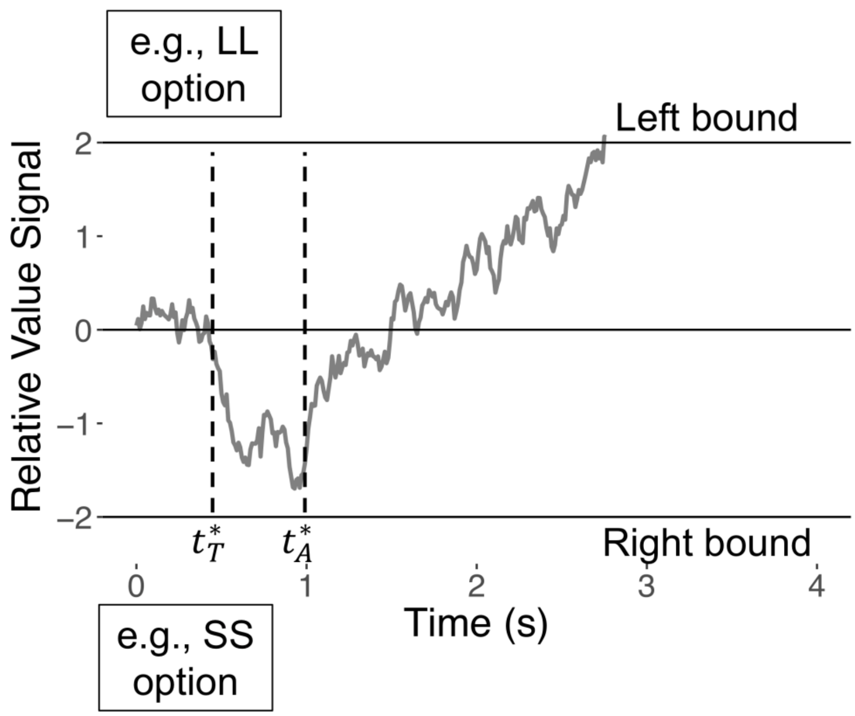
Table of Contents:

**Supplementary Note 1: Attribute-Wise Drift Diffusion Model (DDM).** The drift diffusion model is based on a relative value signal (RVS) that accumulates evidence over time, with a decision being made when the RVS hits a bound. In the simulated run below, the gray line represents the RVS, which starts equidistant between the two options and accumulates evidence via a noisy process. In this example, the time latency ($t^*_T$) occurs first, which drives evidence accumulation toward the smaller, sooner option which has the preferable, shorter time. However, when the amount latency is reached ($t^*_A$), information about amount begins influencing the RVS, shifting the accumulation process toward the larger later option. This shift shows how a faster latency for one attribute can bias choice early in the decision process. In addition, the rate at which the RVS moves, the drift rate, is determined by the value difference between the options as moderated by the subject-specific and attribute-specific drift slopes (i.e., the relative weighting a given subject places on amount or time). Therefore, a shift in direction (i.e., a change in drift rate) could be due to a large difference in the attribute values for the LL compared to SS option and/or a larger drift slope for amount compared to time. Finally, this example shows that moving the bounds inward could have led to a SS choice given the early evidence accumulation toward that SS option.



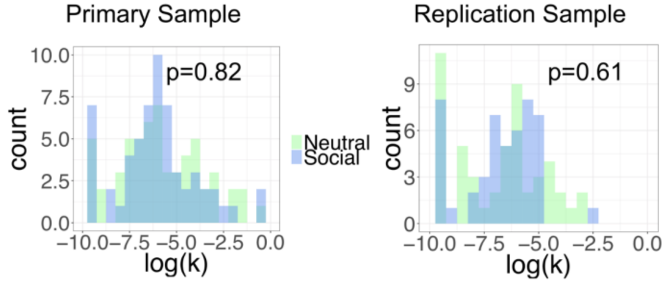

Supplementary Figure 1: Attribute-wise Drift-Diffusion Model (DDM) example. This figure shows the overall form of the drift diffusion model with an example drawn from a simulated trial (e.g., a choice between the LL option on the left side of the screen, and the SS option on the right side of the screen). The gray line represents the relative value signal (RVS), which represents the process of evidence accumulation over time; when that RVS hits a bound, a decision is made. The RVS incorporates both the drift slopes for amount and time and the value differences between the attributes. The horizontal lines represent the decision bounds, and the vertical lines represent the time and amount latencies.

**Supplementary Note 2: Analysis Workflow.** Our approach to analysis included exploratory as well as confirmatory analyses. We initially analyzed our primary sample to determine key results of interest, and then we evaluated whether each of those results replicated in a second independent sample. We first step fit behavior using the canonical hyperbolic model. This allowed us to confirm that fitted discount rates described choice through two manipulation checks: (1) participants tended to choose options with higher estimated subjective value, and (2) choices involving options with similar subjective values generated longer response times than those involving options with very different values (Supplementary Figure 3). We also examined the relationship between looking and choice to ensure that our eye gaze data were related to choice (Supplementary Figures 4 and 5).

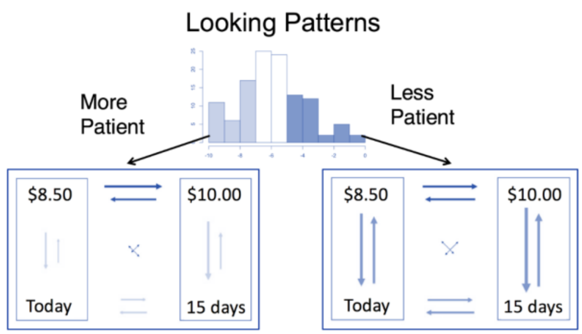
Our first analysis of interest in our primary sample was look for differences across our social and neutral conditions. There were no differences in discount rate in either sample across condition, so we collapsed across this measure for all subsequent analyses. Second, we wanted to test for a relationship between survey-measured impulsivity (ABIS[6]) and intertemporal impulsivity. In our primary sample, we found the strongest relationship between the non-planning subscale and intertemporal impatience, but this did not replicate, and no subscales were significantly related to choice in our replication sample. Therefore, we did not use the ABIS for any further analyses.

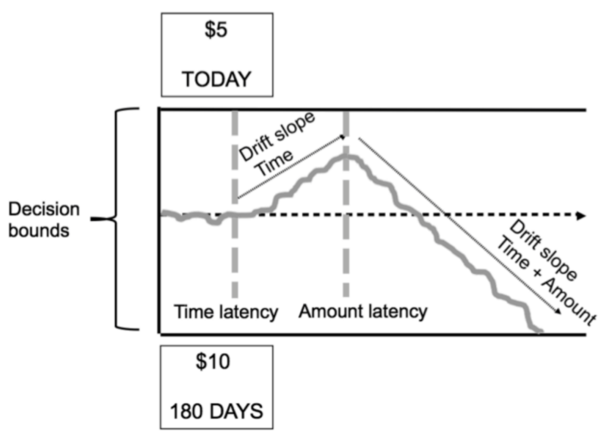
Third, we analyzed our eye tracking data in our primary sample. Because all indices measured (Option Index, Attribute Index, Payne Index) showed significant relationships with the discount rate in our primary sample, we included them in our replication analyses.

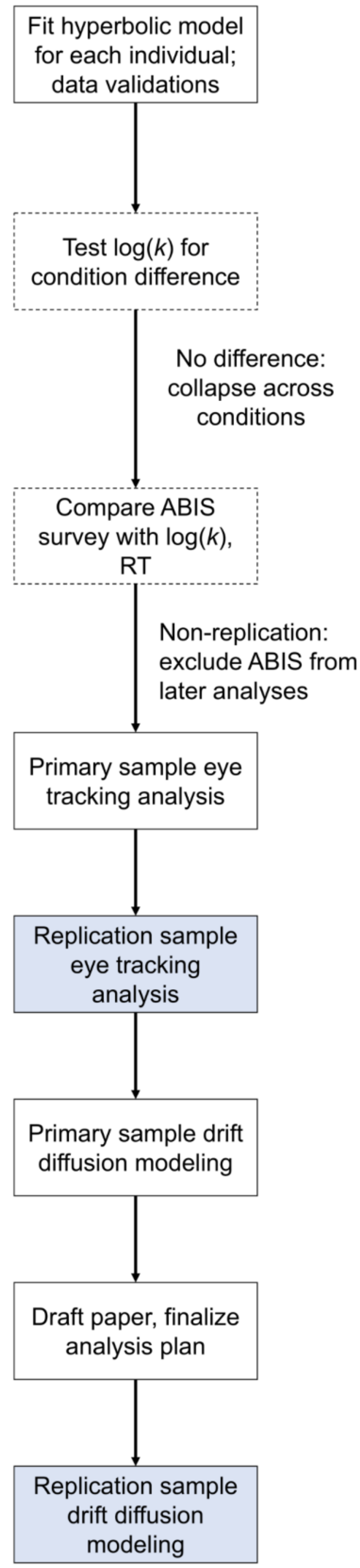
Next, we tested drift diffusion modeling as an attribute-wise model in our primary sample. Before analysis of our replication sample, we drafted our manuscript and finalized our planned analyses. In writing the manuscript, we decided to compare attribute-wise and option-wise models, given that the typical form of intertemporal choice models (e.g., hyperbolic models) assumes an option-wise process. Finally, after our DDM results were finalized in the primary sample, we ran the both attribute-wise and option-wise analyses in our replication sample.




| Correlations (ABIS subscales) | Primary Sample | Replication Sample |
|---|---|---|
| log(k), Nonplanning | r = 0.35*** | r = 0.07 |
| log(k), Attention | r = 0.17 | r = -0.02 |
| log(k), Motor | r = 0.25** | r = 0.05 |
| RT, Motor | r = -0.06 | r = -0.11 |

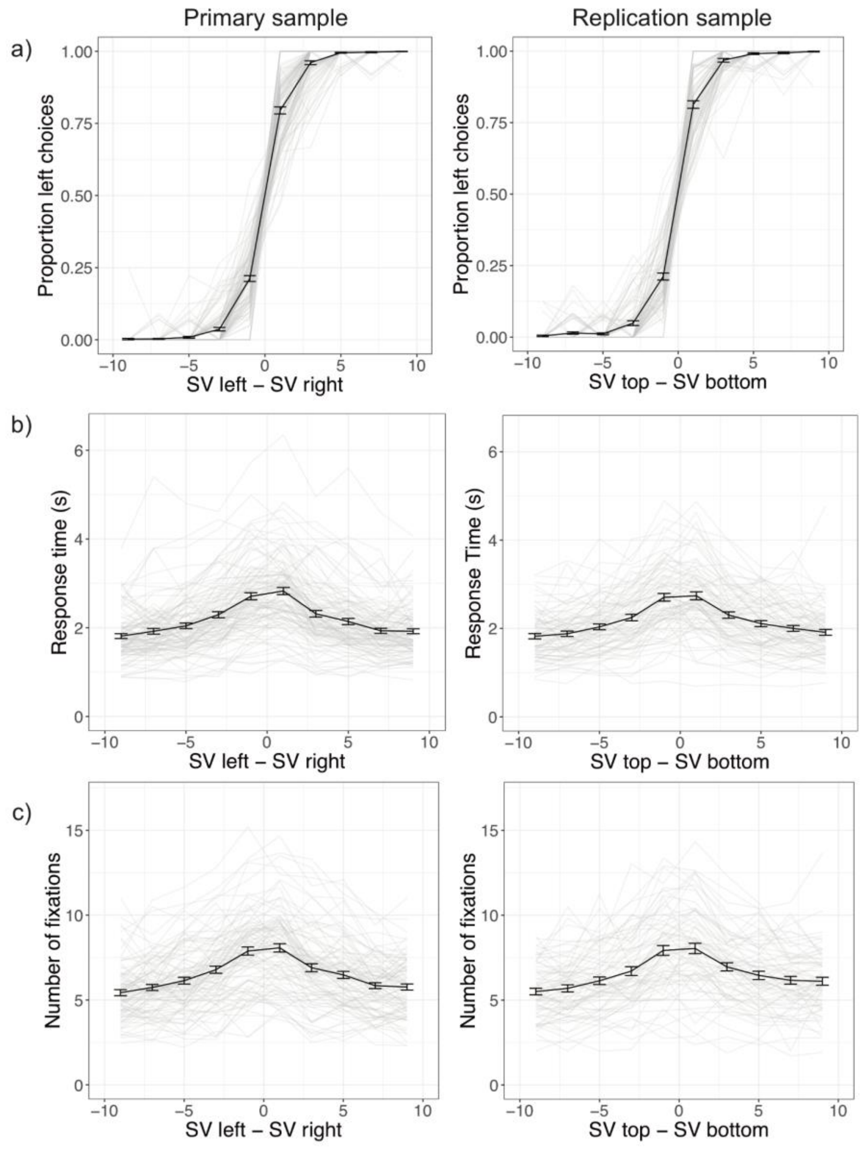


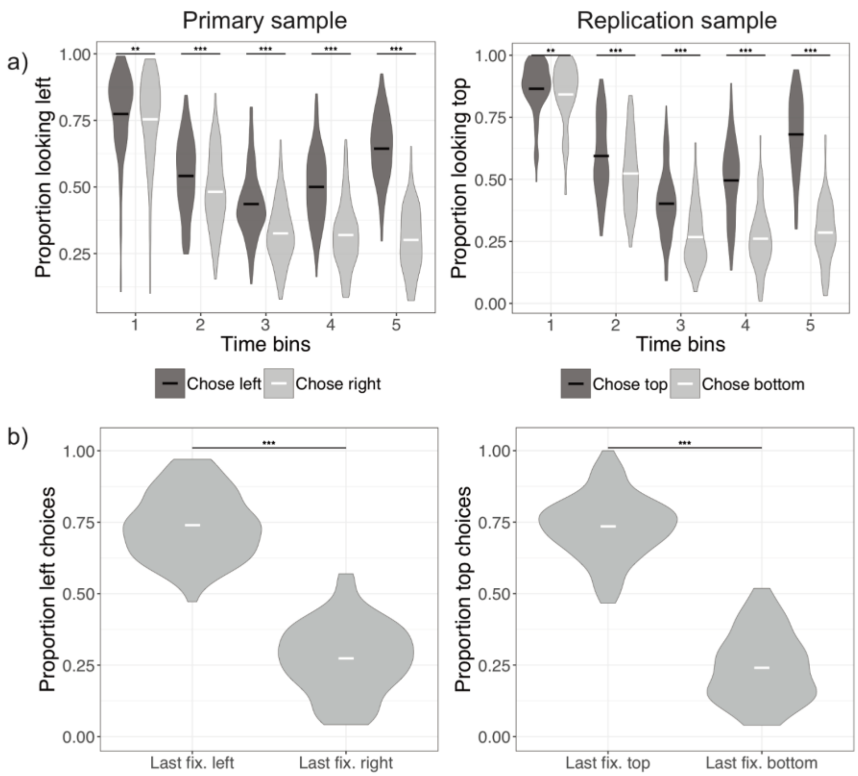

Supplementary Figure 2: Analysis workflow. Rectangles with dotted outlines represent analyses that were unsuccessful or did not replicate. Shaded rectangles represent replication analyses. *p<0.05, **p<0.01, ***p<0.001.

**Supplementary Note 3: Subjective value (SV) predicts choices and response times.** We examined how trial-to-trial variation in subjective value (SV) – as fit by the hyperbolic discounting model – influenced choices, response times, and gaze fixations. As expected, choices followed a logistic shape, such that the proportion of choices to the higher-SV option increased with increasing relative SV. Additionally, trials that had relatively greater differences in SV were associated with faster response times and fewer fixations, while trials where SV was more matched between the options had longer response times and a higher number of fixations. All effects observed in the first sample were replicated in the second sample. We conclude from these manipulation checks that our task had appropriate psychometric properties. While the hyperbolic model may not be the true choice generating process, it explained participants' choices and response times well, and we use it as a comparison to our multi-attribute DDM.



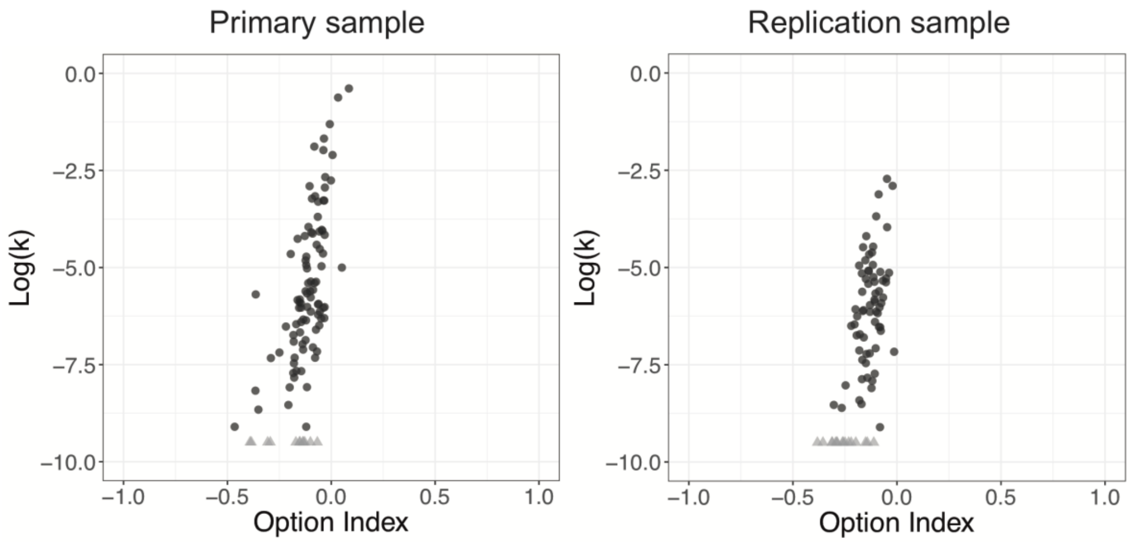

Supplementary Figure 3. Subjective value (SV) corresponds to (a) choices, (b) response time, and (c) eye tracking gaze fixations. Panels (a) and (b) exclude participants not able to be fit to a single discount rate, leaving primary sample N = 105 and replication sample N = 79. Panel (c) excludes participants not able to be fit to a single discount rate or who had insufficient eye-tracking data for analysis, leaving primary sample N = 93 and replication sample N = 68. Light gray lines represent individual subjects; darker lines represent group mean values; error bars are SEM.

**Supplementary Note 4: Eye-tracking data predicts choices.** We examined whether eye-tracking data predicted variation across trials in choices and variation across participants in patience. We partitioned every trial into five time bins, and then measured total looking time to each choice option within each bin. Participants showed a strong initial fixation bias toward the left option (in our primary sample) or the top option (in our replication sample), which likely reflects cultural biases in attention to information positioned at the top left of a display[1]. However, beginning with the second time bin, there was a divergence such that participants increasingly directed their gaze toward the chosen option (Two-sided Welch's paired t-test for second time bin, primary sample $t(104) = 5.25$, $p < 0.001$, Cohen's $d = 0.51$, 95% CI = 0.24 – 0.79; replication sample $t(84) = 6.16$, $p < 0.001$, Cohen's $d = 0.67$, 95% CI = 0.36 – 0.98). The location of the final fixation was a strong predictor of choice; participants chose the last-fixated option on approximately 75% of trials (Two-sided Welch's paired t-test, primary sample $t(104) = 21.48$, $p < 0.001$, Cohen's $d = 2.10$, 95% CI = 1.76 – 2.44; replication sample: $t(84) = 22.92$, $p < 0.001$, Cohen's $d = 2.49$, 95% CI = 2.08 – 2.89). While preferences may drive attention to more highly valued choices, this result could also reflect a gaze cascade effect wherein even without a difference in preference, attention to an option increases the accumulation of evidence for choosing that option, making it more likely to be chosen[2–4]. Again, all effects were fully replicated in both samples.



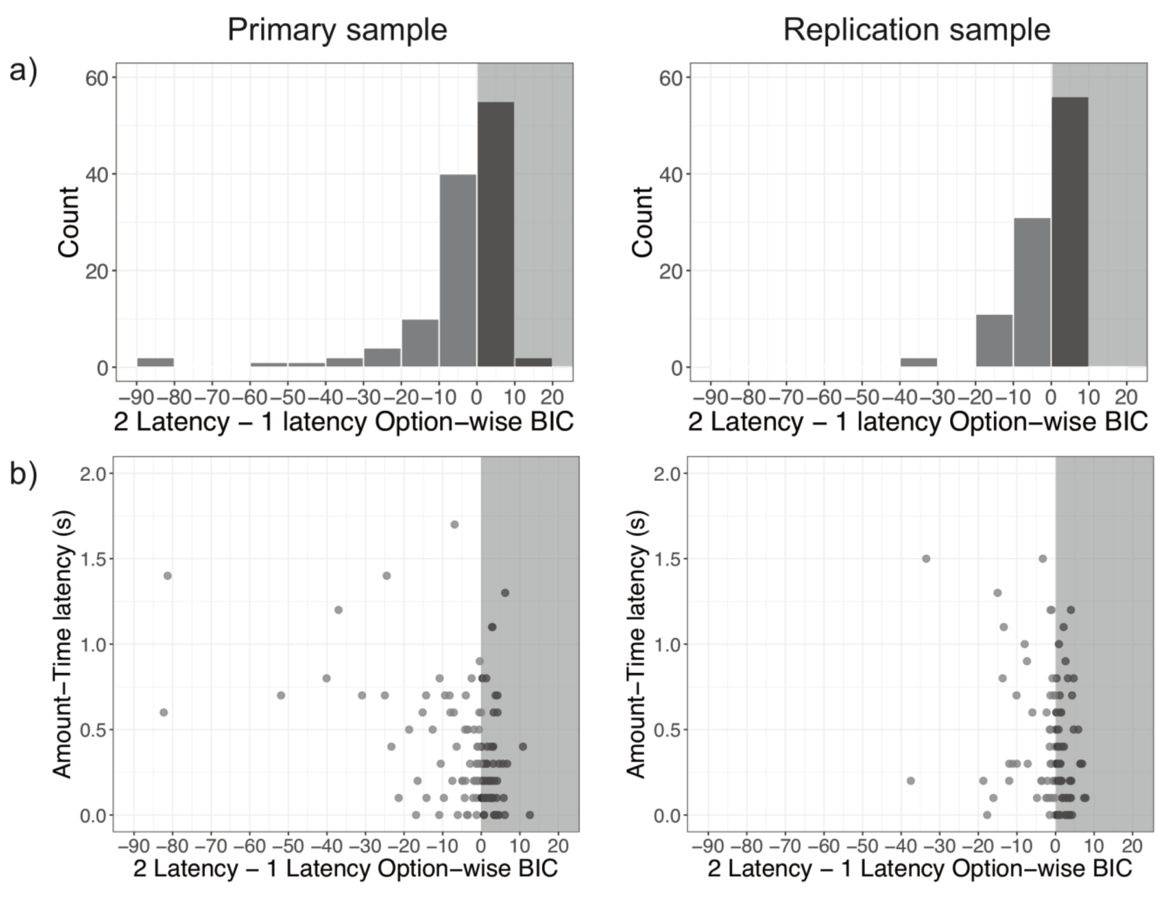

Supplementary Figure 4. Validation of eye-tracking data as a predictor of choice. We examined the relationship between gaze location and eventual choices in all participants with sufficient eye-tracking data (primary sample N = 105; replication sample N = 85 participants). a) We first split trials into five equal time bins according to whether participants chose the left or right option (top or bottom, in replication sample). Participants' eye gaze began to predict their eventual choice by the second time bin in both samples. b) Next, we split trials according to whether the final fixation was to the left or right option (top or bottom, in replication sample). Final fixation location was a strong predictor of eventual choice. Violin plots show data density, and horizontal lines illustrate means. *p < 0.05, **p < 0.01, ***p < 0.001.

**Supplementary Note 5: Option Index values correlate with intertemporal patience.** While the above results linked eye gaze to specific choices, there could also be trait effects such that looking time predicts overall patience across trials. We found a strong positive correlation between participants' *Option Index* and their fitted *k* values, such that those participants whose gaze was most biased toward the LL option exhibited the greatest patience in their choices (two-sided Pearson's product-moment correlation: primary sample $t(91) = 8.22$, $p < 0.001$, $r = 0.65$, 95% CI = 0.52 – 0.76); the same effect was observed in our replication sample ($t(66) = 4.08$, $p < 0.001$, $r = 0.45$, 95% CI = 0.24 – 0.62). This supported the validity of our eye tracking measures by showing that participants tended to look preferentially at options they chose and suggested that individual differences in intertemporal choice reflect an interaction between preference and gaze.



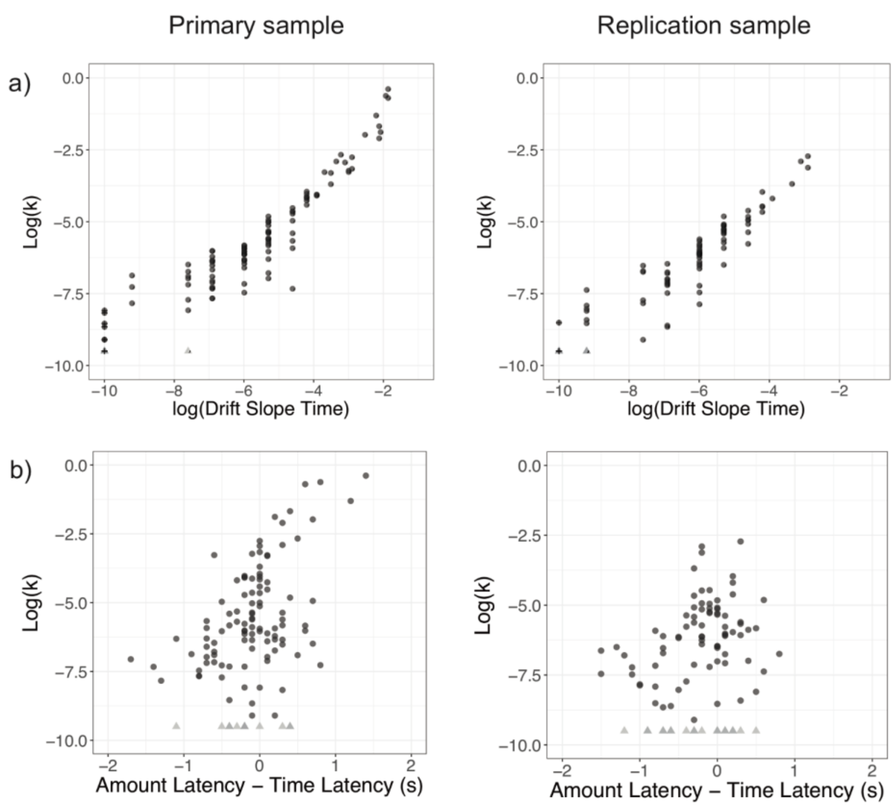

Supplementary Figure 5. Option Index correlates with discount rate. Participants who were not able to be fit to a single discount rate are excluded from statistics. Participants with all patient choices are displayed in light gray triangles at -9.5 on the y-axis for illustration. Primary sample N=105 displayed, 93 used for analysis, replication sample N=84 displayed, 68 used for analysis. The Option Index indicates whether participants looked proportionally more at the SS option (index > 0) or LL option (index < 0).

**Supplementary Note 6: Using separate latencies for amount and time.** We compared option-wise models with a single latency versus models with separate latencies for amount and time. We ran this supplementary analysis because amount and time are dependent on each other in the option-wise model, so splitting up amount and time would not necessarily improve the model fit. We found that overall, there is not a significant difference between which model fits the data better (two-sided exact binomial tests: primary sample 59/117, $p = 1$, 95% CI = 0.41 – 0.60; replication sample 43/100, $p = 0.19$, 95% CI = 0.33 – 0.53). However, we also found that the larger the difference in latencies in the two-latency model, the better that model fit a given subject in our primary sample (two-sided Kendall's rank correlation tau: primary sample $z(115) = -3.22$, $p = 0.0013$, tau = -0.21, 95% CI = -0.35 – -0.08) with a marginally better fit in our replication sample $z(98) = -1.96$, $p = 0.0502$, tau = -0.14, 95% CI = -0.28 – -0.001). Moreover, a subset of subjects had markedly better fits when using the two-latency model. Therefore, we used the two-latency option-wise model in the main paper and show comparisons between the models below.



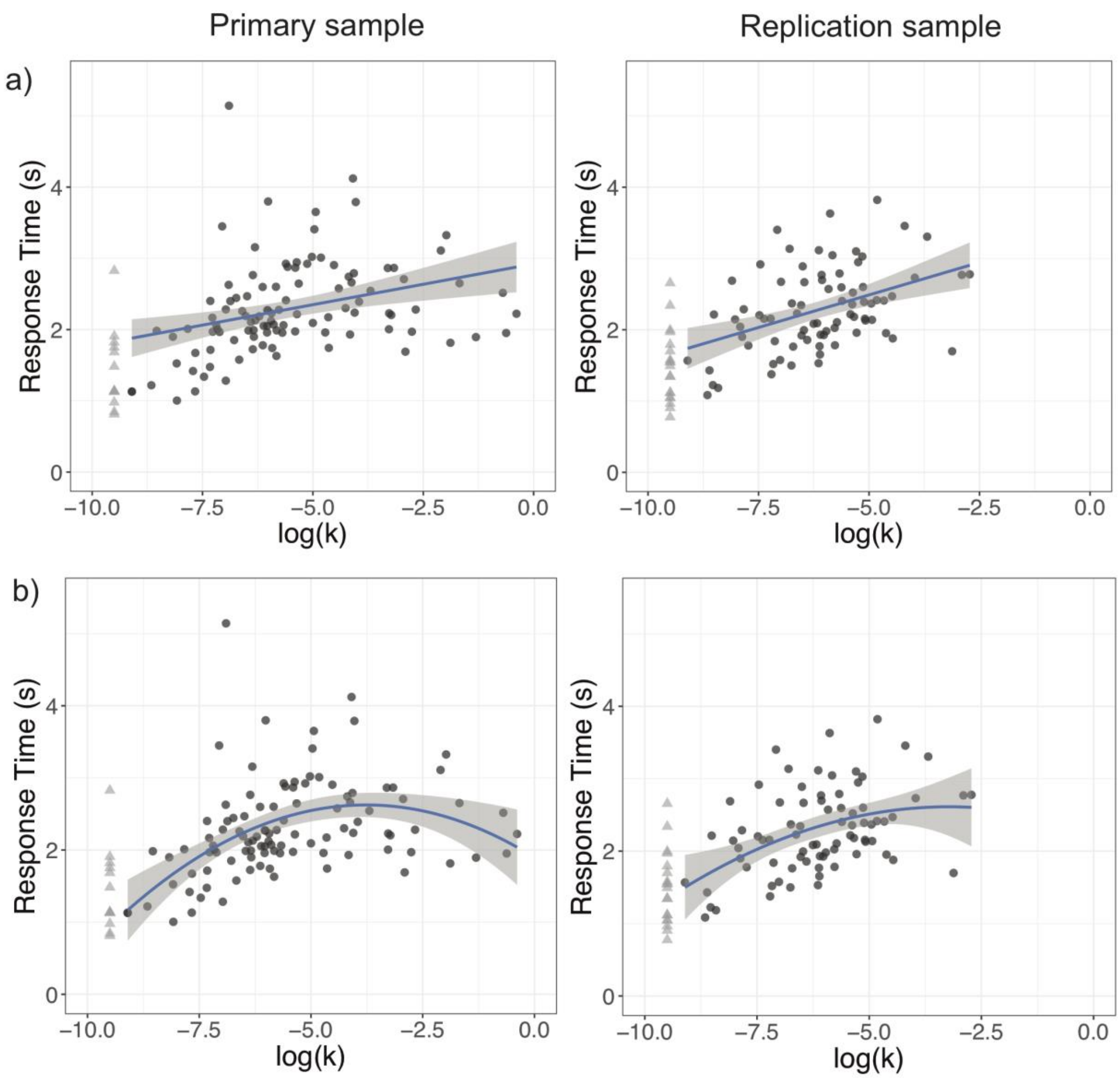

Supplementary Figure 6: Option-wise modeling comparison of one vs. two latencies. Primary sample N = 117, replication sample N = 100. a) A histogram of the difference in BIC for each participant across models. b) The difference in BIC is compared with the difference in the absolute value of amount and time latencies in the two-latency option-wise model. Gray shading indicates values better fit by the single-latency model, whereas no shading indicates values better fit by the two-latency model (lower BIC values indicate better fit).

**Supplementary Note 7: Option-wise DDM: Relationships to intertemporal patience.** There was minimal variation in amount drift slopes compared to time drift slopes in our option-wise DDM whereas there was similar variation across both drift slopes in our attribute-wise model (see Supplementary Table 2 for the relative contributions of amount and time). Here we focused on the relationship between the discount rate k and the time drift slope. Because the option-wise time drift slope was not normally distributed, we used the natural-log transformed drift slope, whereas this was not necessary in the attribute-wise model. Time drift slope and discount rate were correlated in the option-wise model such that those with a lower drift slope for time were more patient and those with a higher drift slope for time were less patient (two-sided Pearson's product-moment correlations: primary sample $t(97) = 22.15$, $p < 0.001$, $r = 0.91$, 95% CI = 0.87 – 0.94; replication sample $t(76) = 16.90$, $p < 0.001$, $r = 0.89$, 95% CI = 0.83 – 0.93). Furthermore, we found similar results for latency as in our attribute-wise model such that a temporal advantage for amount relates to more patient choices and a temporal advantage for time relates to less patient choices (two-sided Pearson's product-moment correlations: primary sample $t(103) = 5.65$, $p < 0.001$, $r = 0.49$, 95% CI = 0.33 – 0.62; replication sample $t(77) = 2.93$, $p = 0.0044$, $r = 0.32$, 95% CI = 0.10 – 0.50). Finally, there is a small correlation between discount rate and decision-bounds in the primary sample, (two-sided Pearson's product-moment correlation: $t(103) = 2.36$, $p = 0.02$, $r = 0.23$, 95% CI = 0.04 – 0.40) but no correlation in the replication sample ($t(77) = 1.31$, $p = 0.19$, $r = 0.15$, 95% CI = -0.08 – 0.36).



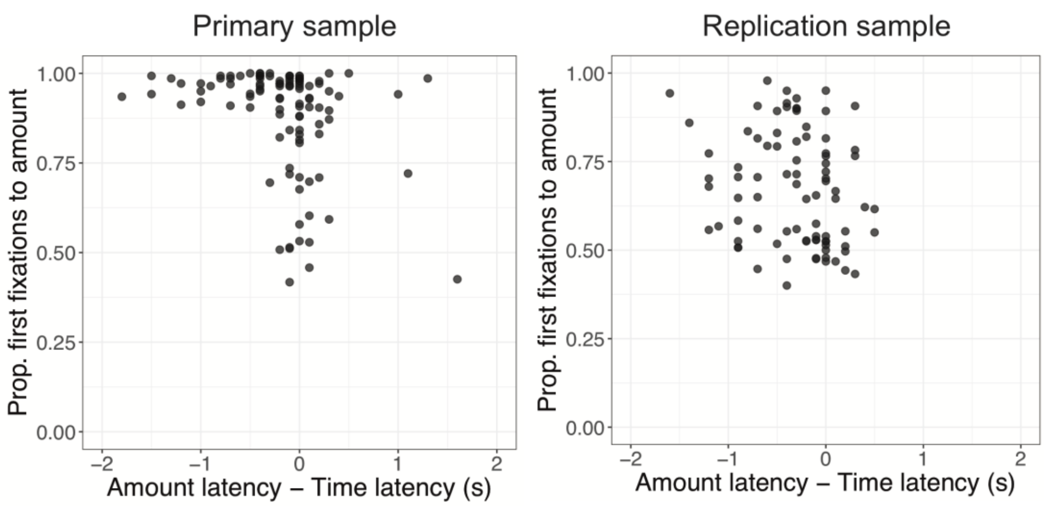

Supplementary Figure 7. Option-wise DDM modeling: relationship to intertemporal patience. Primary sample N = 117 displayed, replication sample N = 98 displayed. a) Participants with time drift rates of 0 are displayed as "+" at -10 on the x-axis for illustration in and excluded from statistics (primary sample N = 99, replication sample N = 78). b) corresponds to figure 3c in the main paper (primary sample N = 105, replication sample N = 79). Participants unable to be fit to a single discount rate are excluded from statistics; those with all patient choices are displayed in light gray triangles at -9.5 on the y-axis for illustration.

**Supplementary Note 8: Relationship between response time and discount rate.** In our primary sample, we found that a quadratic model explained more than twice the variance as a linear model and the coefficient on the quadratic term was significant (Linear model: $F(1,103)=12.2$, $p < 0.001$, Adj. $R^2 = 0.10$; quadratic model: $F(2,102) = 15.42$, $p < 0.001$, Adj. $R^2 = 0.22$, with quadratic coefficient: $b = -2.40$ ($SE = 0.59$), $p < 0.001$), such that both very patient and impatient people make faster choices than those in the middle. However, in our replication sample, there was a minimal change in adjusted R-squared and the coefficient on the quadratic term was not significant (Linear model: $F(1,77) = 17.32$, $p < 0.001$, Adj. $R^2 = 0.17$; quadratic model: $F(2,76) = 9.70$, $p < 0.001$, Adj. $R^2 = 0.18$, with quadratic coefficient $b = -0.72$ ($SE = 0.52$), $p = 0.17$), potentially because of the presence of fewer very impatient people in the replication sample.



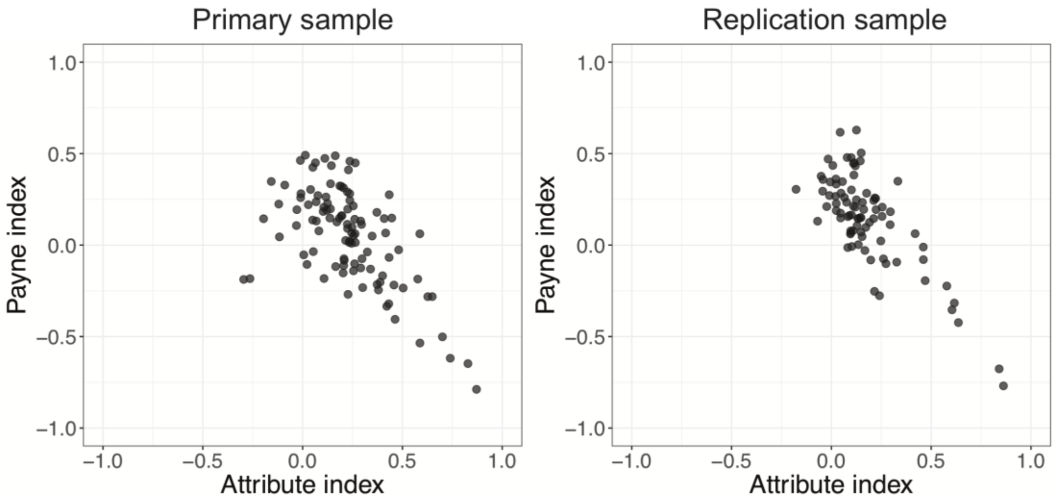

Supplementary Figure 8. Relationship between response time and discount rate. a) Linear models of response time on discount rate. b) Quadratic models of response time on discount rate. Participants unable to be fit to a single discount rate are excluded from statistics, leaving final samples of N = 105 (primary) and N = 79 (replication). Participants with all patient choices are displayed in light gray triangles at -9.5 on the x-axis for illustration.

**Supplementary Note 9: Relationship between initial fixations and attribute latency.** First fixation location was negatively correlated with difference in attribute latency in our primary sample (two-sided Kendall's rank correlation tau: z(103) =-3.40 , p < 0.001, tau = -0.23, 95% CIs = -0.35– -0.12) and in our replication sample (z(83) = -2.16, p = 0.031, tau = -0.16, 95% CIs = -0.30 – -0.03). Those who incorporate time information more quickly into their information gathering process in the DDM also are more likely to look first at amount information.



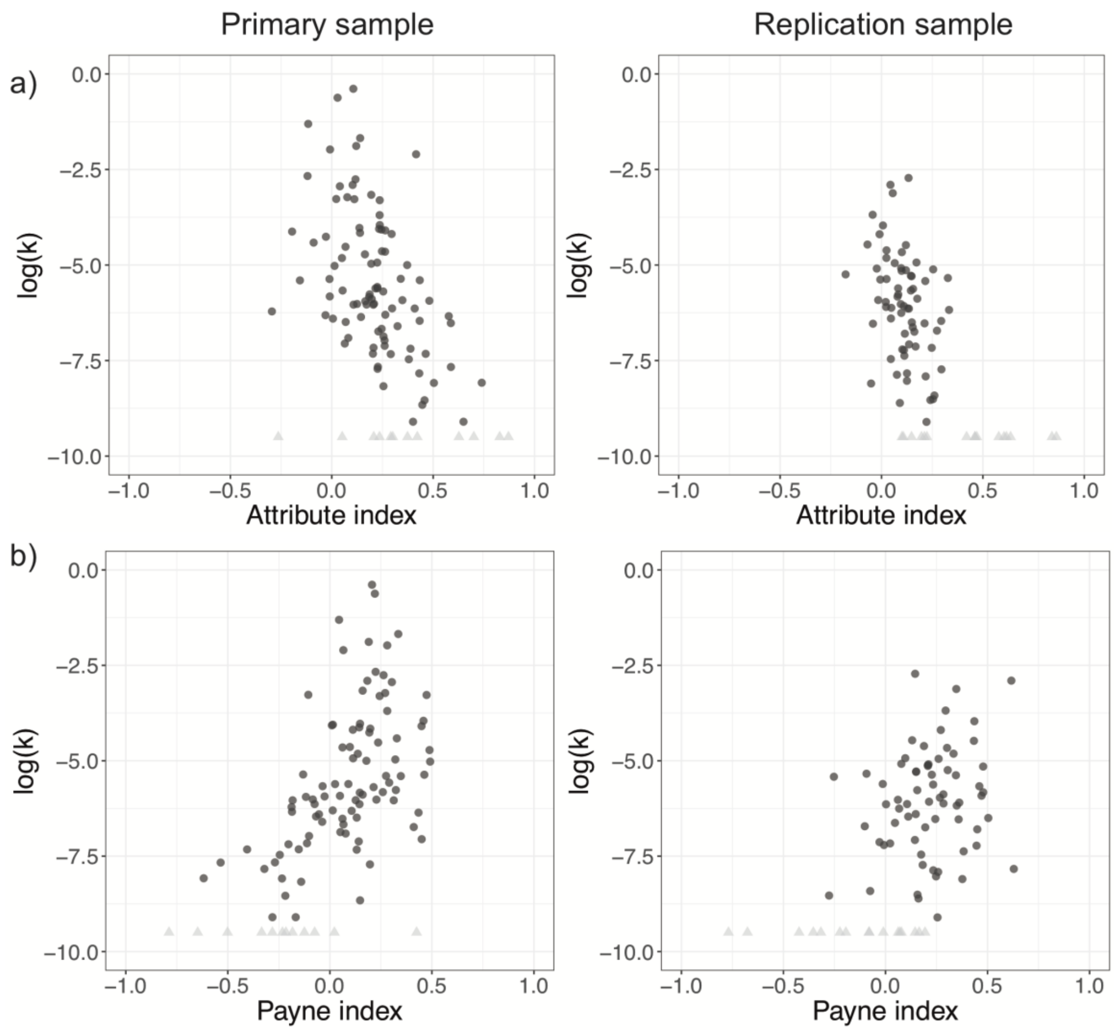

Supplementary Figure 9. Difference in attribute latency correlates with first fixation location. The proportion of fixations to amount compared with time information correlates negatively with the difference in amount and time latency. Primary sample N=105, replication sample = 85 excluding participants with insufficient eye tracking.

**Supplementary Note 10: Relationships between Payne and Attribute indices.** The Payne Index[5] and Attribute Index were negatively correlated such that those who make more attribute-wise comparisons tended to look more at amounts, whereas those who make more option-wise comparisons tended to look more evenly between amounts and times (two-sided Pearson's product-moment correlations: primary sample $t(103) = -7.80$, $p < 0.001$, $r = -0.61$, 95% CI = -0.72 – -0.47; replication sample $t(83) = -11.04$, $p < 0.001$, $r = -0.77$, 95% CI = -0.85 – -0.67).




Supplementary Figure 10. The Payne Index correlates with the Attribute Index. Primary sample N=105, replication sample N=85. Participants with insufficient eye tracking data were excluded. The Attribute Index measures relative looking at amounts (index>0) versus times (index<0). The Payne Index measures relative looking between options (index>0) or between attributes (index<0).

**Supplementary Note 11: Relationships between Attribute/Payne indices and choice behavior.** The Attribute Index correlated negatively with discount rate, such that more patient participants looked more at amounts and less patient participants looked more at times (two-sided Pearson's product-moment correlations: primary sample $t(91) = -5.51$, $p < 0.001$, $r = -0.50$, 95% CI = -0.64 – -0.33; replication sample $t(66) = -3.29$, $p = 0.0016$, $r = -0.38$, 95% CI = -0.56 – -0.15). The Payne Index was significantly correlated with discount rate in the primary sample, but this relationship was not significant in the replication sample (two-sided Pearson's product-moment correlations: primary sample $t(91) = 5.76$, $p < 0.001$, $r = 0.52$, 95% CI = 0.35 – 0.65; replication sample $t(66) = 1.70$, $p = 0.09$, $r = 0.20$, 95% CI = -0.04 – 0.42). This may be due to any of several factors, including reduced variation in discount rate distribution, the change in orientation of information, or a lack of robustness of the result.

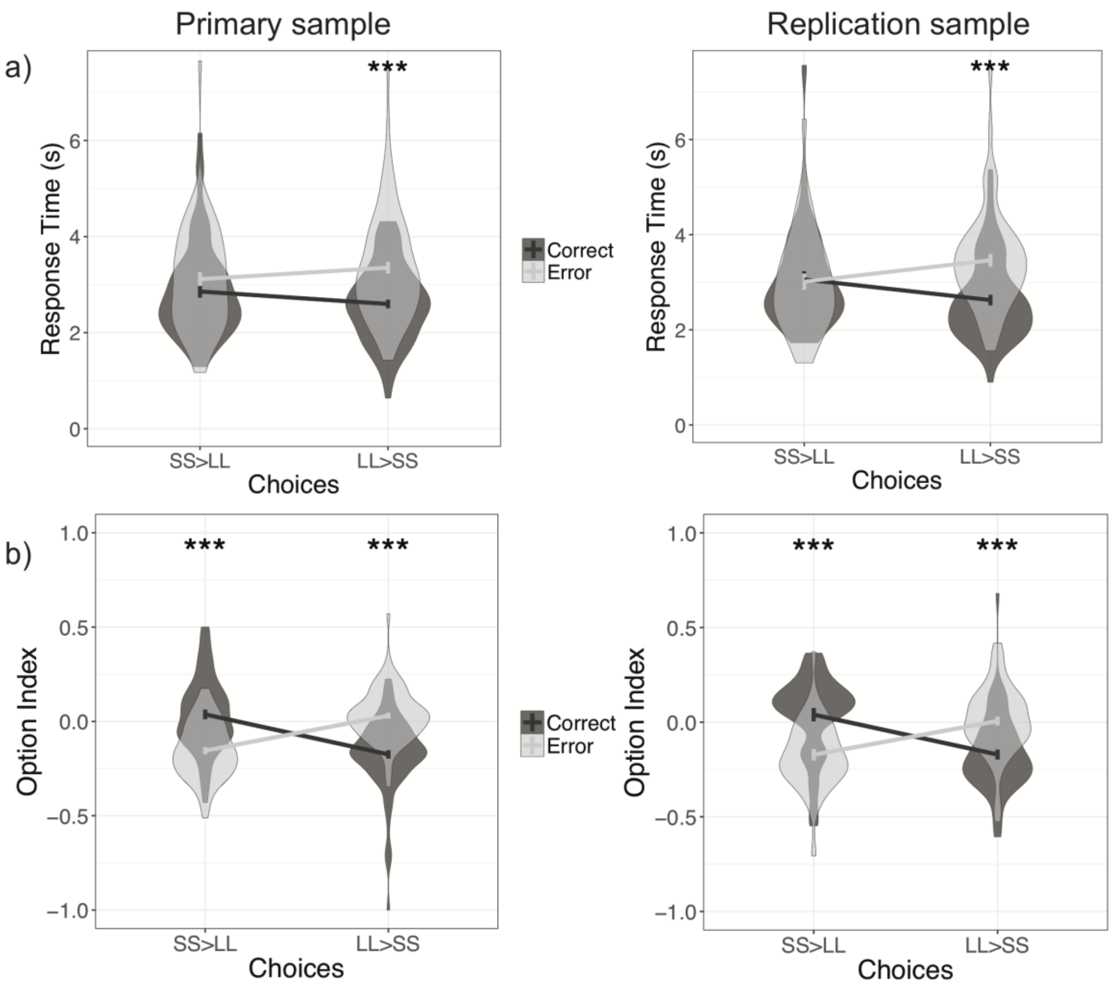



Supplementary Figure 11: Relationship between gaze indices and discount rate. Primary sample N=105 displayed, 93 used for analysis, replication sample N=84 displayed, 68 used for analysis. Participants with insufficient eye tracking data and those unable to be fit to a single discount rate were excluded from statistics. a) The Attribute Index measures relative looking at amounts (index>0) versus times (index<0). b) The Payne Index measures relative looking between options (index>0) or between attributes (index<0). Participants with all patient choices are displayed in light gray at -9.5 on the y-axis for illustration.

**Supplementary Note 12: Analyses of Error Trials.** We examined inter-trial differences in information processing patterns according to choice type (LL or SS) and whether the choices were correct or errors. We defined "error" trials as those on which the lower SV option was chosen as determined by a participant's individual discount rate, *k*. For each error trial, we matched a "correct" trial that was closest in SV difference and on which the participant chose the option with higher SV. After excluding subjects who did not have at least three matched trials across conditions, we had the following sample sizes:

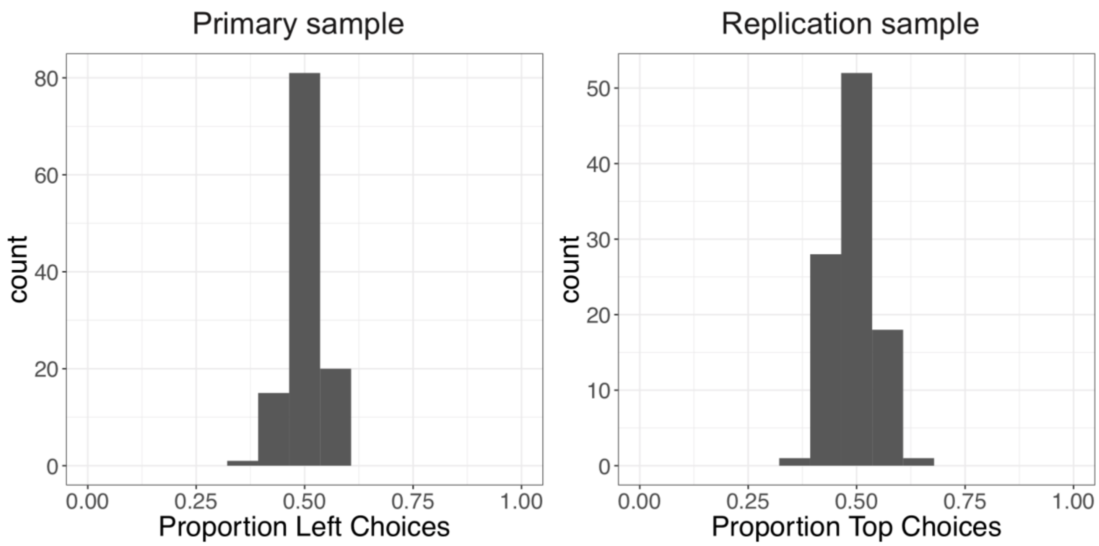
- Response Time, Primary Sample: N=58 for SS>LL, N=79 for LL>SS.
- Response Time, Replication Sample: N=38 for SS>LL, N=60 for LL>SS.
- Option Index, Primary Sample: N=58 for SS>LL, N=79 for LL>SS.
- Option Index, Replication Sample: N=37 for SS>LL, N=60 for LL>SS.

We found a significantly higher response time for errors compared to correct responses on trials in which the LL option had a higher subjective value (two-sided Welch's paired t-tests: primary sample $t(78) = 7.02$, $p < 0.001$, Cohen's $d = 0.50$, 95% CIs = 0.15 – 0.85; replication sample $t(59) = 5.45$, $p < 0.001$, Cohen's $d = 0.73$, 95% CIs = 0.33 – 1.13). Therefore, when people incorrectly chose a SS option over the LL option, they were slower than when correctly choosing the LL option. We also found a difference in the Option Index such that people tended to look more at the option they choose. Therefore, when people chose a LL option in error, they looked more at the LL option on that trial than when correctly choosing the SS option (two-sided Welch's paired t-tests: primary sample $t(57) = -6.17$, $p < 0.001$, Cohen's $d = -0.61$, 95% CIs = -1.09 – -0.13; replication sample $t(36) = -4.71$, $p < 0.001$, Cohen's $d = -1.15$, 95% CIs = -1.81 – -0.49) and the same pattern holds for SS errors compared to correctly choosing the LL option (two-sided Welch's paired t-tests: primary sample $t(78) = 6.74$, $p < 0.001$, Cohen's $d = 0.66$, 95% CIs = 0.31 – 1.02; replication sample $t(59) = 5.63$, $p < 0.001$, Cohen's $d = 0.62$, 95% CIs = 0.23 – 1.02). There were no differences in Attribute and Payne Index scores across error and correct trials.




Supplementary Figure 12. Error analysis: inter-trial differences in information processing. a) On trials where the SV of the LL option was greater than that of the SS option (LL>SS), response times were significantly slower on error trials (i.e., choices of the SS option) than on correct trials. b) We observed an interaction in the Option Index, such that both sorts of errors were associated with increased gaze time toward the subsequently chosen (and lower SV) option. Participants with insufficient eye tracking, those unable to be fit to a single discount rate, and those with insufficient errors were excluded from this analysis. Violin plots show data density, and error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001.

**Supplementary Note 13: Control analysis: Option Position.** We explored the distribution of left and top choices for the primary and replication samples (on a per-subject basis). While there was variation in whether individual subjects predominantly chose left vs. right options (or top vs. bottom in the replication sample), both distributions were centered around 0.5. This indicates that directional biases had minimal if any effect on our results. Given this finding – and the lack of a theoretical reason to expect that spatial bias would confound any analyses – our modeling assumed that there were no significant biases associated with spatial position.




Supplementary Figure 13: Choices were not influenced by option positioning. Histograms of choices by direction: left/right in primary sample, top/bottom in replication sample. Primary sample N=117, replication sample N=100.

**Supplementary Methods: DDM Fitting Procedure.** We initially used coarse grids for each DDM parameter such that very few subjects fit on the ends of the range (<5%). In the subsequent finer grid, drift slopes were fit to the nearest 0.005, latencies were fit to the nearest 100 ms, and bounds were fit to the nearest 0.25. For drift slopes, we fit a range of 2 steps in both directions from the coarser grid. For drift slopes below 0.005, we used a slightly finer grid shown below including a minimum of 0. For latency, we fit a range of 1s in both direction with a minimum latency of 100 ms and a maximum latency of average response time. For bounds, we fit 0.5 in both directions. The full range of values used is shown below, but each subject's finer grid was fit by a subset of these values.

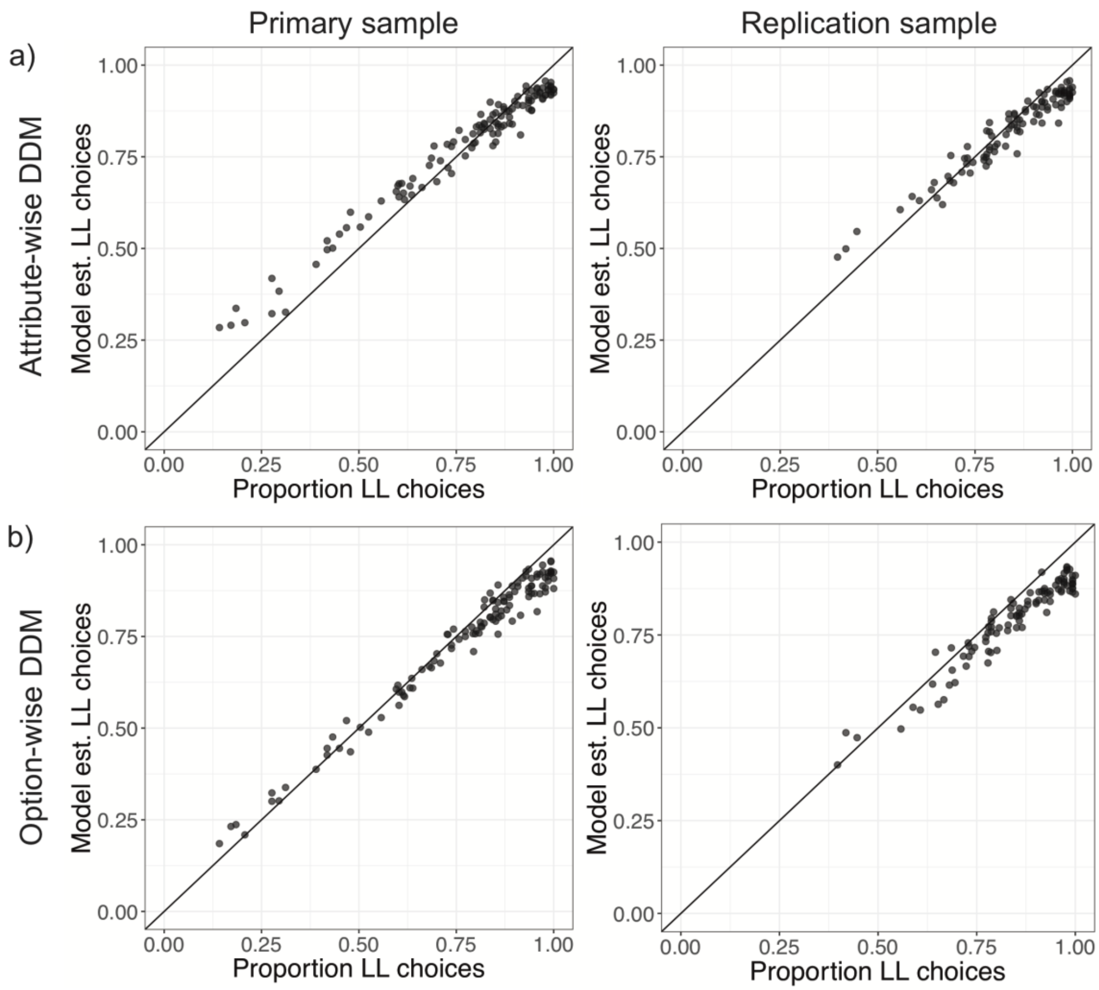
Coarse grid values:

- Attribute-wise model drift slopes: [.0001 .0079 .0156 … .0700]
- Option-wise model drift slopes: [.0001 .0003 .0006 .0013 .0028 .0060 .0131 .0286 .0622 .1354]
- Latency A, Latency T (ms): 5 equally-spaced values from 20 to mean RT
- Bounds: [1 1.5 2 2.5 3]

Finer grid values:

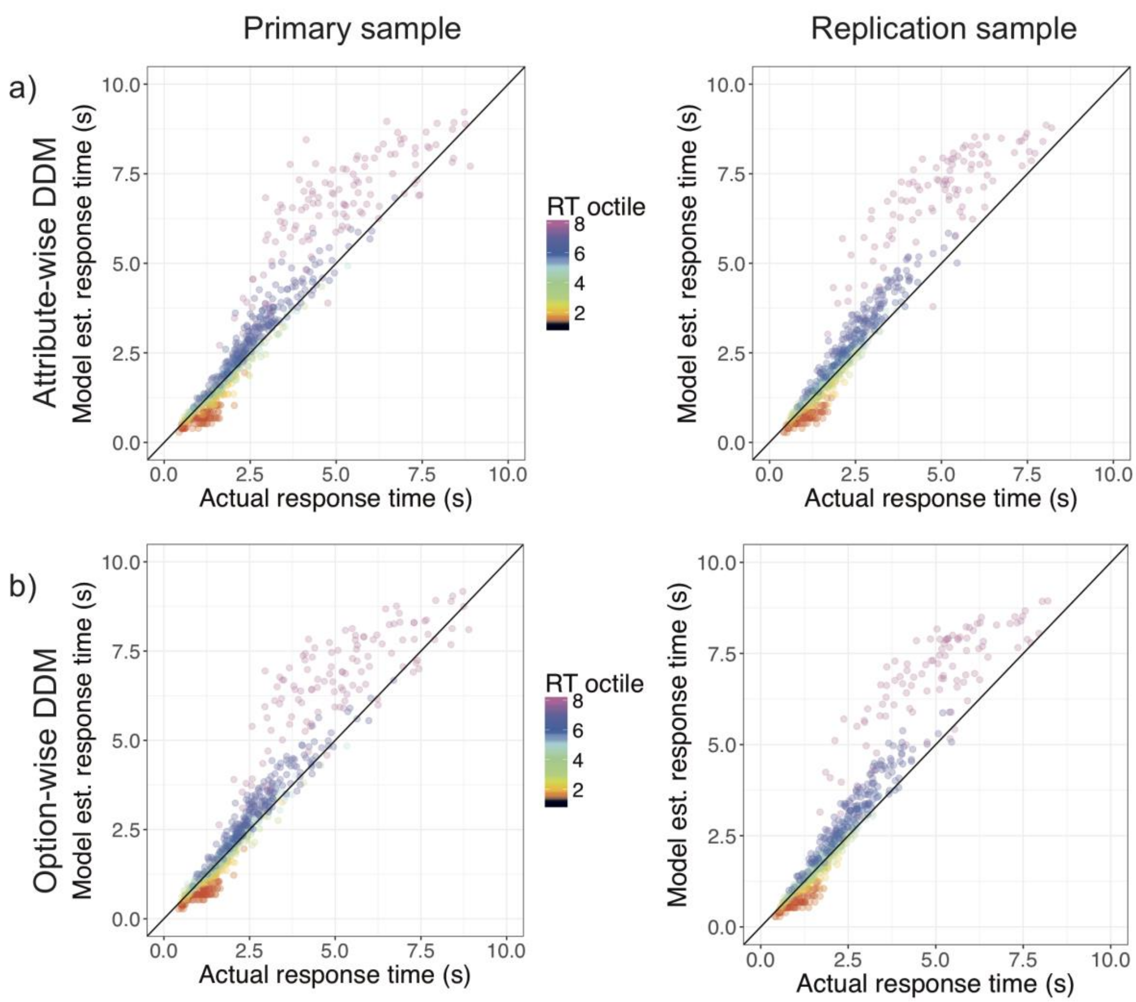
- Attribute-wise model drift slopes: [0 0.0001 0.001 0.005 … 0.085]
- Option-wise model drift slopes: [0 0.0001 0.0005 0.001 0.0025 0.005 … 0.155]
- Latencies: 100 ms increments, from 100 to max of mean RT
- Bounds: [0.5 0.75 1 … 3.5]

**Supplementary Note 14: Accuracy of DDM predictions of LL choices.** We compared the proportion of delayed (LL) choices for an individual to the model predicted proportion for the attribute-wise and option-wise models. For both the attribute-wise model and the option-wise model, the correlation between actual and model-predicted LL choices very high in the primary sample (two-sided Pearson's product-moment correlations: attribute-wise $t(115) = 62.65$, $p < 0.001$, $r = 0.99$, 95% CI = 0.98 – 0.99; option-wise $t(115) = 68.02$, $p < 0.001$, $r = 0.99$, 95% CI = 0.98 – 0.99) and the replication sample (attribute-wise $t(98) = 37.61$, $p < 0.001$, $r = 0.97$, 95% CI = 0.95 – 0.98; option-wise $t(98) = 35.31$, $p < 0.001$, $r = 0.96$, 95% CI = 0.95 – 0.97).



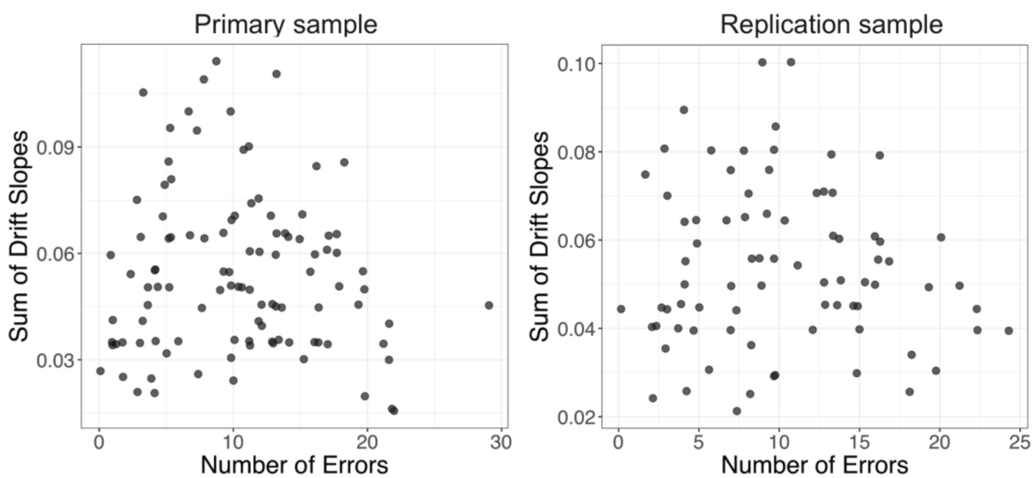

Supplementary Figure 14: Accuracy of DDM predictions of LL choices. Primary sample N=117, replication sample N=100. Line indicates equivalence between model and actual choice proportion.

**Supplementary Note 15: Accuracy of DDM predictions of response times.** We compared the average response times and model predictions of average response times. Both models fit the data well overall (two-sided Pearson's product-moment correlations: attribute-wise primary sample t(934) = 92.87, p < 0.001, r = 0.95, 95% CI = 0.94 – 0.96, option-wise primary sample t(934) = 94.23, p < 0.001, r = 0.95, 95% CI = 0.94 – 0.96; attribute-wise replication sample t(798) = 91.5, p < 0.001, r = 0.96, 95% CI = 0.95 – 0.96, option-wise replication sample t(798) = 93.43, p < 0.001, r = 0.96, 95% CI = 0.95 – 0.96). There are slight deviations in response times at the extremes (first and last octile) but all of the middle octiles lie very close to the unity line.




Supplementary Figure 15: Accuracy of DDM predictions of response times. Primary sample N=117, replication sample N=100. Line indicates equivalence between model and actual response time octiles; each octile (1-8) is shown in a different color.

**Supplementary Note 16: The sum of drift slopes**. The sum of DDM drift slopes vary across individuals. Differences in rate of accumulation could be due to variation in response times or noisy responding. To investigate this, we examined the relationship between the sum of the drift slopes and the number of error trials (when lower SV option was chosen based on a participant's individual discount rate, *k*). There was not a significant correlation between the sum of DDM drift slopes and the number of errors made (Pearson's product-moment correlations: primary sample $t(103) = -0.82$, $p = 0.41$ $r = -0.08$, 95% CI = -0.27 – 0.11; replication sample $t(77) = -0.68$, $p = 0.50$, $r = -0.08$, 95% CI = -0.29 – 0.15). This suggests that different magnitudes of drifts slopes are not simply due to noisiness in choice.




Supplementary Figure 16: Lack of correlation between sum of drift slopes and number of errors. Primary sample N = 105; Replication sample N = 79 (participants unable to be fit to a single discount rate were excluded from analyses and are not displayed). There is no correlation between the number of errors made (as measured by choosing the option with the lower subjective value) and the sum of drift slopes. Values are jittered (.001 vertical jitter) to reduce over-plotting.

**Supplementary Note 17: Contributions of model parameters to interindividual variability in patience.** We used a linear regression of model parameters from the attribute-wise and option-wise DDM onto the discount rate using standardized coefficients to enable the direct comparison of the size of the coefficients and their importance in explaining the dependent variable. Results are similar across the two model specifications with difference in drift slope most associated with intertemporal patient, latency difference showing smaller but significant effects, and decision bounds being unrelated to patience in both samples.

| Measure | Primary Sample | Replication Sample |
|---|---|---|
| **Regression of attribute-wise parameters on log(*k*)** | | |
| ***Fit measures*** | *Adj.* $R^2 = 0.86$<br>$F(3,101) = 208.7$<br>$p < 0.001$ | *Adj.* $R^2 = 0.85$<br>$F(3,75) = 144.4$<br>$p < 0.001$ |
| **Drift slope** | -0.80 ($p < 0.001$) | -0.81 ($p < 0.001$) |
| **Attribute latency** | 0.30 ($p < 0.001$) | 0.29 ($p < 0.001$) |
| **Decision bounds** | -0.01 ($p = 0.80$) | -0.03 ($p = 0.46$) |
| **Regression of option-wise parameters on log(*k*)** | | |
| ***Fit Measures*** | *Adj.* $R^2 = 0.91$<br>$F(3,95) = 320.6$<br>$p < 0.001$ | *Adj.* $R^2 = 0.88$<br>$F(3,74) = 181.7$<br>$p < 0.001$ |
| **Drift Slope Difference** | -0.83 ($p < 0.001$) | -0.89 ($p < 0.001$) |
| **Latency Difference** | 0.25 ($p < 0.001$) | 0.29 ($p < 0.001$) |
| **Decision bounds** | 0.06 ($p = 0.041$) | 0.03 ($p = 0.50$) |

Supplementary Table 1: Regression of attribute-wise and option-wise DDM parameters on discount rate. The largest influence in explaining discount rate is difference in drift slope followed by difference in latency. Note: because drift slopes in the option-wise model were log-transformed, drift slopes of 0 that cannot be log-transformed were excluded. Participants unable to be fit to a single discount rate were also excluded from analysis. Standardized betas are reported.

**Supplementary Note 18: Attribute-specific contributions to intertemporal patience.** We conducted separate linear regressions for the attribute-wise and option-wise models, each examining the contributions of DDM model parameters for amount and for time to predict intertemporal patience. In the attribute-wise model, amount and time made independent and similarly-sized contributions. Drift slopes had the largest impact, followed by latencies; there was no effect of decision bounds. However, in the option-wise model, the impact of the time drift slope was much greater than that of amount. Thus, the option-wise model has a higher impact of time drift slope, whereas the attribute-wise model has relatively even contributions of amount and time.

| Measure | Primary Sample | Replication Sample |
|---|---|---|
| **Regression of attribute-wise parameters on log(*k*)** | | |
| ***Fit measures*** | *Adj.* $R^2 = 0.87$<br>$F(5,99) = 142.2$<br>$p < 0.001$ | *Adj.* $R^2 = 0.88$<br>$F(5,73) = 114.6$<br>$p < 0.001$ |
| **Drift Slope Amount** | -0.59 ($p < 0.001$) | -0.52 ($p < 0.001$) |
| **Drift Slope Time** | 0.56 ($p < 0.001$) | 0.71 ($p < 0.001$) |
| **Latency Amount** | 0.32 ($p < 0.001$) | 0.28 ($p < 0.001$) |
| **Latency Time** | -0.15 ($p < 0.001$) | -0.22 ($p < 0.001$) |
| **Decision bounds** | -0.06 ($p = 0.13$) | -0.09 ($p = 0.10$) |
| **Regression of option-wise parameters on log(*k*)** | | |
| ***Fit measures*** | *Adj.* $R^2 = 0.93$<br>$F(5,93) = 243.1$<br>$p < 0.001$ | *Adj.* $R^2 = 0.88$<br>$F(5,72) = 112.8$<br>$p < 0.001$ |
| **Drift Slope Amount** | -0.12 ($p < 0.001$) | -0.15 ($p < 0.001$) |
| **Drift Slope Time** | 0.78 ($p < 0.001$) | 0.87 ($p < 0.001$) |
| **Latency Amount** | 0.32 ($p < 0.001$) | 0.21 ($p = 0.0013$) |
| **Latency Time** | -0.19 ($p < 0.001$) | -0.29 ($p < 0.001$) |
| **Bounds** | -0.02 ($p = 0.68$) | -0.001 ($p = 0.99$) |

Supplementary Table 2: Regression of DDM parameters on discount rate: separating amount and time contributions. Note: because drift slopes in the option-wise model were log-transformed, drift slopes of 0 that cannot be log-transformed were excluded. Participants unable to be fit to a single discount rate were also excluded from analysis. Standardized betas are reported.

## Supplementary References


1. Orquin, J. L. & Mueller Loose, S. Attention and choice: A review on eye movements in decision making. *Acta Psychol. (Amst).* **144,** 190–206 (2013).
2. Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* **13,** 1292–1298 (2010).
3. Mullett, T. L. & Stewart, N. Implications of visual attention phenomena for models of preferential choice. *Decision* **3,** 231–253 (2016).
4. Shimojo, S., Simion, C., Shimojo, E. & Scheier, C. Gaze bias both reflects and influences preference. *Nat. Neurosci.* **6,** 1317–1322 (2003).
5. Payne, J. W. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ. Behav. Hum. Perform.* **16,** 366–387 (1976).
6. Coutlee, C. G., Politzer, C. S., Hoyle, R. H. & Huettel, S. A. An abbreviated impulsiveness scale constructed through confirmatory factor analysis of the Barratt Impulsiveness Scale Version 11. *Arch. Sci. Psychol.* **2,** 1–12 (2014).