Supplementary Procedure

Symmetrical outcome-revaluation task (SORT)

Participants were instructed that they were completing a computerised game, the *Starving Pedestrian Game*, where they played a hungry pedestrian and had to collect ice creams (and points) from an ice cream delivery truck to satisfy their hunger.

Demonstration phase. The demonstration phase introduced participants to the task, with example stimuli and outcomes, different to the subsequent phases. The stimuli were four scooters with an associated logo and delivery man, and outcomes were four different types of pizza slices (Margherita, Pepperoni, Vegetarian, and Hawaiian). The demonstration phase consisted of 2 blocks of 8 trials with differing stimuli (scooters) and outcomes (pizza slices) presented throughout. Participants were tasked with earning points by collecting pizzas; however, some pizzas were valuable (satisfied their hunger, +1 point) and others were nonvaluable (made them feel sick, -1 point). Participants were told that each of the four scooters carried a different type of pizza slice and that they had to learn the stimulus-outcome (S-O) associations by trial and error. The S-O associates remained constant throughout the demonstration phase. At the start of the phase, participants were presented an outcome-value instruction screen identifying which two pizzas were valuable (presented in a green frame), and which two were nonvaluable (presented in a red frame). Pizzas were collected by responding (pressing the spacebar) when the scooter and delivery man (stimulus) appeared on the screen and participants were instructed to respond quickly before the pizza (outcome) appeared, as responses after this were “late responses” and not awarded points. Participants received feedback indicating the number of correct, incorrect, and missed responses as well as their total points score at the end of each block.

Pavlovian learning phase. For the remaining phases, the stimuli were represented by eight trucks, each with a different coloured symbol on the side. Outcomes (O) were represented by four

ice creams (Cornetto, Magnum, Rocket, soft serve ice cream cone). The Pavlovian learning phase provided participants the opportunity to form stimulus-outcome (S-O) associations between the trucks and ice creams. This phase consisted of 2 blocks of 8 trials. Participants were told that the same game instructions from the demonstration phase applied but that they should forget about the scooters and pizzas, and they were instead introduced to the truck stimuli and ice cream outcomes. Participants were instructed to gradually learn which ice cream was delivered by each truck based on its logo. They were presented all eight trucks to study and informed that each ice cream was delivered by two different trucks. Each trial presented an ice cream atop one of its associated trucks for one second and participants were encouraged to focus on these associations to start forming S-O links between trucks and ice creams. After each Pavlovian learning block participants completed a Stimulus-Outcome (S-O) Knowledge Test where they were presented four of the trucks and one ice cream, and instructed to match the ice cream to the truck that delivers it as quickly as possible.

Instrumental training phase. The training phase consisted of 24 blocks of 16 trials, with valuable and nonvaluable stimuli presented an equal number of times. At the start of each block, participants were informed of the ice cream values specific to that block (outcome-value instruction screen). During training, only four trucks featured in each block, meaning that each ice cream was carried by one truck per block. Each trial began with an ice cream truck with a logo and a vendor (stimulus) appearing in the centre. The logo on each truck signified which type of ice cream it carried. Like the demonstration phase, participants had to respond (press the spacebar) to collect the ice cream if it was carrying a valuable ice cream. If the truck was carrying a nonvaluable ice cream, they did not need to respond. The inter-stimulus-interval (ISI; response window) was either 500ms for younger or 800ms for older adults. The end of the response window was signified by the vendor disappearing and there was a 200ms delay interval before the outcome ice cream was presented atop the truck for 700ms. Inter-trial-intervals (ITIs) varied between 1000-1500ms.

During training, four trucks always contained valuable ice creams, while the other four always contained nonvaluable ice creams. This response consistency allowed participants to form direct S-R links between the trucks and either responding or not-responding. Unlike the demonstration phase, participants did not receive immediate feedback on their performance accuracy in-between trials. Instead, they had to determine this independently using their knowledge of the current values of the ice creams. During training, participants only received feedback between each block (same feedback from learning phase, displayed for 5 seconds). After every eight blocks, participants could take an optional self-paced break of up to 30 seconds. After the last block of training, participants completed a third S-O Knowledge Test and the Self-Report Behavioural Automaticity Index (SRBAI; Gardner, Abraham, et al., 2012).

Test phase. The test phase consisted of two blocks of 32 trials. Each of the eight trucks were presented four times per block, and trials were randomised per 16. During the test phase, each block began with the outcome-value instruction screen indicating valuable and nonvaluable ice creams. Test trials ran similarly to training; however, some key manipulations were made. First, ice creams displayed atop the trucks at the end of trials were covered and no longer visible. In addition, all eight trucks were now featured in each block, meaning that each ice cream was carried by two trucks per block. Most importantly, some ice cream outcome-values changed for each block. This created four different trial types; still-valuable, upvalued, still-nonvaluable, and devalued. Collecting valuable and nonvaluable ice creams still resulted in the respective gain and loss of 1 point, but participants did not receive immediate feedback on performance accuracy between blocks. Instead, they were presented with response accuracy at the end of the test phase, so they did not continue to learn S-R associations during testing. The test phase consisted of a fast and slow ISI. The fast condition was 500ms/800ms (younger/older adults) and the slow condition was 800ms/1100ms (younger/older adults).

Self-Report Behavioural Automaticity Index (SRBAI)

The SRBAI (Gardner, Lally, et al., 2012) is a four-item scale where participants reflected on how behaviour is initiated; "before I realize I am doing it", "automatically", "without having to consciously remember", or "without thinking". Participants rated their responses on a scale of 1 (strongly disagree) to 7 (strongly agree), and a mean score was calculated for each item. The SRBAI measured how automatic participants felt their response or nonresponse to each stimulus truck was. The SRBAI has demonstrated reliability, convergency, and predictive validity for detecting health habits and has been previously used to assess automaticity in the context of instrumental learning tasks (Gardner, Lally, et al., 2012; Watson et al., 2022).

Supplementary Results

SORT S-O Knowledge

Table 1

Means and Standard Deviations on SORT S-O Knowledge by Age Group

Timepoint	Younger adults		Older adults	
	M	*SD*	*M*	*SD*
T1	72.50	21.95	51.79	19.48
T2	79.50	18.71	67.26	24.20
T3	98.50	4.15	81.55	20.77

Note. $N = 46$ ($n = 25$ younger adults, $n = 21$ older adults). S-O knowledge measured as accuracy (%) halfway through Pavlovian learning (T1), the end of Pavlovian learning (T2), and the end of training phase (T3).

SORT Training Phase ANOVA

Table 2

Means and Standard Deviations on SORT Training by Age Group

Trial	Younger adults		Older adults	
	M	*SD*	*M*	*SD*
Valuable Go				
Block 1	79.50	17.35	58.93	20.11
Block 2	88.75	14.77	64.29	14.41
Block 3	86.75	17.71	65.48	21.97
Block 4	91.50	13.12	78.57	17.63
Block 5	88.00	15.09	79.17	13.60
Block 6	91.75	19.82	81.25	16.54
Block 7	92.75	12.33	83.04	20.84
Block 8	89.75	17.57	84.23	16.49
Block 9	91.75	18.02	87.80	12.73
Block 10	89.50	18.55	87.20	10.91
Block 11	91.50	20.64	89.88	11.09
Block 12	92.50	17.40	91.07	13.49
Nonvaluable NoGo				
Block 1	77.00	23.71	82.44	15.26
Block 2	88.00	14.54	75.89	16.80
Block 3	88.50	12.84	82.14	16.33
Block 4	91.25	10.05	85.72	11.72
Block 5	91.25	14.77	88.69	12.44
Block 6	93.5	12.04	91.37	10.17
Block 7	96.5	7.23	91.07	13.77

Block 8	94.25	6.49	88.10	14.10
Block 9	97.00	4.46	90.18	14.06
Block 10	96.00	6.72	91.67	13.60
Block 11	97.50	4.77	94.05	11.44
Block 12	98.00	4.32	94.05	10.36

Note. N = 46 (n = 25 younger adults, n = 21 older adults). SORT training accuracy was measured as the % of correct responses on valuable Go and nonresponses on nonvaluable NoGo trials, collapsed per 2 blocks.

SORT Test Phase ANOVA

Table 3

Means, Standard Deviations, and Pairwise Comparisons on SORT Test by Age Group

Trial type	Younger adults		Older adults		Pairwise comparisons		
	M	*SD*	*M*	*SD*	M_{diff}	*p*	95% Confidence Interval
Short condition							
Still valuable	71.75	16.84	71.50	21.43	.25	.964	(-10.71, 11.21)
Upvalued	42.50	24.74	44.25	27.54	-1.75	.814	(-16.64, 13.14)
Still nonvaluable	87.50	10.36	92.50	12.63	-5.00	.133	(-11.57, 1.57)
Devalued	66.50	17.01	72.00	21.73	-5.50	.324	(-16.60, 5.60)
Long condition							
Still valuable	85.25	13.61	83.96	17.41	1.29	.772	(-7.60, 10.18)
Upvalued	79.75	20.91	50.00	30.57	29.75^{***}	<.001	(14.60, 44.64)
Still nonvaluable	88.96	14.01	90.50	13.40	-1.54	.693	(-9.33, 6.25)
Devalued	71.75	20.34	73.75	28.64	-2.00	.777	(-16.13, 12.13)

Note. N = 50 (n = 25 in each age group). SORT test accuracy was measured as the % of correct responses on valuable and nonresponses on nonvaluable trials. Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults. M_{diff} = (younger – older). Pairwise comparisons are Bonferroni adjusted for multiple comparisons.

***indicates significance at $p < .001$.

Table 4

Pairwise comparisons on SORT test accuracy between response windows by age group

Trial type	Pairwise comparisons					
	Younger adults			Older adults		
	M_{diff}	*p*	95% Confidence Interval	M_{diff}	*p*	95% Confidence Interval

Still valuable	-13.5***	.001	(-21.43, -5.57)	-12.46***	.003	(-20.39, -4.53)
Upvalued	-37.25***	<.001	(-46.88, -27.62)	-5.75	.236	(-15.38, 3.88)
Still nonvaluable	-1.46	.611	(-7.19, 4.27)	2.00	.486	(-3.73, 7.73)
Devalued	-5.25	.161	(-12.67, 2.17)	-1.75	.638	(-9.17, 5.67)

Note. N = 50 (*n* = 25 in each age group). M_{diff} = (short – long). Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults. Pairwise comparisons are Bonferroni adjusted for multiple comparisons. ***indicates significance at $p < .001$.

Table 5

Pairwise comparisons on SORT test accuracy between value-congruency by age group

Trial types compared	Pairwise comparisons					
	Younger adults			Older adults		
	M_{diff}	*p*	95% Confidence Interval	M_{diff}	*p*	95% Confidence Interval
Short condition						
Go	29.25***	<.001	(19.22, 39.29)	27.25***	<.001	(17.22, 37.29)
NoGo	21.00***	<.001	(13.42, 28.58)	20.50***	<.001	(12.92, 28.08)
Long condition						
Go	5.50	.293	(-4.91, 15.91)	33.96***	<.001	(23.55, 44.37)
NoGo	17.21***	<.001	(8.93, 25.50)	16.75***	<.001	(8.47, 25.04)

Note. N = 50 (*n* = 25 in each age group). Go M_{diff} = (still valuable - upvalued). NoGo M_{diff} = (still nonvaluable - devalued). Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults. Pairwise comparisons are Bonferroni adjusted for multiple comparisons. ***indicates significance at $p < .001$.

Table 6

Pairwise comparisons on SORT test accuracy between response type by age group

Trial types compared	Pairwise comparisons					
	Younger adults			Older adults		
	M_{diff}	*p*	95% Confidence Interval	M_{diff}	*p*	95% Confidence Interval
Short condition						
Congruent	-15.75***	<.001	(-24.09, -7.14)	-21.00***	<.001	(-29.34, -12.66)
Incongruent	-24.00***	<.001	(-34.70, -13.30)	-27.75***	<.001	(-38.45, -17.05)
Long condition						
Congruent	-3.71	.212	(-9.61, 2.19)	-6.54*	.031	(-12.44, -.64)
Incongruent	8.00	.117	(-2.09, 18.09)	-23.75***	<.001	(-33.84, -13.66)

Note. N = 50 (*n* = 25 in each age group). Congruent M_{diff} = (still valuable – still nonvaluable). Incongruent M_{diff} = (upvalued - devalued). Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults. Pairwise comparisons are Bonferroni adjusted for multiple comparisons. ***indicates significance at $p < .001$. ***** represents a significant difference at $p < .05$.

Working memory ANCOVA

To control for the effects of participant's working memory capabilities, we added digit spans backwards scores as a covariate to the analysis. After controlling for working memory, the four-way interaction remained significant, $F(1,47) = 8.38$, $p = .006$, $\eta p^2 = .15$. The main effects also remained unchanged with a non-significant effect of age-group, $F(1,48) = .40$, $p = .531$, $\eta p^2 = .01$, suggesting that both younger and adults continued to perform similarly overall. The main effect of value-congruency was still significant as well ,$F(1,48) = 13.59$, $p < .001$, $\eta p^2 = .22$, with both age groups continuing to perform slips of actions on the incongruent trials ($M_{diff} = -21.43$, $p < .001$, 95% CI [-26.18, -16.68]). Both age groups also continued to perform more accurately when given longer to respond in the long response window condition ($M_{diff} = 9.43$, $p < .001$, 95% CI [5.86, 13.00]) as demonstrated by the significant main effect of response window, $F(1,48) = 4.83$, $p = .033$, $\eta p^2 = .09$. The main effect of response also remained significant, $F(1,48) = 7.24$, $p = .010$, $\eta p^2 = .13$, with participants performing more accurately on trials where a response was not required (NoGo) compared to when a response was required (Go; $M_{diff} = 14.31$, $p < .001$, 95% CI [9.68, 18.95]).

Table 7

Pairwise Comparisons on SORT Test by Age Group when Controlling for working memory

Trial type	Pairwise comparisons		
	M_{diff}	*p*	95% Confidence Interval
Short condition			
Still valuable	.27	.963	(-11.24, 11.77)
Upvalued	.16	.984	(-15.33, 15.65)
Still nonvaluable	-4.81	.167	(-11.70, 2.09)

Devalued	-5.40	.356	(-17.05, 6.26)
Long condition			
Still valuable	1.57	.736	(-7.75, 10.90)
Upvalued	30.33***	<.001	(14.71, 45.95)
Still nonvaluable	-1.08	.791	(-9.25, 7.09)
Devalued	-2.00	.788	(-16.83, 12.83)

Note. N = 50 (*n* = 25 in each age group). M_{diff} = (younger – older). Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults.

***represents a significant difference at $p < .001$.

Table 8

Pairwise comparisons on SORT test accuracy between response windows by age group controlling for working memory

Trial type	Pairwise comparisons					
	Younger adults			Older adults		
	M_{diff}	*p*	95% Confidence Interval	M_{diff}	*p*	95% Confidence Interval
Still valuable	-13.63*	.002	(-21.81, -5.46)	-12.33*	.004	(-20.50, -4.15)
Upvalued	-36.60***	<.001	(-46.46, -26.71)	-6.41	.198	(-16.30, 3.46)
Still nonvaluable	-1.59	.590	(-7.50, 4.31)	2.13	.471	(-3.77, 8.04)
Devalued	-5.20	.178	(-12.85, 2.45)	-1.80	.638	(-9.45, 5.85)

Note. N = 50 (*n* = 25 in each age group). M_{diff} = (long – short). Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults. ***represents a significant difference at $p < .001$. *represents a significant difference at $p < .05$.

Table 9

Pairwise comparisons on SORT test accuracy between value-congruency by age group controlling for working memory

Trial types compared	Pairwise comparisons					
	Younger adults			Older adults		
	M_{diff}	*p*	95% Confidence Interval	M_{diff}	*p*	95% Confidence Interval
Short condition						
Go	28.31***	<.001	(18.06, 38.55)	28.20***	<.001	(17.95, 38.44)
NoGo	21.04***	<.001	(12.23, 28.86)	20.46***	<.001	(12.64, 28.27)
Long condition						
Go	5.35	.321	(-5.37, 16.08)	34.12***	<.001	(23.39, 44.83)
NoGo	17.44***	<.001	(8.91, 25.97)	16.52***	<.001	(7.99, 25.06)

Note. N = 50 (*n* = 25 in each age group). Go M_{diff} = (still valuable - upvalued). NoGo M_{diff} = (still nonvaluable - devalued). Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults. Pairwise comparisons are Bonferroni adjusted for multiple comparisons. ***indicates significance at $p < .001$.

Table 10

Pairwise comparisons on SORT test accuracy between response type by age group controlling for working memory

Trial types compared	Pairwise comparisons					
	Younger adults			Older adults		
	M_{diff}	*p*	95% Confidence Interval	M_{diff}	*p*	95% Confidence Interval
Short condition						
Congruent	-15.84***	<.001	(-24.43, -7.25)	-20.91***	<.001	(-29.50, -12.32)
Incongruent	-23.10***	<.001	(-34.04, -12.16)	-28.65***	<.001	(-39.60, -17.71)
Long condition						
Congruent	-3.80	.215	(-9.88, 2.29)	-6.45*	.038	(-12.54, -.370)
Incongruent	8.29	.115	(-2.10, 18.68)	-24.04***	<.001	(-34.43, -13.65)

Note. N = 50 (*n* = 25 in each age group). Congruent M_{diff} = (still valuable – still nonvaluable). Incongruent M_{diff} = (upvalued - devalued). Short condition was 500ms/800ms for younger/older adults. Long condition was 800ms/1100ms for younger/older adults. Pairwise comparisons are Bonferroni adjusted for multiple comparisons. ***represents a significant difference at $p < .001$. *represents a significant difference at $p < .05$.

Speed of processing correlations

Table 11

Speed of Processing Correlations for Younger Adults

	SORT Test Accuracy			
	Short response window		Long response window	
	Go trials	NoGo trials	Go trials	NoGo trials
Variable				
Symbol Digit Modalities Test	.08	.04	-.02	.22
Speed of Processing discrimination (ms)	-.45*	-.27	.03	-.15

Note. N = 25. Symbol Digit Modalities Test, measured by the number of correct substitutions in 90 seconds, related to psychomotor speed, working memory, and organic cerebral dysfunction. Speed of Processing discrimination task as measured by response reaction time (ms) to visual stimuli presented on screen, related to visual processing speed. Sort test accuracy measured as difference scores on still valuable minus upvalued (Go trials) and accuracy on still nonvaluable minus devalued (NoGo trials)**. ***represents a significant difference at $p < .05$.

Table 12

Speed of Processing Correlations for Older Adults

	SORT Test Accuracy			
	Short response window		Long response window	
Variable	Go trials	NoGo trials	Go trials	NoGo trials
Symbol Digit Modalities Test	-.09	-.29	-.32	-.41*
Speed of Processing discrimination (ms)	-.04	-.03	.11	.16

Note. N = 25. Symbol Digit Modalities Test, measured by the number of correct substitutions in 90 seconds, related to psychomotor speed, working memory, and organic cerebral dysfunction. Speed of Processing discrimination task as measured by response reaction time (ms) to visual stimuli presented on screen, related to visual processing speed. Sort test accuracy measured as difference scores on still valuable minus upvalued (Go trials) and accuracy on still nonvaluable minus devalued (NoGo trials)**. ***represents a significant difference at $p < .05$.

S-O Knowledge correlations

Table 13

S-O Knowledge Correlations for Younger Adults

	SORT Test Accuracy	
	Short response window	Long response window

	Go trials	NoGo trials	Go trials	NoGo trials
Variable				
S-O Knowledge	.19	.15	-.06	.29

Note. N = 25. S-O Knowledge presented as a %, measured at the end of training. Sort test accuracy measured as difference scores on still valuable minus upvalued (Go trials) and accuracy on still nonvaluable minus devalued (NoGo trials)**.**

Table 14

S-O Knowledge Correlations for Older Adults

	SORT Test Accuracy			
	Short response window		Long response window	
	Go trials	NoGo trials	Go trials	NoGo trials
Variable				
S-O Knowledge	-.19	-.36	-.33	-.60**

Note. N = 21. S-O Knowledge presented as a %, measured at the end of training. Sort test accuracy measured as difference scores on still valuable minus upvalued (Go trials) and accuracy on still nonvaluable minus devalued (NoGo trials)**. ****represents a significant difference at $p < .01$.

Mood correlations

Table 15

DASS Correlations for Younger Adults

Variable	SORT Test Condition			
	Short response window		Long response window	
	Go trials	NoGo trials	Go trials	NoGo trials
Depression Anxiety Stress Scale				
Depression subscale (/42)	.20	.14	.34	.47*
Anxiety subscale (/42)	-.01	-.15	.17	.12

Stress subscale (/42)	.10	-.02	-.01	.12

Note. N = 25. Sort test accuracy measured as difference scores on still valuable minus upvalued (Go trials) and accuracy on still nonvaluable minus devalued (NoGo trials)**. ***represents a significant difference at $p < .05$.

Table 16

DASS Correlations for Older Adults

Variable	SORT Test Condition			
	Short response window		Long response window	
	Go trials	NoGo trials	Go trials	NoGo trials
Depression Anxiety Stress Scale				
Depression subscale (/42)	.07	.39	.53**	.25
Anxiety subscale (/42)	.11	.11	.44*	-.01
Stress subscale (/42)	.13	.32	.32	-.04

Note. N = 25. Sort test accuracy measured as difference scores on still valuable minus upvalued (Go trials) and accuracy on still nonvaluable minus devalued (NoGo trials)**. ***represents a significant difference at $p < .05$. ****** represents a significant difference at $p < .01$.