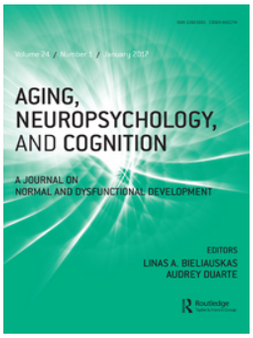

Routledge
Taylor & Francis Group

Aging, Neuropsychology, and Cognition
A Journal on Normal and Dysfunctional Development

ISSN: 1382-5585 (Print) 1744-4128 (Online) Journal homepage: http://www.tandfonline.com/loi/nanc20

Positive explorers: modeling dynamic control in normal aging

Brian D. Glass & Magda Osman

To cite this article: Brian D. Glass & Magda Osman (2017) Positive explorers: modeling dynamic control in normal aging, Aging, Neuropsychology, and Cognition, 24:1, 62-79, DOI: 10.1080/13825585.2016.1171290

To link to this article: https://doi.org/10.1080/13825585.2016.1171290

Published online: 04 May 2016.

AGING, NEUROPSYCHOLOGY, AND COGNITION, 2017
VOL. 24, NO. 1, 62–79
http://dx.doi.org/10.1080/13825585.2016.1171290

Routledge
Taylor & Francis Group

Positive explorers: modeling dynamic control in normal aging

Brian D. Glass[a] and Magda Osman[b]

[a]Department of Computer Science, University College London, University of London, London, United Kingdom; [b]Biological and Experimental Psychology, School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom

ABSTRACT
Situations in which there are multiple changes occurring all at once and which demand complex decisions to be made are common throughout life, but little is known about how normal aging influences performance on these types of scenarios. To determine performance differences associated with normal aging, we test older and younger adults in a dynamic control task. The task involves the control of a single output variable over time via multiple and uncertain input controls. The Single Limited Input, Dynamic Exploratory Responses (SLIDER) computational model, is implemented to determine the behavioral characteristics associated with normal aging in a dynamic control task. Model-based analysis demonstrates a unique performance signature profile associated with normal aging. Specifically, older adults exhibit a positivity effect in which they are more influenced by positively valenced feedback, congruent with previous research, as well as enhanced exploratory behavior.

ARTICLE HISTORY
Received 28 April 2015
Accepted 18 March 2016

KEYWORDS
Dynamic decision making; learning; normal aging; computational modeling

Throughout life, people are confronted with dynamic and complex scenarios that they must learn to control in order to survive and thrive. By dynamic we mean that we experience changes in an event of which the source of change could be the result of our actions (exogenous changes to the outcome), or the result of internal changes inherent to the workings of the causal system we are interacting with in which the event takes place (endogenous changes to the outcome), or both (Brehmer, 1992; Funke, 1992; Osman, 2010). As people age normally, cognitive changes can lead to differences in the approach toward and performance in decision-making tasks (Park & Schwarz, 2012; Salthouse, 2010). Such complex dynamic environments are analogous to many real-life situations (Fisk, Rogers, Charness, Czaja, & Sharit, 2012; Giger & Markward, 2011). For example, we make several distinct health choices on a daily basis which influence our overall health and well being in uncertain ways. These types of environments, such as the human body, are often noisy, and the specific influence of the various choices is often unclear or unspecified; that is, the source of observable changes can seem to be ambiguous (e.g., the result of our dietary habits, age-related changes, or genetically

CONTACT Brian D. Glass b.glass@ucl.ac.uk Department of Computer Science, University College London, Gower Street, WC1E 6BT London, United Kingdom

predisposed factors). Consider the toil of managing a healthy diet, with the goal of remaining sufficiently nourished and satisfied with one's diet. The individual can make adjustments to several different factors, including the choice of foods, exercise routines, medication type and dosage, and other lifestyle choices. These factors all combine and interact in a way such that a single and often noisy feedback variable is returned: the current body mass index. In fact, some factors may not play a role at all, such as the intake of a multivitamin or switching between brands of bottled water. The question we ask in this study is whether these cognitive changes can be measured in a decision-making task that is designed to promote processes necessary for exerting control in a dynamic environment.

In the case of dynamic control, it is important for the individual to be able to juggle these factors, determine which have a detrimental or beneficial effect, and control them in an effective manner in order to reach and maintain a goal. In this way, dynamic decision-making environments combine a variety of important cognitive activities such as prediction, causal learning, monitoring, reasoning, feedback learning (Osman, 2010). In the laboratory, the typical set-up used to examine dynamic control involves the following basic features (for review see, Osman, 2010, 2014). Participants are shown a dynamic system (a set of input–output variables—typically continuous) and told that on each trial they are required to manipulate the input variables in order to bring the output variables to a specific target level. Even if they have reached the target level, they must still maintain it for a specified length of trials, because of the endogenous fluctuations to the output value. Thus, mimicking situations in real life, the task becomes hard to achieve because there is no advanced knowledge of the underlying input–output relationship is (e.g., linear, nonlinear), the causal structure that connects inputs to output needs to be inferred, and as mentioned before, the observed dynamic changes to the output of the system can either be exogenous (the result of participants actions), endogenous (the result of internal mechanisms in the system), or a combination of both. Thus, the aim is to use information available from the task to develop strategies, by which we mean people '…implement actions, observe their effect, and use this feedback to decide whether further action is needed or not" (Kerstholt, 1996, p. 274). To facilitate learning, controlling the output to a target value may carry a separate reward (i.e., a monetary incentive). That is, successful interventions on the inputs that lead to an output value that reaches the target value, or incrementally gets closer to the target value from trial to trial, may in turn be rewarded (Osman, 2012).

As has been commonly reported in work on dynamic control tasks, one critical activity involved in successfully controlling a dynamic environment is interpreting feedback (typically output feedback), because this helps inform the action selection process, i.e., what decision to make next, based on what happened previously. This involves appropriately dealing with positive feedback (e.g., more optimal body mass index) and negative feedback (e.g., less optimal body mass index; Osman, 2012). Though one should note that this particular activity is not only specific to dynamic control tasks in which people are required to make decisions, feedback learning is also critical to judgment tasks such as the Multiple Cue Probability learning (MCPL) task (Hammond & Summers, 1965; Smedslund, 1955). In MCPL tasks, there are noisy cue–outcome relations, which participants learn through the presentation of cue values which they use along with outcome feedback in order to predict outcomes (for review of similarities

and differences between MCPL and Control tasks, see Osman & Speekenbrink, 2012). Thus, feedback here as with dynamic control tasks is critical for updating one's knowledge in order to refine one's predictions and adapt one's strategies (i.e., which inputs to manipulate in a control task) in order to more effectively control an outcome. In the case of MCPL tasks, there is evidence to suggest that even under the noisiest cue–outcome relations, elderly (aged 65–75) and very elderly (aged 76–90) participants can learn from feedback to predict an outcome, as well as young participants (aged 20–30) (Chasseigne, Mullet, & Stewart, 1997). However, although younger populations benefited, elderly populations did not benefit from feedback in MCPL versions in which the cue–outcome relationship contained a combination of negative and positive associations (Chasseigne et al., 2004, 1997).

Turning specifically to dynamic situations, which is the focus of the present study, understanding how feedback is interpreted also allow for differences in strategy searching, namely along the exploitation and exploration dimension. This dimension is characterized by the trade-off between sampling new strategies which may have a lower expected value (i.e., exploring) versus relying on actions that one believes will result in the highest expected value (i.e., exploiting). In terms of reinforcement learning (Sutton & Barto, 1998), a decision maker seeking to maximize rewards may do so by either picking the available choice with the highest expected value (i.e., exploit) or attempt to explore the decision space by selecting from options which currently may have lower expected values (i.e., explore). In dynamic decision-making paradigms, exploration is often preferred due to the complex nature of the decision space which may not be immediately apparent or which may change over time. However, if the decision maker has an accurate internal representation of the reward space, then exploitation is preferred (Dam & Körding, 2009; Miler, 2009; Stahlman, Roberts, & Blaisdell, 2010). Moreover, these cognitive aims are all combined in one complex decision-making environment. Understanding how the underlying mechanisms involved in engaging in these sorts of task, and especially determining the role of normal aging, is critical to gain insights into how different populations adapt to uncertain circumstances.

Although complex decision-making environments tap into many different cognitive mechanisms, they can be designed for laboratory use in a manner that keeps them tractable to behavioral analysis and computational modeling. In this paper, a computational model is developed for a dynamic decision making task in order to quantitatively characterize how normal aging might affect decision-making in dynamic environments. While previous research has focused on classic paradigms such as category learning, task switching, and single-response choice procedures, little is known about normal aging in dynamic control tasks for which the participant controls multiple input variables in an uncertain task environment. These uncertain task environments are characterized by a noisy relationship between the actions the agent takes and the effects on the system. Moreover, the system may have autonomous components over which the agent has no control. The present research contrasts older adult and younger adult performance in a dynamic control task designed to simulate such real-life dynamic decision making environments (Osman & Speekenbrink, 2011).

Prior research has indicated that normal aging might affect decision making in dynamic environments in different ways. It has been shown that older adults appear to suffer from executive control deficits (Braver et al., 2001; Kray, Li, & Lindenberger,

2002; Ortega, Gómez-Ariza, Román, & Bajo, 2012). However, emerging evidence suggests that older adults can utilize compensatory strategies to return performance to or beyond baseline levels (Glass, Chotibut, Pacheco, Schnyer, & Maddox, 2012; Huang, Polk, Goh, & Park, 2012; Worthy, Gorlick, Pacheco, Schnyer, & Maddox, 2011). Older adults have also been shown to adjust their decision strategies to match task demands (Mata, Schooler, & Rieskamp, 2007). Indeed, Mata et al. (2007) speculate that older adults may have an advantage in adaptive strategy selection because of their richer knowledge of mapping appropriate strategies that correspond to the particular structure of the environment, because they have been exposed to a greater variety of task environments compared to younger adults. On this basis, it may be the case that in situations of dynamic control, older adults will be sensitive to situations in which they should adopt an exploratory strategy because they have had prior exposure to situations in the past to know when to explore more and when to exploit more. This complements a key argument made by Salthouse (2012); he suggests that in many cases, the lack of age-related deficits, particularly in real-world job-related contexts, in which people are exposed to many complex dynamic situations, can be explained by the compensation of the wealth of knowledge of strategic problem solving that older adults have to draw on. In addition, it may also be the case that when older adults perform tasks in which they experience the cumulative consequences of their decisions and actions, as they would in real life (Salthouse, 2012), they are more vigilant and so are likely to be more motivated to perform well. In addition, given the long debate regarding the predictability of cognitive measures on performance on complex tasks, to date it is still unclear whether this can be used to predict age-related effects outside other factors (Salthouse, 2012). To this end, there is an unreliable association between scores on cognitive ability measures and performance on dynamic control tasks (Osman, 2010), so there is no reason to assume that cognitive decline in executive functions would lead to predict impairments in performance in control tasks. Thus, older adults may demonstrate differential behavior patterns by relying on the use of compensatory strategies. These compensatory strategies may result from known differences in cognitive processing which occurs in the course of normal aging; these differences include biases in the exploration–exploitation trade-off as well as the positivity effect, both of which are defined below.

Normal aging is also thought to have a complex impact on the exploration–exploitation trade-off in decision-making (Hills, Mata, Wilke, & Samanez-Larkin, 2011). This trade-off is characterized by the inherent structure in many decision spaces which allow the decision-maker to shift or adapt strategic policies in lieu of persisting with one or a small handful of policies (Macready & Wolpert, 1998). Older adults who engage in complex decision-making environments often demonstrate differential behavior given the nature of the decision space. For example, older adults demonstrate exploitative behavior and relative performance difficulty when the dynamic decision-making environment does not feature a strong contingency between choice and reward (Worthy et al., 2011). However, when introduced to a dynamic decision-making environment which features a strong choice-reward contingency, older adults demonstrate exploratory behavior and a performance enhancement over younger adults. Older adults have also been shown to engage in less exploratory strategies in tasks involving information foraging (Chin et al., 2012; Mata, Wilke, & Czienskowski, 2013). The present study seeks to contribute to the understanding of the role of normal aging in the exploration–exploitation trade-off by

utilizing computational modeling of a dynamic decision-making task which involves multiple input sources and a hidden and complex underlying task mechanism, similar to many real-world scenarios. Crucially, the present study examines dynamic control-based decision-making, in which participants are required to make decisions in a dynamic environment for which they have to control a dynamic output reliably over time. The aspect of control makes our work different from previous work examining the impact of aging on dynamic decision-making (Worthy et al., 2011).

As well as developing compensatory strategies, and approaches to information search, another hallmark of the effect of normal aging on cognitive processing is the positivity effect. This is characterized by the propensity for older adults to allocate an attentional preference away from negatively valenced stimuli and toward positively valenced stimuli (Carstensen & Mikels, 2005). This positivity bias can result in differences in goal acquisition in older versus younger individuals. For example, older individuals are more likely to seek goals which are emotionally rewarding, whereas younger individuals are more prone to set information-seeking goals in an effort toward skill acquisition. Older adults have demonstrated an attentional preference toward positive versus negative features in simulated purchasing choices, as well as displaying a better memory for positive features (Kennedy & Mather, 2007). Preferences of this kind lead to the development of differential economic choice strategies between younger and older adults. For example, older adults have been found to focus on positive attributes when evaluating product alternatives, leading to different product choices than younger adults as well as increased choice satisfaction (Kim, Healey, Goldstein, Hasher, & Wiprzycka, 2008).

Prior research has found that older adults are more sensitive to negative feedback in a probabilistic selection task in comparison to a younger group (Simon, Howard, & Howard, 2010). This suggests that in certain circumstances, the positivity effect may be construed as a reaction to negative feedback, perhaps in an effort to seek positive reinforcement. It is unknown how this effect will play out in a dynamic decision-making task, although there is substantial prior evidence that differences between positive and negatively valenced feedback will have a differential impact under normal aging.

Taken together, in a dynamic decision-making environment, older individuals may develop strategies on the fly in response to positive reward feedback, as opposed to other aspects of the task. Additionally, it is unclear whether to expect older adults to demonstrate increased exploration behavior. In previous work, older adults increased exploration in tasks with high choice-reward contingencies. In a complex decision-making environment, this contingency may be clouded, resulting in exploitative behavior. However, the task at hand has additional dynamic properties which may prohibit a direct comparison to other discrete choice tasks.

In order to quantitatively gauge the complex relationships between normal aging, the exploration–exploitation trade-off, and positive feedback sensitivity, a computational model is implemented to assess individual behavioral characteristics and strategies in a dynamic control task. Moreover, dynamic control tasks are commonly taken to be examples of tasks that have high degree of validity (Funke, 1992, 1993), because in many day to day situations in the real world, we not only have to track changes in events over time, but we typically have to assert control over them. Therefore, we aim to provide valuable new insights into the impact of normal aging on dynamic control performance.

Methods

Participants

Twenty-nine younger participants aged 18 to 25 (M = 22.3, SD = 5.4) and 17 older participants aged 61 to 75 (M = 67.92, SD = 5.03) participated in the dynamic control task. The younger participants were recruited from the Queen Mary, University of London undergraduate community and paid £6 ($9.50). The older adults were recruited as a healthy control group via the National Hospital for Neurology and Neurosurgery. Older adults received the same flat fee as the younger participants, but also received reimbursement for their travel. Unlike the younger participants that were recruited on campus, the older adults were required to travel to the National Hospital in order to take part in the experiment which is why they received extra funds compared with the younger participants. To qualify for the healthy adult participation pool, the older adults completed the Beck Depression Inventory-II (Beck, Steer, Ball, & Ranieri, 1996) and Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975). All scores fell within the normal cutoff range for both the Mini-Mental State Examination (greater than 27) and Beck Depression Inventory-II (less than 18). None of the older adults had a history of neurological or physical or psychiatric illness, head injury, or drug or alcohol abuse.

Procedure

In the present dynamic control task, participants were told that they were conducting a medical test on a patient by injecting three different hormone levels (labeled A, B, and C), with the aim of maintaining a safe level of neurotransmitter release. The participant attempts to control a single output value toward a goal. To do so, on each trial, the participant chooses values (ranging from 0 to 100) for three separate inputs. These input values are then combined via the dynamic control equation (Equation 1), then summed with the output value plus some normally distributed random noise (standard deviation = 8). In this way, the participant's input selections guide the output value. The output value is initialized at 178 with a goal value of 62 and a "safe range" (±10 around the goal value).

$$y(t) = y(t-1) + 0.65x_1(t) - 0.65x_2(t) + e(t) \tag{1}$$

where $y(t)$ is the output on trial t, x_1 is the positive input, x_2 is the negative input, and e is an error term randomly sampled from a normal distribution with a mean of 0 and SD of 8. Participants were instructed to try to direct the output value into the "safe range" around the goal value.

The dynamic control equation was designed such that one input has a positive impact on the output value, one input has a negative impact, and a third input has no impact (Osman & Speekenbrink, 2011). The impact (i.e., positive or negative direction) of the input is not labeled or available to the participant; thus, the participant must learn to control the output value based solely on the resulting movement of the cumulative output value on each trial. After each trial, the input values are reset to 0. The participant can then freely select input values for each of the three inputs before confirming the choices. A critical feature of this control task is that the output value can swing below

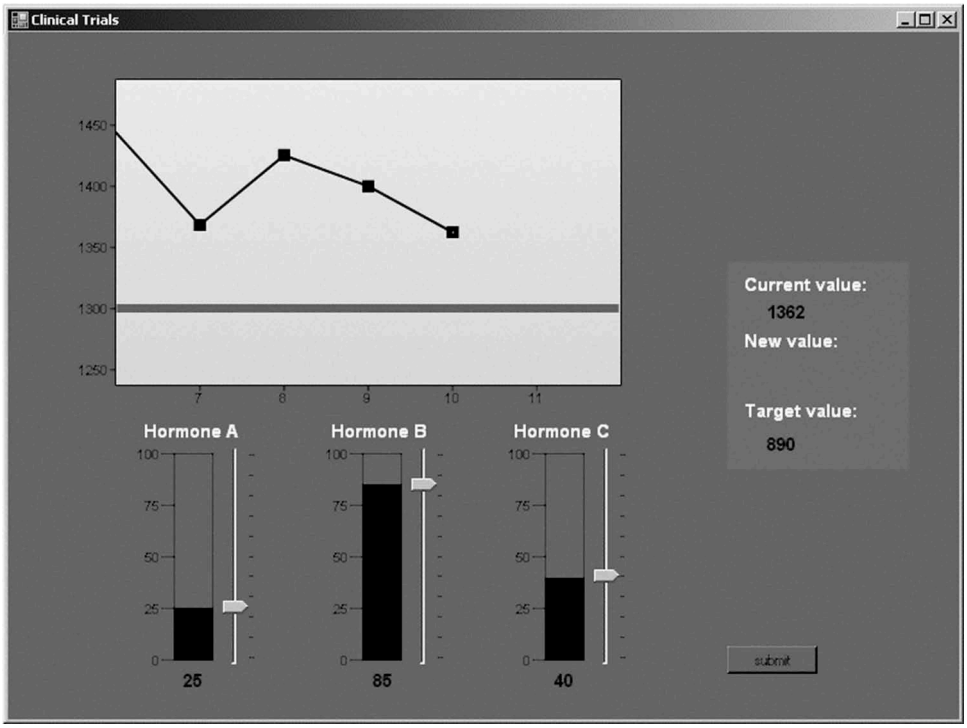

Figure 1. Screenshot of the DDM task.

the target, meaning the participant must dynamically adapt in order to maximize performance (see Figure 1).

For the learning phase, participants engaged in a single learning block of 100 trials. After this initial learning phase, all participants completed 2 test blocks of 20 trials each. The first test block was a familiar goal test in which the starting value and goal criterion were equivalent to the learning phase. The second test block was a transfer test with a different starting value and goal value than earlier training leading up to this block (i.e., a new goal test). At the beginning of each block, the control task was reset to the initial state.

Computational model

A computational model of behavior in the dynamic control task was implemented to determine behavioral characteristics of individual participants. The model, Single Limited Input, Dynamic Exploratory Responses (SLIDER), is based on memory trace reinforcement learning (Osman, Glass, & Hola, 2015). After each trial, a reinforcement history for each of the three inputs is updated according to whether the input choices resulted in the discrepancy between achieved outcome value and goal value increasing (unsuccessful trial) or decreasing (successful trial). Conversely, on an unsuccessful trial, the output value moved further from the target. On the following trial, the reinforcement history becomes the basis for a probabilistic action selection function using Luce's choice.

The SLIDER model features three free parameters: an exploitation parameter governing the action selection function and two memory-updating reinforcement strengths (one for successful trials and one for unsuccessful trials). To evaluate the model, the model's probability of selecting the human participant's input choice are combined across all trials and all three inputs into a single model fit value. The model is fit to an individual participant's responses by an optimization procedure that determines the parameter values which minimizes the computed maximum likelihood fit value. The Nelder–Mead simplex algorithm is used to find the optimal set of model parameters.

Memory-updating reinforcement strengths

After each trial, the computational model determines whether the input values it selected resulted in a successful or unsuccessful trial. For each input, a Gaussian curve[1] with a mean equal to the chosen input value is constructed (Equation 2).

$$P_{\text{update}}(v) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{v-v_p}{\sigma}\right)^2} \tag{2}$$

where P_{update} (v) is the probability of selecting a value of v when the previous selected value was v_p. This array of values is then added (successful trial) or subtracted (unsuccessful trial) to the input's former reinforcement history. A free parameter, γ_s (γ_{positive} for successful trials and γ_{negative} for unsuccessful trials) determines the relative weight of the updating summation. For example, if the memory-updating positive reinforcement strength (γ_{positive}) is 0.8, then after a successful trial, the reinforcement history is updated such that 80% of the new reinforcement history reflects the current input value choice and 20% reflects the previous reinforcement history (Equation 3).

$$P_{\text{History}}(v) = [(1-\gamma_s)P(v)] + [\text{sign}(R)\cdot\gamma_s\cdot P_{\text{update}}(v)] \tag{3}$$

where $P_{\text{History}}(v)$ is the input selection probability history for input value v, γ_s is the memory-updating reinforcement strength for feedback (positive or negative), and R is the change in the difference between the output value and the goal from the previous trial. Thus, if the sign of R is negative, then the output value has moved further from the goal (an unsuccessful trial) and the negative feedback learning parameter is implemented. Conversely, if the sign of R is positive, then the output value has moved toward the goal (a successful trial) and the positive feedback learning parameter is implemented. The input selection probability history is stored as a discrete probability function, such that each potential input response (0 to 100) is associated with a probability of selection (all summing to 1).

In summary, there are two memory-updating reinforcement strengths, one for positive feedback (output value moved toward goal) and one for negative feedback (output value moved away from goal). Each strength represents the weight with which current choices impact choice history (see Figure 2). There are two history memory vectors, one for when the output value is below the target goal and one for when it is above. Thus, the model learns to dynamically control the output value toward the target range.

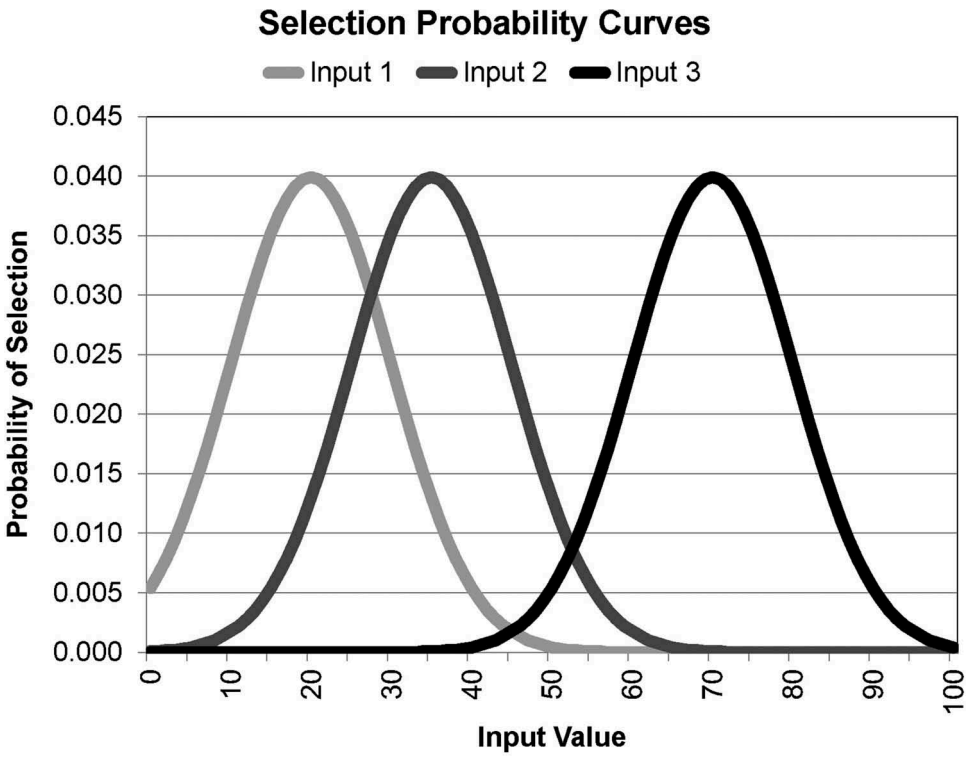

Figure 2. Sample probability density curves of selecting a given value for a given input. Over the course of a block, the curves will alter in various ways depending on the model parameters, trial success, and uncertainty inherent in the output value.

Exploitation parameter

On each trial, the computational model evaluates the reinforcement history of each input to generate the probability of selecting each of the 101 input value options (with action choices of 0 to 100). To do so, a Softmax decision rule (Equation 4) assigns probabilities to each action choice by considering the association strengths assigned by the previous steps. The equation's exploitation parameter, K, determines the level of determinism in the decision process or the certainty in picking the action with the highest expected value (Daw & Doya, 2006). As K approaches ∞, the process is more likely to choose the most probable option. As K approaches 0, the equation is more likely to pick a less probable option.

$$P_{\text{Final}}(v_i) = \frac{e^{\left[P_{\text{History}}(v_i)\cdot K\right]}}{\sum_{j=0}^{100} e^{\left[P_{\text{History}}(v_i)\cdot K\right]}} \tag{4}$$

where P_{Final} (v_i) is the final probability of selecting input value v_i, K is the exploitation parameter, and v_j are all the input values from 0 to 100 for given input. We use the SLIDER model in order to measure characteristics of trial-by-trial decision making which are dependent on the reinforcement history provided by experience on previous trials.

To accomplish this, we fit free model parameters to the empirical data in order to specify levels of exploration, negative feedback sensitivity, and positive feedback sensitivity.

Results

Behavioral analysis

The primary performance measure for this dynamic control task was the trial-by-trial optimality of the input choices in the test blocks in order to determine the role of normal aging in a familiar goal and new goal situation. By considering the optimal input actions that will maximize the output value's movement toward the target, the optimal selections can be computed for each trial. In terms of Equation 1, this was done by solving for which inputs (x_1 and x_2) would move the output value y closest to the target value (y), assuming a noise term (e) of 0. The difference between the optimal selections and the actually chosen selections results in an optimality score for each participant. Figure 3 shows that the younger group tended to select slightly more optimal responses in both test blocks, although the difference was not statistically significant. A 20 (trial) × 2 (older, younger) × 2 (familiar goal, new goal) repeated measures ANOVA revealed only a main effect of trial number, $F(19, 513) = 15.42$, $p < 0.001$. This indicated that within-block learning occurred, as is apparent in the learning curves in Figure 3.

At first blush, it may seem that the older group performed similarly to the younger group. However, further analysis of the strategies used by both groups demonstrates critical differences in the way the older adults completed the dynamic control task. The strategy analysis considered four different types of input changes: varying none, varying one input, varying two inputs, and varying all three inputs. Figure 4 illustrates the input varying strategies for both groups on both the familiar goal and new goal tests. A 2 (older, younger) × 2 (familiar goal, new goal) × 4 (strategy type) repeated measures ANOVA revealed a main effect of age condition, $F(5, 123) = 5.99$, $p = 0.019$. Furthermore, there was a significant Block by Strategy Type by Age interaction, $F(5, 123) = 2.84$, $p = 0.04$. This pattern indicates older adults varied somewhat fewer inputs on each trial, t

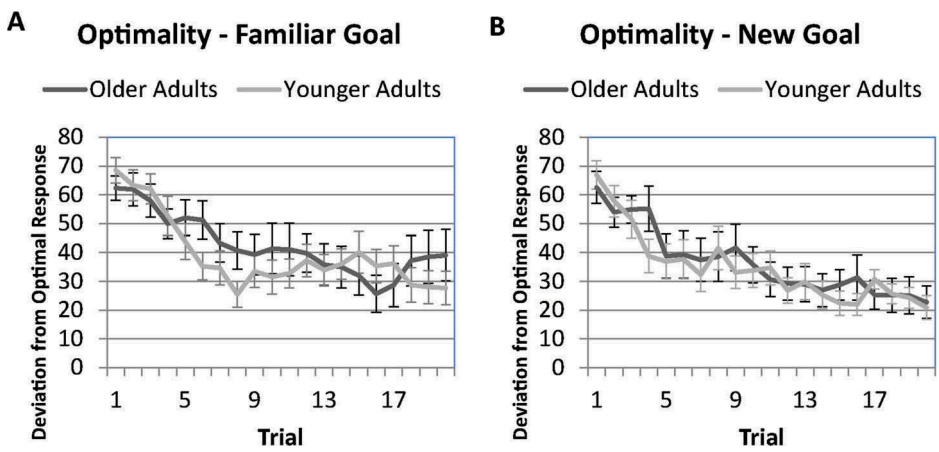

Figure 3. Deviation from optimal choices for the familiar goal (A) and new goal (B) tests (SE 1±).

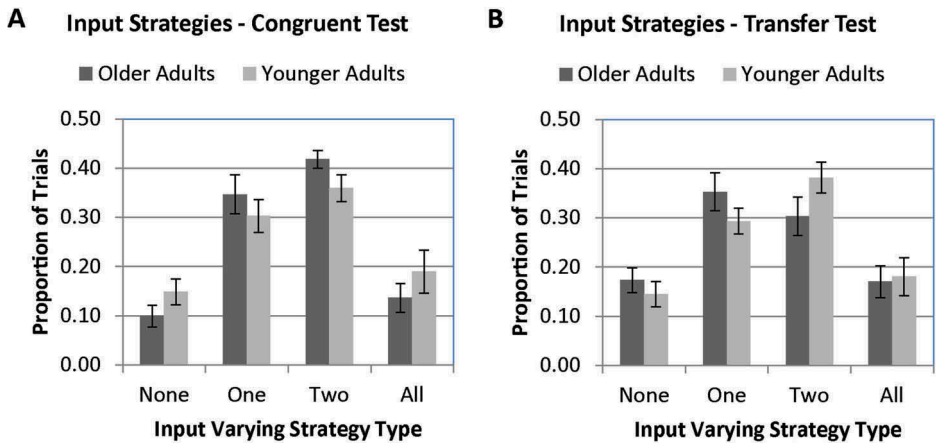

Figure 4. Mean number of input inputs varied (i.e., given a value higher than zero) for the familiar goal (A) and new goal (B) tests (SE 1±).

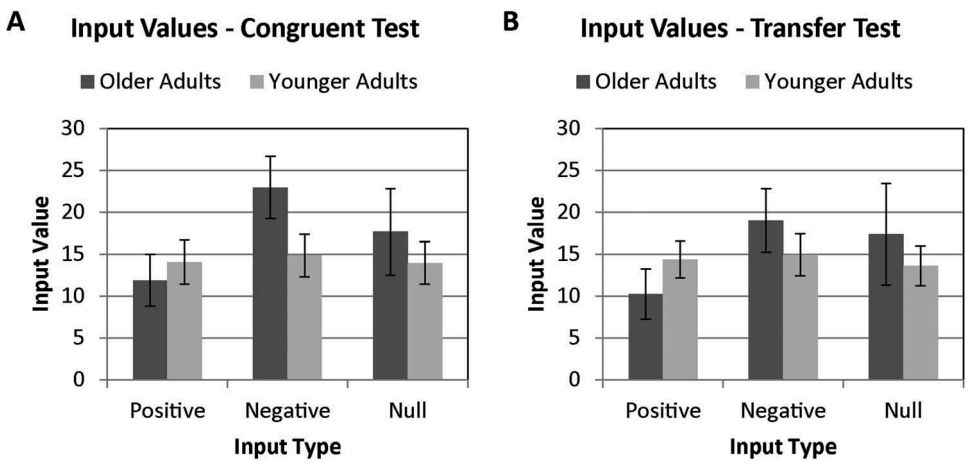

Figure 5. Mean values chosen (i.e., location of the inputs from 0 to 100) for the familiar goal (A) and new goal (B) tests (SE 1±).

(43) = −0.31, p = 0.76, although post hoc t-tests were not statistically significant (see Figure 4).

The younger and older groups also differed in the values selected for the inputs. Figure 5 reports the mean input values (between 0 and 100) selected for each of the three input types. A 2 (older, younger) × 2 (familiar goal, new goal) × 3 (input type) repeated measures ANOVA revealed a main effect of input type, $F(2, 82) = 5.03$, $p = 0.009$, as well as a input type × age interaction, $F(2, 82) = 4.08$, $p = 0.02$. This pattern suggests that overall participants chose to manipulate the negative input slightly more than the positive and null inputs, although post hoc t-tests were not statistically significant.

Taken together, analyses of surface level behavior suggest the older group differed from the younger group in the way they developed strategies in the dynamic control

task, but not with regards to performance. However, the nature of the underlying cognitive processes which led to these patterns of behavior remains elusive using only basic metrics of behavior available from the task. In order to distill psychologically relevant characteristics of the processes involved in the dynamic decision-making task performance, a computational reinforcement learning model of the dynamic control task was fit to individual participant data.

Model-based analysis

Task behavior was fit to the computational model using an optimization procedure that attempted to minimize the difference between observed trial-by-trial input value selections and the expected input value selections as determined by the model. This was done by considering the probabilities given to the various input values for each input on a given trial. The optimization procedure attempted to determine best fitting free parameters (exploitation parameter, positive and negative reinforcement sensitivity parameter) that maximized the probability that the model would select the same input values as the human participant on a given trial.

The exploitation parameter offers insight into how a decision on any given trial relates to learning history from previous trials. This offers insights which differ from considering only the exact choices an individual makes, since the same set of actions may be deemed exploratory or exploitative depending on the context of the previous choices made. In general, there was an alignment with exploitation parameters and the selection of higher input values. For example, the best-fit exploitation parameter was negatively correlated with the size of the input values during the familiar goal block (positive input, $R = -0.17$, $p = 0.27$; negative input, $R = -0.41$, $p < 0.01$; and neutral input, $R = -0.30$, $p < 0.05$) and directionally negatively correlated during the new goal block (positive input, $R = -0.11$, $p = 0.47$; negative input, $R = -0.18$, $p = 0.24$; and neutral input, $R = -0.11$, $p < 0.47$). This trend suggests that exploitative choice is somewhat associated with more conservative action selection.

Figures 6 and 7 reports the mean best-fit parameter values for the younger and older groups. A 2 (older, younger) × 2 (familiar goal, new goal) repeated measures ANOVA was conducted on the best fitting exploitation parameter. There was a main effect of age group, $F(1, 41) = 4.37$, $p < 0.042$, such that the younger group's performance was better fit by higher exploitation parameters. For the positive sensitivity parameter, there was also a main effect of age group, $F(1, 41) = 8.77$, $p = 0.005$, such that the older group's performance was better fit by learning parameters more sensitive to positive feedback. There were no significant effects for the negative sensitivity parameter, which were all quite low. In short, the older group's performance was better fit with model parameters associated with higher exploration and higher positive feedback sensitivity.

In the familiar goal block, choice optimality (i.e., deviation from optimal response) was correlated positively with exploitation parameter for older adults ($R^2 = 0.26$, $p < 0.05$), but not for younger adults ($R^2 = 0.003$, $p = 0.77$). Thus, older adults who made more exploratory choices also achieved better performance. In the new goal block, exploitation parameter was not associated with better performance for either age group.

Partial least squares regression determined the fit of the SLIDER model to the behavioral data. In both age groups, the SLIDER model was compared to a baseline

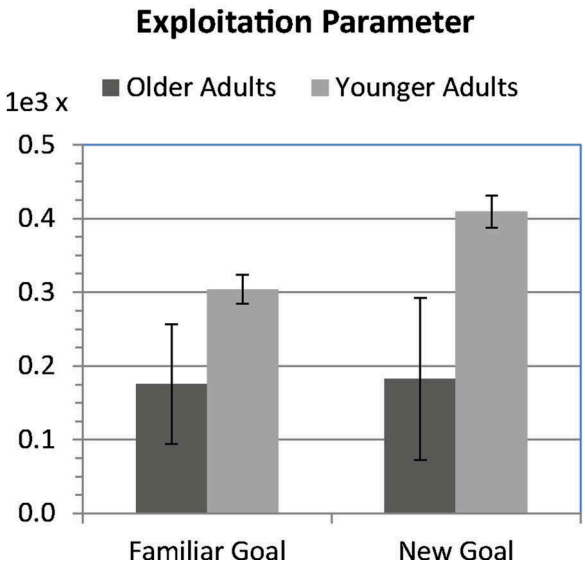

Figure 6. Best fitting exploitation parameter (lower values suggest increased exploration) recovered from the model based analysis (SE 1±).

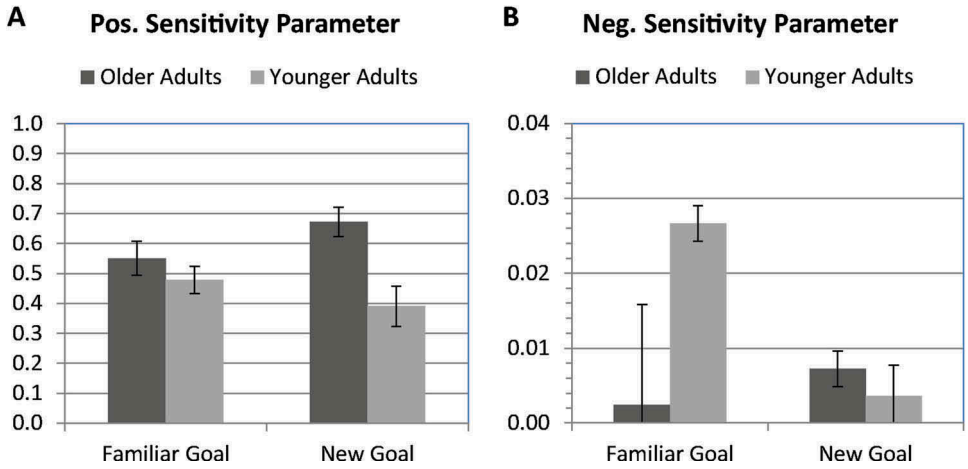

Figure 7. Best fitting positive (A) and negative (B) sensitivity parameters recovered from the model based analysis (SE 1±).

model. First, a simple autocorrelation determined the amount of variance explained by simply considering the response to the previous trial (resulting in a baseline model). Second, a correlation between observed and model predicted responses was performed to determine the amount of variance explained by the computational model. Finally, Akaike's Information Criterion (AIC) was calculated in order to compare the baseline model with the computational model (Akaike, 1974; Symonds & Moussalli, 2011). The computational model explained more of the variance (R^2 = 0.50) than both the simple autocorrelation method (R^2 = 0.37) as well as a random responder method (which by

definition would select the correct state on only 1 out of 100 responses, putting R^2 close to 0). The AIC is used to punish models for additional free parameters, with lower AIC values signifying a better model fit. For the computational model, the average AIC = 586.2 (assuming 4 free parameters). For the baseline model, the average AIC = 599.1 (assuming no free parameters). Thus, the computational model is preferred over the baseline model for assumptions of free parameters up to 10.

Discussion

The present study examined the role of normal aging and learning style in a dynamic control task using a novel application of computational modeling. Standard behavioral analysis revealed older adults potentially utilized an alternative strategy in completing the dynamic control task than younger adults. Our initial predictions were based on prior research in normal aging. With the relatively robust nature of the positivity effect in other research, we expected older adults to be better fit by models with learning parameters that were more sensitive to positive feedback. Indeed, we found this to be the case. We also predicted that it could be possible that older adults would be better fit by models with higher exploitation parameters, given that older adults have been shown to be more exploitative in situations with low choice-reward contingency. In addition, as mentioned earlier, work examining the impact of age and task difficulty on judgment performance in MCPL tasks suggests that elderly populations struggle to learn more complex cue–outcome contingencies, particularly from feedback (e.g., Chasseigne et al., 2004, 1997; Mutter & Williams, 2004). Of note is the work by Chasseigne (Chasseigne et al., 2004, 1997) which suggests that elderly populations struggle to learn negative cue–outcome relations when embedded in more complex MCPL tasks, but have little difficulty showing comparable performance accuracy to younger adults when it comes to learning positive linear relationships. Though MCPL and dynamic control tasks have many differences (Osman & Speekenbrink, 2012), they also share some properties in common—in particular feedback learning, and to this end, the findings from the present study compliment the work by Chasseigne (Chasseigne et al., 2004, 1997). In both cases, the evidence suggests that there is a difference in the way elderly populations process negative feedback as compared to positive feedback or negative relations compared to positive ones.

Considering general task performance based on immediately available surface behavior (e.g., input values chosen, output value performance), these styles did not lead to differences in choice optimality. However, the learning styles did appear to have an impact on which input values were manipulated, and the values they chose when they decided to manipulate the inputs. The interpretation of the underlying reasons for these differences can provide a clearer connection with well-established psychological constructs. To this end, we developed model-based analyses which permitted the measurement of strategic exploration–exploitation trade-off, as well as differential reinforcement learning sensitivity to positive and negative feedback.

A computational model of the task revealed specific behavioral characteristics associated with normal aging. In the familiar goal block, older adults demonstrated more exploratory behavior and somewhat more reliance on recent and positive success. On the transfer block in which the target output value that participants had to reach and

maintain, older adults continued to demonstrate exploratory behavior while younger adults became more exploitative. Older adults continued to demonstrate more reliance on recent and positive success, as evidenced by higher positive sensitivity parameters.

One possible interpretation of this pattern of results is that older adults were able to achieve the final performance profile of younger adults (as measured by deviation from optimal responses) by relying on compensatory mechanisms to engage the task. Specifically, the older adults were more exploratory, which has been shown to lead to enhanced performance in dynamic decision-making tasks with choice–reward contingency (Worthy et al., 2011). Additionally, older adults remained more influenced by positive feedback on both tests. This interpretation is supported by previous research which has shown that older adults are able to achieve the performance levels of younger adults via a compensatory strategy (Glass et al., 2012; Worthy & Maddox, 2012) by successfully engaging in adaptive strategy selection to match task demands (Mata et al., 2012).

Another interpretation of the current results is that older adults approached the task by utilizing alternative mechanisms which may be enhanced in older adults. For example, older adults exhibit a positivity effect characterized by superior emotional processing of positively valenced content and contingencies (Carstensen & Mikels, 2005; Mutter & Williams, 2004). In terms of the current task setup, this could mean that older adults were more sensitive to successful trials relative to unsuccessful trials. This could account for the older adults' higher learning rate sensitivity parameter for positive feedback, but not for negative feedback. Thus, when older adults encountered successful trials, their learning rate parameters increased such that prior knowledge was discounted. In this interpretation, older adults differed in their overall strategy due to specific enhancements associated with normal aging. This interpretation is supported by the positive learning rate sensitivity parameter remaining higher for older adults than younger adults in both the familiar goal and new goal tasks.

These above interpretations are informed by correlating the exploitation parameter with task performance. When the goal remained familiar, exploratory choices were associated with increased performance in older adults, but not younger adults. This suggests that older adults were able to use exploration to their advantage when the goal remained familiar. Once the goal changed in the final test block, older adults were able to match the performance of younger adults, although this was not associated with the exploration parameter. Thus, these results suggest that older adults may be able to utilize exploration as a compensatory strategy, but this compensation may be localized to dynamic decision-making scenarios in which the goal is familiar.

Future research should determine whether the differences in strategies used by older adults to complete the dynamic control task are simply the result of slower overall learning rates or due to differences in underlying cognitive mechanisms associated with normal aging. Future work should incorporate manipulations to test these interpretations, such as limiting feedback types to determine whether the aging positivity effect can account for performance differences. Dynamic decision making tasks such as the one used in the present study are an effective way to investigate decision making in situations which mimic real-world dynamic environments. The present study speaks to the effectiveness of using dynamic decision-making tasks to reveal the impact of aging on day-to-day complex decision-making.

Notes

1. Here, sigma values were fixed to 10, equivalent to 10% of the available range of choices, although the results are robust to other sigma values.

Acknowledgements

The authors thank Agata Ryterska for help with data collection for this project.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control IEEE Transactions on*, *19*, 716–723.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of beck depression inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, *67*, 588–597. doi:10.1207/s15327752jpa6703_13

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., … Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*, 746–763. doi:10.1037/0096-3445.130.4.746

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, *81*, 211–241. doi:10.1016/0001-6918(92)90019-A

Carstensen, L. L., & Mikels, J. A. (2005). At the intersection of emotion and cognition. Aging and the positivity effect. *Current Directions in Psychological Science*, *14*, 117–121. doi:10.1111/cdir.2005.14.issue-3

Chasseigne, G., Ligneau, C., Grau, S., Gall, A. L., Roque, M., & Mullet, E. (2004). Aging and probabilistic learning in single- and multiple-cue tasks. *Experimental Aging Research*, *30*, 23–45. doi:10.1080/03610730490251469

Chasseigne, G., Mullet, E., & Stewart, T. R. (1997). Aging and multiple cue probability learning: The case of inverse relationships. *Acta Psychologica*, *97*, 235–252. doi:10.1016/S0001-6918(97)00034-6

Chin, J., Payne, B., Battles, A., Fu, W. T., Morrow, D., & Stine-Morrow, E. (2012). Information foraging in the unknown patches across the life span. In *Proceedings of the 34th annual meeting of the Cognitive Science Society* (pp. 1404–1409). Sapporo: Cognitive Science Society.

Dam, G., & Körding, K. (2009). Exploration and exploitation during sequential search. *Cognitive Science*, *33*, 530–541. doi:10.1111/cogs.2009.33.issue-3

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*, 199–204. doi:10.1016/j.conb.2006.03.006

Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2012). *Designing for older adults: Principles and creative human factors approaches*. Boca Raton, FL: CRC press.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198. doi:10.1016/0022-3956(75)90026-6

Funke, J. (1992). Dealing with dynamic systems: Research strategy, diagnostic approach and experimental results. *German Journal of Psychology*, *16*, 24–43.

Funke, J. (1993). Microworlds based on linear equation systems: A new approach to complex problem solving and experimental results. In G. Strube & K.-F. Wender (Eds.), *The cognitive psychology of knowledge* (pp. 313–330). Amsterdam: Elsevier.

Giger, J. T., & Markward, M. (2011). The need to know caregiver perspectives toward using smart home technology. *Journal of Social Work in Disability & Rehabilitation*, *10*, 96–114. doi:10.1080/1536710X.2011.571529

Glass, B. D., Chotibut, T., Pacheco, J., Schnyer, D. M., & Maddox, W. T. (2012). Normal aging and the dissociable prototype learning systems. *Psychology and Aging*, *27*, 120–128. doi:10.1037/a0024971

Hammond, K. R., & Summers, D. A. (1965). Cognitive dependence on linear and nonlinear cues. *Psychological Review*, *72*, 215–224. doi:10.1037/h0021798

Hills, T. T., Mata, R., Wilke, A., & Samanez-Larkin, G. R. (2011). Exploration and exploitation in memory search across the lifespan. In L. Carlson, et al. (Eds.) *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 991–996). Boston, MA: Cognitive Science Society.

Huang, C.-M., Polk, T. A., Goh, J. O., & Park, D. C. (2012). Both left and right posterior parietal activations contribute to compensatory processes in normal aging. *Neuropsychologia*, *50*, 55–66. doi:10.1016/j.neuropsychologia.2011.10.022

Kennedy, Q., & Mather, M. (2007). Aging, affect and decision making. In K. D. Vohs, R. F. Baumeister, & G. Loewenstein (Eds.), *Do emotions help or hurt decision making? A hedgefoxian perspective* (pp. 245–265). New York, NY: Russel Sage Foundation.

Kerstholt, J. (1996). The effect of information costs on strategy selection in dynamic tasks. *Acta Psychologica*, *94*, 273–290. doi:10.1016/S0001-6918(96)00011-X

Kim, S., Healey, M. K., Goldstein, D., Hasher, L., & Wiprzycka, U. J. (2008). Age differences in choice satisfaction: A positivity effect in decision making. *Psychology and Aging*, *23*, 33–38. doi:10.1037/0882-7974.23.1.33

Kray, J., Li, K. Z. H., & Lindenberger, U. (2002). Age-related changes in task-switching components: The role of task uncertainty. *Brain and Cognition*, *49*, 363–381. doi:10.1006/brcg.2001.1505

Macready, W. G., & Wolpert, D. H. (1998). Bandit problems and the exploration/exploitation tradeoff. *Evolutionary Computation IEEE Transactions on*, *2*, 2–22.

Mata, R., Pachur, T., Von Helversen, B., Hertwig, R., Rieskamp, J., & Schooler, L. (2012). Ecological rationality: A framework for understanding and aiding the aging decision maker. *Frontiers in Neuroscience*, *6*, 19.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, *22*(4), 796.

Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: Is there a reduction in exploration with aging. *Frontiers in Neuroscience*, *7*(53), doi:10.3389/fnins.2013.00053.

Miler, K. (2009). Doomed to repeat the successes of the past: History is best forgotten for repeated choices with nonstationary payoffs. *Memory & Cognition*, *37*, 985–1000. doi:10.3758/MC.37.7.985

Mutter, S. A., & Williams, T. W. (2004). Aging and the detection of contingency in causal learning. *Psychology and Aging*, *19*, 13–26. doi:10.1037/0882-7974.19.1.13

Ortega, A., Gómez-Ariza, C. J., Román, P. E., & Bajo, M. T. (2012). Memory inhibition, aging and the executive deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *38*, 178–186.

Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin*, *136*, 65–86. doi:10.1037/a0017815

Osman, M. (2012). The role of feedback in dynamic decision making. *Frontiers in Decision Neuroscience: Human Choice and Adaptive Decision Making*, *6*, 56.

Osman, M. (2014). *Future-minded: The psychology of agency and control*. Basingstoke, UK: Palgrave MacMillan.

Osman, M., Glass, B. D., & Hola, Z. (2015). Approaches to learning to control dynamic uncertainty. *Systems*, *3*, 211–223. doi:10.3390/systems3040211

Osman, M., & Speekenbrink, M. (2011). Cue utilization and strategy application in stable and unstable dynamic environments. *Cognitive Systems Research*, *12*(3), 355–364.

Osman, M., & Speekenbrink, M. (2012). Prediction and control in a dynamic environment. *Frontiers in Psychology*, *3*, 1–12. doi:10.3389/fpsyg.2012.00068

Park, D. C., & Schwarz, N. (Eds.). (2000). *Cognitive aging: A primer*. Philadelphia, PA: Psychology Press.

Salthouse, T. (2012). Consequences of age-related cognitive declines. *Annual Review of Psychology*, *63*, 201–226. doi:10.1146/annurev-psych-120710-100328

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, *16*, 754–760. doi:10.1017/S1355617710000706

Simon, J. R., Howard Jr, J. H., & Howard, D. V. (2010). Adult age differences in learning from positive and negative probabilistic feedback. *Neuropsychology*, *24*, 534–541. doi:10.1037/a0018652

Smedslund, J. (1955). *Multiple probability learning*. Oslo: Akademisk Forlag.

Stahlman, W. D., Roberts, S., & Blaisdell, A. P. (2010). Effect of reward probability on spatial and temporal variation. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*, 77.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT press.

Symonds, M. R., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, *65*, 13–21. doi:10.1007/s00265-010-1037-6

Worthy, D. A., Gorlick, M. A., Pacheco, J. L., Schnyer, D. M., & Maddox, W. T. (2011). With age comes wisdom: Decision making in younger and older adults. *Psychological Science*, *22*, 1375–1380. doi:10.1177/0956797611420301

Worthy, D. A., & Maddox, W. T. (2012). Age-based differences in strategy-use in choice tasks. *Frontiers in Neuroscience*, *5*, 1–10. doi:10.3389/fnins.2011.00145