age Group	177.16	1	23.47	< .0001	[1.23, 4.04]
Condition	0.97	1	0.12	.7208	[-1.38, 1.43]
Age Group * Condition	1.20	1	0.15	.6904	[-2.39, 1.59]

Discussion

The purpose of this thesis was to investigate how prevalence-induced concept change differentially affected the judgements of older adults across two conceptual domains: perception and ethics. We hypothesized that older adults would either be less sensitive (H1) or more sensitive (H2) to prevalence-induced concept change than younger adults. Our results support the first hypothesis, demonstrating that older adults were **less** sensitive to prevalence-induced concept change in their judgements about the colours of dots and were not significantly affected by the phenomenon in their ethical judgements about fictitious research proposals.

These results dovetail nicely with a body of research demonstrating that older adults have greater difficulty than younger adults abandoning past behaviours in favour of new behaviours despite changes in the environment (Eppinger, Hämmerer, & Li, 2011). In line with this view, as Wilson's (2018) computational model (described in the introduction) highlights, prevalence-induced concept change could also be thought of as a form of implicit learning, where the underlying statistic of a stimuli (e.g., the average blueness or ethicality) is implicitly estimated based on recently seen exemplars (cf. Cleeremans, Destrebecqz, & Maud Boyer, 1998). From this perspective, older adults may have more difficulty learning these latent states of stimuli and default to their original responses (Howard & Howard, 2013). As Nassar and colleagues (2016) suggest, this difficulty might in part be due to a reduced sensitivity to uncertainty in the environment. Furthermore, this reduced sensitivity might result from a failure to update learning without an explicit environmental cue (e.g., a surprise). Indeed, our results are consistent with this view. That is, insofar as prevalence-induced concept change can be thought of as an implicit learning process (Wilson, 2018), so too can we say that a reduced sensitivity to prevalence-induced concept change might result from a failure to learn from an uncertain and subtly changing environment (Nassar et al., 2016).

In the tasks used in this study, this same failure may result in perseverative behaviour (repeating the same responses; Burckner et al., in prep.), which would in turn drive prevalence-induced concept change down, just as we see in our behavioural data. To examine whether perseveration is a driving mechanism behind the age differences we observed in this study, Wilson's (2018) computational model could be applied to the current data. In this model, two key parameters of interest are estimated: the effect of past stimulus and the effect of previous response on current response. Interestingly, these parameters act as opponent processes, such that a greater effect of past stimuli increases prevalence-induced concept, whereas a greater effect of past response reduces prevalence-induced concept change. Thus, based on past literature and the current behavioural results, we would expect that responses from older adults in our sample were, on average, driven more by past response than younger adults (i.e., older adults perseverate more). Indeed, this interpretation would line up with neurocognitive work, suggesting that deficient dopaminergic modulation of the prefrontal cortex's attention regulation mechanisms leads to less distinctive mental representations among older adults (Li, Lindenberger, & Sikström, 2001). This would further imply that as people age, their representations of past stimuli become weaker and less specific. Thus, due to this dysregulation of dopamine pathways, older adults may rely less on their (impoverished) representation of past stimuli and instead rely more readily on their previous responses (i.e., engage in perseverative behaviour; Eppinger, et al., 2013; de Boer et al., 2017).

If perseveration is the(/a) mechanism underlying older adults' reduced sensitivity to the phenomenon however, it would be rather interesting that a feature of healthy ageing generally regarded as maladaptive—a difficulty in adapting behaviour to changing and uncertain conditions; deficient dopamine modulation—would in this case be protective against some of the more problematic effects of prevalence-induced concept change in the real-world. Take, for instance, some ethical judgements, where one might argue that judgements ought to remain relatively stable over time.

However, there is an alternative explanation as to why we observe differences between younger and older adults in the current sample that needs to be considered: Beyond just perseveration, older adults' longer response times themselves might also have contributed—or been the result of some other process that contributed—to reduced prevalence-induced concept change. At the end of his paper, Wilson (2018) briefly remarked that the smallest behavioural effects in Levari's et al. (2018) original data were observed in the Ethics Task, which also had the longest response times. These differences in response time inadvertently increased the amount of time between stimuli across the tasks (approximately 850 ms between dots in Study 1-5 and 5 s between research proposals in Study 7). Our results replicate this finding in the young adults, such that young adults showed reduced prevalence-induced concept change in the Ethics Task compared to the Dots Task. Furthermore, we also found an even smaller effect of prevalence-induced concept change (or complete lack of one in the Ethics Task) in older adults, who also showed even longer response times across both tasks (see Figure 3). Thus, the question arises: Are the effects of prevalence-induced concept change observed in these tasks affected by the speed at which responses are made?

The observed differences in response times between younger and older adults can be explained by two different (but not mutually exclusive) hypotheses, which in turn might elucidate the mechanisms underlying the differences in sensitivity to prevalence-induced concept change. First, older adults in our sample might be exhibiting general slowing, a well-known cognitive phenomenon in healthy ageing whereby peoples' response times slow with age (Verhaeghen & Cerella, 2002). If this were the case, slower responses among older adults in our sample would simply be demonstrating a basic feature of healthy ageing. Furthermore, if differences in sensitivity to prevalence-induced concept change are in part due to general slowing, it might suggest that the paradigms we used are very sensitive to differences in timing between the stimuli. If so, it would be important to re-evaluate to what degree prevalence-induced concept change is task-dependent and if there were ways to measure it independently of response time.

However, a second explanation as to why older adults differ in terms of response time from younger adults might be that they engage in a speed-accuracy trade-off (Starns & Ratcliffe, 2012; Salthouse, 1979). From this perspective, it would not necessarily be the case that older adults are limited in their ability to respond quickly per se, but rather prioritise accuracy over response speed. Thus, in the context of our tasks, older adults would spend more time judging each exemplar to attempt to maximise "accuracy" (or perhaps something like internal consistency in the case of the Ethics Task). Were this the case, it would suggest that the effects of prevalence-induced concept change might be resisted if one put effort into slow, deliberate, and accurate response (cf. Kahneman's (2011) System 2 thinking). Furthermore, it would open up the possibility that the effects elicited by current tasks used to measure prevalence-induced concept change are not necessarily sensitive to increased response time in and of itself (i.e., greater distance between stimuli) but rather to different decision-making strategies altogether, such as those that trade speed for accuracy.

Exploring which of these interpretations are borne out by future data is important not only for elucidating how prevalence-induced concept change differentially affects younger and older adults' judgements, but it is also crucial for better understanding how it plays out in the real-world. This is because real-world judgements are often made on "stimuli" that are hours, days, or months apart in time and, even when judgements are made on "stimuli" that are close in time, people may have motivation to weigh their choices carefully and deliberately (e.g., in political discussions). As such, it is important to verify how robust the observed lab effects are to differences in timing between stimuli and response strategies should we wish to generalise from the lab to real-world decision-making (cf. Yarkoni, 2019).

Furthermore, this future research has substantial implications for prevalence-induced concept change as a phenomenon more broadly. Levari's et al. (2018) initial interpretation of their results was that prevalence-induced concept change, as the name suggests, affects *concepts themselves*. While the authors provide this phenomenon with a new name, prevalence-induced concept change can be accounted for by an older theory in cognitive psychology: the prototype theory of concepts (e.g., Homa, Rhoads, & Chambliss, 1979; Rosch & Mervis, 1975). According to this theory stimuli are categorised by referring to an abstracted mental prototype of a concept (e.g., a labrador may stand in for a prototype to the concept "dog"). These prototypes are then themselves susceptible to update from environmental information, such as, in this case, the prevalence (Smith & Minda, 1998; Minda & Smith, 2001). Prototype theory, however, runs into issues, particularly in how it is usually supported empirically; that is through categorisation tasks. As Magolis (1994) points out, it is very difficult to determine if such tasks truly tap into people's concepts (their representations) or merely describe how they make judgements about concepts (their behaviour). In the case of prevalence-induced concept change, if it is the latter, then it is dubious to call the phenomenon "concept" change, when really what is changing as the prevalence of exemplars decrease is the criterion people use for categorising certain stimuli. For instance, past work using similar linguistic binary categorisation tasks—categorising stimuli as a "word" or "non-word"—has revealed that when the proportion of exemplars is asymmetrical (more non-words than words), error rates increase (Wagenmakers et al., 2008). From the interpretation Levari et al. (2018) offer, this would suggest that the "concept" of "word" has changed due to this asymmetry, since participants become more likely to judge words as non-words (i.e., they are making more errors). This, of course, would be a nonsensical explanation of these data. For similar reasons, Levari's et al. (2018) original interpretation that concepts themselves change alongside the prevalence of exemplars may not be warranted. In other words, with current methods for measuring prevalence-induced concept change in the lab, it is difficult to distinguish between concept change and task-dependent response bias.

However, another, albeit less grandiose, interpretation of the current findings (and Levari's et al. (2018) findings for that matter) is that concepts remain stable, but *judgements* about concepts change as the prevalence of exemplars of a concept shifts. That is, our concept of blue might remain stable despite changes in the environment, but whether and how we decide to apply that concept to a stimulus might change. If this were the case, it is reasonable to assume that decision-making strategies like a speed-accuracy trade-off may affect such decisions.

However, it is less intuitive to suggest that—presumably deliberate—strategies like this affect the content of our basic conceptual apparatuses, as Levari's et al. (2018) interpretation could suggest. Moreover, this interpretation may allow for a lower-level understanding of prevalence-induced concept change, as a cognitive phenomenon that affects humans' criteria for making judgements rather than one that changes higher-order concepts themselves. Therefore, elucidating whether prevalence-induced concept change is affected by explicit decision-making strategies has not only important implications for understanding how the phenomenon affects older adults, but also for properly understanding what is meant by prevalence-induced concept change to begin with.

In summary, the current results suggest that older adults are less sensitive to prevalence-induced concept change than younger adults. The potential real-world implication of these findings are context-dependent, such that in some cases it can be adaptive for one's judgments to be sensitive to a changing world, however in others it can be harmful. For instance, when fresh fruits are not available in supermarkets, it is adaptive for customers to adjust their judgements of "fresh" to include foods that they might otherwise exclude. However, in moral and social decisions, it may be beneficial to hold our judgements to a higher ideal rather than shift them alongside changing trends. As such, the fact that older adults demonstrate less sensitivity to prevalence-induced concept change can be a benefit in some situations and more problematic in others.

To further explore how prevalence-induced concept change affects judgements, both in old age and in general, we present three avenues for future research: (a) the implementation of Wilson's (2018) computational model to better understand how age differences in sensitivity to prevalence-induced concept change arise as a result of sequential choice effects such as perseveration; (b) further exploration of the degree to which prevalence-induced concept change is sensitive to the timing between stimuli; and (c) investigating whether longer response times among older adults in these tasks are due to general slowing, a speed-accuracy trade-off, or a combination of both.

Though much more work needs to be done, the current study points to the fact that as we age, how we make judgements from our concepts might become more rigid as we face a changing world. This notion is important when considering the degree to which older adults' use of concepts, such as right, fair, free, and so on will come to affect the future direction of our society. Indeed, as we age, it seems that our judgements about concepts remain more stable, even if the world around us presents us with continued reason to change them. It is in this sense that the quote at the beginning of this thesis earns its relevance: The more things (our age and our environment) change, the more they (our judgements) stay the same.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bruckner, R., Nassar, M., Li, S.C., & Eppinger, B. (in prep.). *Normative and heuristic determinants of learning across the lifespan*.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*(10), 406–416. https://doi.org/10.1016/S1364-6613(98)01232-7

Cole, B. L., Lian, K. Y., & Lakkis, C. (2006). The new Richmond HRR pseudoisochromatic test for colour vision is better than the ishihara test. *Clinical and Experimental Optometry*, *89*(2), 73–80. https://doi.org/10.1111/j.1444-0938.2006.00015.x

de Boer, L., Axelsson, J., Riklund, K., Nyberg, L., Dayan, P., Bäckman, L., & Guitart-Masip, M. (2017). Attenuation of dopamine-modulated prefrontal value signals underlies probabilistic reward learning deficits in old age. *Elife, 6*, e26424. 10.7554/eLife.26424

Dittmar, M. (2001). Changing colour preferences with ageing: A comparative study on younger and older native Germans Aged 19–90 Years. *Gerontology, 47*(4), 219–226. 10.1159/000052802

Eppinger, B., Hämmerer, D., & Li, S. C. (2011). Neuromodulation of reward-based learning and decision making in human aging. *Annals of the New York Academy of Sciences, 1235*, 1. https://doi.org/10.1111/j.1749-6632.2011.06230.x.Neuromodulation

Eppinger, B., Heekeren, H. R., & Li, S. C. (2015). Age-related prefrontal impairments implicate deficient prediction of future reward in older adults. *Neurobiology of Aging*, *36*(8), 2380–2390. https://doi.org/10.1016/j.neurobiolaging.2015.04.010

Eppinger, B., Nystrom, L. E., & Cohen, J. D. (2012). Reduced sensitivity to immediate reward during decision-making in older than younger adults. *PLoS ONE*, *7*(5), 1–10. https://doi.org/10.1371/journal.pone.0036953

Eppinger, B., Walter, M., Heekeren, H. R., & Li, S. C. (2013). Of goals and habits: Age-related and individual differences in goal-directed decision-making. *Frontiers in Neuroscience*, *7*, 1–14. https://doi.org/10.3389/fnins.2013.00253

Fiorentini, A., Porciatti, V., Morrone, M. C., & Burr, D. C. (1996). Visual Ageing: Unspecific Decline of the Responses to Luminance and Colour. *Vision Research, 36*(21), 3557–3566. doi:10.1016/0042-6989(96)00032-6

Hämmerer, D., Müller, V., & Li, S. C. (2014). Performance monitoring across the lifespan: Still maturing post-conflict regulation in children and declining task-set monitoring in older adults. *Neuroscience and Biobehavioral Reviews*, *46*(P1), 105–123. https://doi.org/10.1016/j.neubiorev.2014.06.008

Hämmerer, D., Schwartenbeck, P., Gallagher, M., FitzGerald, T. H. B., Düzel, E., & Dolan, R. J. (2019). Older adults fail to form stable task representations during model-based reversal inference. *Neurobiology of Aging*, *74*, 90–100. https://doi.org/10.1016/j.neurobiolaging.2018.10.009

Hollenbeck, J. R., & Wright, P. M. (2017). Harking, sharking, and tharking: making the case for post hoc analysis of scientific data. *Journal of Management, 43*(1), 5–18. https://doi.org/10.1177/0149206316679487

Homa, D., Rhoads, D., & Chambliss, D. (1979). Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory, 5*(1), 11. https://doi.org/10.1037/0278-7393.5.1.11

Howard, J. H., & Howard, D. V. (2013). Aging mind and brain: Is implicit learning spared in healthy aging? *Frontiers in Psychology*, *4*, 1–6. https://doi.org/10.3389/fpsyg.2013.00817

Kahneman, D. (2011). *Thinking Fast and Slow*. Macmillan.

Kant, I. (1998). *Critique of Pure Reason*. Cambridge: Cambridge University Press.

Kerchner, G. A., Racine, C. A., Hale, S., Wilheim, R., Laluz, V., Miller, B. L., & Kramer, J. H. (2012). Cognitive Processing Speed in Older Adults: Relationship with White Matter Integrity. *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0050425

Levari, D. E., Gilbert, D. T., Wilson, T. D., Sievers, B., Amodio, D. M., & Wheatley, T. (2018). Prevalence-induced concept change in human judgment. *Science*, *360*(6396), 1465–1467. https://doi.org/10.1126/science.aap8731

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological Assessment* (5th ed.). Oxford University Press.

Li, S. C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: from neuromodulation to representation. Trends in cognitive sciences, 5(11), 479-486. https://doi.org/10.1016/S1364-6613(00)01769-1

Lindenberger, U., & Mayr, U. (2014). Cognitive aging: is there a dark side to environmental support? *Trends in cognitive sciences, 18*(1), 7-15. https://doi.org/10.1016/j.tics.2013.10.006.Cognitive

Margolis, E. (1994). A reassessment of the shift from the classical theory of concepts to prototype theory. *Cognition, 51*(1), 73-89. https://doi.org/10.1016/0010-0277(94)90009-4

Mayr, U., Spieler, D. H., & Hutcheon, T. G. (2015). When and why do old adults outsource control to the environment? *Psychology and Aging*, *30*(3), 624–633. https://doi.org/10.1037/a0039466

Mayr, U., Spieler, D. H., & Kliegl, R. (2001). *Ageing and Executive Control*. Psychology Press.

Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: the effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(3), 775. https://doi.org/10.1037/0278-7393.27.3.775

Murman, D. L. (2015). The Impact of Age on Cognition. *Seminars in Hearing*, *36*(3), 111–121. https://doi.org/10.1055/s-0035-1555115

Nassar, M. R., Bruckner, R., Gold, J. I., Li, S. C., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, *7*(May 2015). https://doi.org/10.1038/ncomms11609

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive psychology, 7*(4), 573-605. https://doi.org/10.1016/0010-0285(75)90024-9

Salthouse, T. A. (1979). Adult age and the speed-accuracy trade-off. *Ergonomics*, *22*(7), 811–821. https://doi.org/10.1080/00140137908924659

Samanez-Larkin, G. R., & Knutson, B. (2017). Decision making in the ageing brain: Changes in affective and motivational circuits. *Nat Rev Neuroscience*, *176*(1), 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6), 1411. https://doi.org/10.1037/0278-7393.24.6.1411

Sparrow, E. P., & Spaniol, J. (2016). Age-Related Changes in Decision Making. *Current Behavioral Neuroscience Reports*, *3*(4), 285–292. https://doi.org/10.1007/s40473-016-0091-4

Spieler, D. H., Mayr, U., & LaGrone, S. (2006). Outsourcing cognitive control to the environment: Adult age differences in the use of task cues. *Psychonomic Bulletin and Review*, *13*(5), 787–793. https://doi.org/10.3758/BF03193998

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*(2), 377–390. https://doi.org/10.1037/a0018022

Statistics Canada. (2019). *What will the population of Canada look like in 2068?* Retrieved December 2, 2019, from Statistics Canada website: https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2019050-eng.htm

Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: a review of meta-analyses. *Neuroscience and Biobehavioral Reviews*, *26*, 1–9. https://doi.org/10.1016/S0149-7634(02)00071-4

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of memory and language, 58*(1), 140-159. https://doi.org/10.1016/j.jml.2007.04.006

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wilson, R. C. (2018). *Sequential choice effects predict prevalence - induced concept change*. Retrieved from: https://psyarxiv.com/75bpy/

Yarkoni, T. (2019). *The Generalizability Crisis*. Retrieved from: https://psyarxiv.com/jqw35

Appendix A

Example of HRR Colour Vision Test Plate

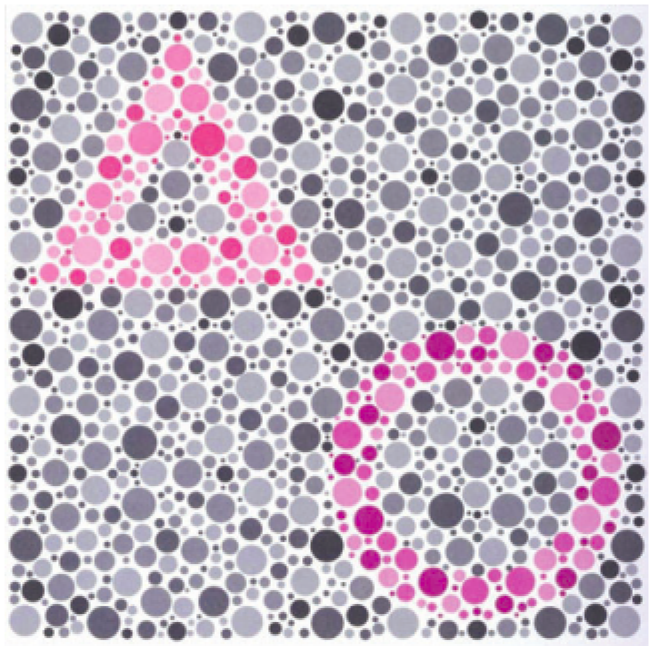

Figure 1. Example of a colour plate in the HRR Colour Vision Test. Taken from Cole, Lian, and Lakkis (2006).

Appendix B

Instructions for the Dots Task

Welcome to this study! We’re interested in studying how people perceive and identify colors.

In this task, you will see dots presented on the screen one at a time, in a variety of colors. Your task in this study will be to identify blue dots.

When you see a blue dot on the screen, press the “A” key. For all other dots, press the “L” key.

The dots will be presented in series with breaks in between. This means that you will see a series of dots, have a short break, and then another series of dots, until you have seen 20 series.

Some of the series you see may have a lot of blue dots, and others may have only a few. There’s nothing for you to count or keep track of -- your only task is to identify blue dots.

You should do your best to answer quickly and accurately during the study. However, if you make a mistake and hit the wrong button at any point, just keep going.

Now you will complete a brief practice series so you can get used to the task.

You have now completed the practice series. If you have any questions, you can ask the experimenter now.

Otherwise, you're ready to begin the study.

After each block:

Series complete.

Please take a short break. We’ll start the next series in a moment.

Appendix C

Instructions for the Ethics Task

Welcome to this study! We're interested in studying how people make ethical decisions about scientific experiments.

Many scientific experiments involve some risk for the participants because they can cause psychological distress or physical harm. Universities have to make difficult ethical decisions about whether or not to allow experiments to be conducted.

Today, you will read about various experiments that could be conducted on human beings. We simply want to know whether you think scientists SHOULD or SHOULD NOT be allowed to conduct each of these experiments.

Because this is an ethical decision, there are no right or wrong answers. We simply want your personal decision for each study.

Here are some things to keep in mind as you make your decisions.

1) All of the experiments you will read about will be conducted on adults who have volunteered to take part in exchange for money.

2) All of the experiments are part of research on human behavior.

3) When scientists must lie to the participants either before or during the experiment, they always tell the participants the truth when the experiment is over.

4) Participants are always free to withdraw and can stop participating at any time they wish.

In the task, you will see descriptions of experiments presented on the screen, one at a time.

When you read a description of an experiment that you would not allow to be conducted, press the "A" key. For all other experiments, press the "L" key.

The experiments will be presented in series, with breaks in between. This means that you will read a series of experiments, have a short break, and then another series of experiments, until you have seen 10 series.

Some of the series you see may have a lot of unethical experiments, and others may have only a few. There's nothing for you to count or keep track of -- your only task is to approve or reject each experiment.

You should do your best to answer quickly and accurately during the study. However, if you make a mistake and hit the wrong button at any point, just keep going.

Now you will complete a brief practice round so you can get used to the task.

You have now completed the practice round. If you have any questions, you can ask the experimenter now. Otherwise, you're ready to begin the study.

After each block:

Series complete.

Please take a short break. We'll start the next series in a moment.

Appendix D

Recruitment Materials

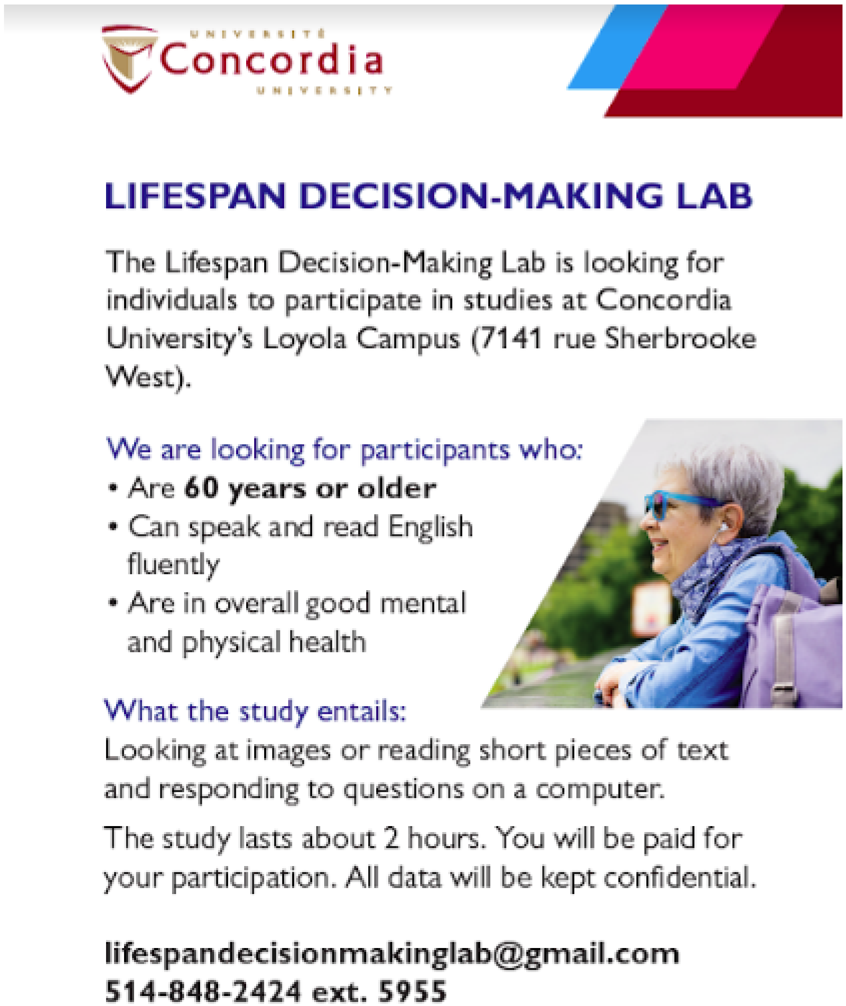

Figure 2. Example of a recruitment postcard distributed to older adults in the community.

Study Name	Categorization Study
Study Type	**Standard (lab) study** This is a standard lab study. To participate, sign up, and go to the specified location at the chosen time.
Credits	2 Credits
Duration	120 minutes
Sign-Up Restrictions	You must have signed up or completed ALL of these studies: Categorization Study: PREREQUISITE QUESTIONNAIRE
Abstract	This study involves completing various computer tasks to assess how young adults form categories.
Description	This study seeks to understand how younger adults make decisions about categories. It involves completing a short questionnaire at-home, a vision task, and a series of computer tasks. The at-home questionnaire asks some simple demographics question to ensure that you are eligible to participate in the study and that we can use your data. This questionnaire is available on Participation Pool as "Categorization Study: PREREQUISITE". You must complete it before being able to sign up for this study. The vision task is done when you arrive at the lab. It's short test to ensure that your colour vision is normal, as some of the computer tasks you will do require you to distinguish between colours. The computer tasks are short judgment and cognitive tasks. They are completed in our lab on a computer using a mouse and keyboard. The specifics of the tasks would be explained during your session, but they all involve making certain judgments and decisions about things you see on the screen. The whole study should take 1.5 - 2 hours and you will be given 2 credits upon completion.
Eligibility Requirements	Complete the "Categorization Study: PREREQUISITE" on Participation Pool
Researcher	Lifespan Decisionmaking Office: SP350.00 5148482424
Principal Investigator	Ben Eppinger
Deadlines	Deadlines that occur on a Saturday or Sunday will be moved back to Friday Sign-Up: 12 hour(s) before the appointment Cancellation: 12 hour(s) before the appointment

Figure 3. Participation pool page for this study, termed “Categorization Study”

Appendix E

Consent Form

INFORMATION AND CONSENT FORM

Study Title: Prevalence-induced Concept Change in Older Adults

Researcher(s): Dr. Ben Eppinger; Sean Devine

Researchers' Contact Information: Dept. of Psychology, Concordia University 7141 Sherbrooke St. Ouest, Montréal, H4B 1R6 (514-848-2424 x2397; ben.eppinger@concordia.ca).

Faculty Supervisor: Dr. Ben Eppinger

Faculty Supervisor's Contact Information: 514-848-2424 x2397; ben.eppinger@concordia.ca

Source of funding for the study: CIHR MC001-2017: CRC in cognitive neuroscience of decision-making in healthy human aging.

You are being invited to participate in the research study mentioned above. This form provides information about what participating would mean. Please read it carefully before deciding if you want to participate or not. If there is anything you do not understand, or if you want more information, please ask the researcher.

A. PURPOSE

The purpose of this study is to investigate how younger and older adults make decisions about categories.

B. PROCEDURES

If you participate, you will be asked to complete a demographic questionnaire, as well as several computerized tasks. The whole session should take about 2 hours and will require you to complete various tasks on a computer in the lab.

C. RISKS AND BENEFITS

There are no direct advantages for participation in this study, and this research is not intended to benefit you personally.

The principal disadvantage of participating in this study is the time it will take you to participate in the testing session. There is a slight risk that you might feel a loss of confidence in your ability to perform tests of cognitive skill, because of the perception that you did not perform as well as you expected. Importantly, each of the tests are calibrated to be at a suitable level of difficulty for individuals in your age group. If you wish to discuss any concerns, you may contact Dr. Ben Eppinger PhD, the research project leader (514-848-2424 ext. 2397), who can provide advice concerning follow-up consultation with your family physician.

D. CONFIDENTIALITY

We will gather the following information as part of this research:

- Demographic information (including information on your history, background, and any physical or cognitive problems you may have).
- Computer-based tasks aimed at examining your judgment about categories.

We will not allow anyone to access the information, except people directly involved in conducting the research. We will only use the information for the purposes of the research described in this form. The information gathered will be coded. That means that the information will be identified by a code. The researcher will have a list that links the code to your name. We will protect the information by storing it in a locked and secure location at Concordia University, and in the case of digital information, in a password protected file on a secure server. We intend to publish the results of the research. However, it will not be possible to identify you in the published results. We will destroy the information five years after the end of the study.

F. CONDITIONS OF PARTICIPATION

You do not have to participate in this research. It is purely your decision. If you do participate, you can stop at any time. You can also ask that the information you provided not be used, and your choice will be respected. If you decide that you don't want us to use your information, you must tell the researcher before leaving the testing session.

Participants will receive a compensation of $10.00 per hour. If you withdraw before the end of the research, you will receive $10.00.

To make sure that research money is being spent properly, auditors from Concordia or outside will have access to a coded list of participants. It will not be possible to identify you from this list.

Appendix F

Mixed-Effects Regression Tables

Table 1

Output from overall mixed-effects regression in The Dots Task

Source	**Estimate**	**SE**	***p-value***
Intercept	-5.72	0.15	< .0001
Age Group	1.74	0.22	< .0001
Condition	1.16	0.18	< .0001
Trial	-1.06	0.24	< .0001
Colour	15.15	0.15	< .0001
Age Group * Condition	-2.07	0.27	< .0001
Age Group * Trial	1.06	0.52	.0400
Age Group * Colour	-5.41	0.20	< .0001
Condition * Trial	-2.64	0.30	< .0001
Condition * Colour	-4.41	0.18	< .0001
Trial * Colour	0.53	0.19	.0060
Age Group * Condition * Trial	-2.30	0.71	.0012
Age Group * Condition * Colour	5.21	0.27	< .0001
Age Group * Trial * Colour	-0.50	0.29	.0832
Condition * Trial * Colour	17.43	0.22	< .0001
Age Group * Condition * Trial * Colour	8.48	0.37	< .0001

Table 2

Output from mixed-effects regression for young adults only, in the Dots Task

Source	**Estimate**	**SE**	***p-value***
Intercept	-3.98	0.18	< .0001
Condition	-0.91	0.25	.0002
Trial	0.001	0.25	.9937
Colour	9.74	0.20	< .0001
Condition * Trial	-4.89	0.34	< .0001
Condition * Colour	0.82	0.31	.0008
Trial * Colour	0.01	0.36	0.9678
Condition * Trial * Colour	25.74	0.90	< .0001

Table 3

Output from mixed-effects regression for older adults only, in the Dots Task

Source	**Estimate**	**SE**	***p-value***
Intercept	-5.72	0.19	< .0001
Condition	1.17	0.23	< .0001
Trial	-1.10	0.22	< .0001
Colour	15.15	0.26	< .0001
Condition * Trial	-2.63	0.26	< .0001
Condition * Colour	-4.42	0.34	< .0001
Trial * Colour	0.61	0.28	.0285
Condition * Trial * Colour	17.44	0.30	< .0001

Table 4

Output from overall mixed-effects regression in The Ethics Task

Source	**Estimate**	**SE**	***p-value***
Intercept	-2.61	0.24	< .0001
Age Group	-0.60	0.34	.0798
Condition	0.21	0.35	.5478
Trial	-0.87	0.47	.0617
Ethicality	0.75	0.05	< .0001
Age Group * Condition	-2.05	0.55	.0002
Age Group * Trial	-0.25	0.68	.7163
Age Group * Ethicality	0.20	0.07	.0043
Condition * Trial	-0.52	0.75	.4864
Condition * Ethicality	-0.03	0.07	.6470
Trial * Ethicality	0.18	0.08	.0308
Age Group * Condition * Trial	-4.69	1.23	.0001
Age Group * Condition * Ethicality	0.60	0.12	< .0001
Age Group * Trial * Ethicality	0.25	0.13	.0450
Condition * Trial * Ethicality	0.25	0.14	.0700
Age Group * Condition * Trial * Ethicality	0.90	0.25	.0004

Table 5

Output from mixed-effects regression for younger adults only, in the Ethics Task

Source	**Estimate**	**SE**	***p-value***
Intercept	-3.28	0.27	< .0001
Condition	-1.79	0.45	< .0001
Trial	-1.02	0.59	.0835
Ethicality	0.96	0.05	< .0001
Condition * Trial	-5.53	1.10	< .0001
Condition * Ethicality	0.56	0.10	< .0001
Trial * Ethicality	0.43	0.10	< .0001
Condition * Trial * Ethicality	1.19	0.22	< .0001

Table 6

Output from mixed-effects regression for older adults only, in the Ethics Task

Source	**Estimate**	**SE**	***p-value***
Intercept	-2.58	0.23	< .0001
Condition	0.14	0.33	.6624
Trial	-0.93	0.40	.0221
Ethicality	0.74	0.05	< .0001
Condition * Trial	-0.37	0.68	.5859
Condition * Ethicality	-0.01	0.07	.8406
Trial * Ethicality	0.19	0.08	.0197
Condition * Trial * Ethicality	0.21	0.14	0.1324