Learning decouples accuracy and reaction time for rapid decisions in a transitive inference task

Abbreviated Title: Diffusion modeling of rapid decisions in transitive inference

Fabian Munoz[1,2], Greg Jensen[3], Maxwell Shinn[4], Yelda Alkan[1,2], John D. Murray[5], Herbert S. Terrace[6], Vincent P. Ferrera[*,1,2,7,8]

1. Department of Neuroscience, Columbia University Medical Center, New York, NY, 10032,
2. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, 10027
3. Reed College, Portland, OR, 97202
4. UCL Queen Square Institute of Neurology, University College London, London WC1E 6BT, UK
5. Department of Psychology, Dartmouth University, Hanover, NH 03755
6. Department of Psychology, Columbia University, New York, NY 10027
7. Department of Psychiatry, Columbia University Medical Center, New York, NY, 10032
8. Kavli Institute for Brain Sciences, Columbia University, New York, NY, 10027

*Corresponding author email address: vpf3@cumc.columbia.edu

Keywords: sequential sampling, drift-diffusion model, drift rate, collapsing boundary, leaky integration

Abstract

The accumulation of evidence over time, formalized in the drift diffusion model (DDM), has become a prevalent model of deliberative decision-making. Here, we adapted this model to understand better the role of latent variables in a serial learning task where decisions were made rapidly and did not show typical accuracy and response time patterns. We fit behavioral data using PyDDM (Shinn et al., 2020). We trained macaque monkeys (N = 3) on a transitive inference transfer task. They learned the implied order in a ranked list of 7 novel pictures in each behavioral session, indicating their choices by making saccadic eye movements. They reliably learned each new list order within 200-300 trials with asymptotic accuracy of around 80-90% correct. Their responses showed a symbolic distance effect, with 60% accuracy for adjacent list items and 90% accuracy for the largest symbolic distance. Although performance accuracy improved with learning and symbolic distance, reaction times were nearly invariant. Nevertheless, both accuracy and reaction time were well fit by the generalized drift-diffusion model. The fits were achieved by simultaneously increasing both the evidence accumulation rate and a collapsing bound to capture the shape of the reaction time distributions. These results indicate that decision-making during learning and transfer in a TI task may be characterized by a "variable collapsing bound" DDM. These results point to a unique dynamical regime of the DDM framework during serial learning.

Introduction

Perceptual decision-making is often characterized as a process of deliberative evidence accumulation. Such models provide a bridge linking behavior to neurophysiology. A well-studied neuro-computational paradigm for decision-making is the drift-diffusion model (DDM, Ratcliff 1978), which derives from the statistical theory of sequential sampling (Wald, 1945). Drift-diffusion models a process that repeatedly samples the relative likelihood of each decision alternative until sufficient evidence supports making a choice with a given level of accuracy. These models therefore provide an explanatory framework for the frequently observed empirical relationship between performance accuracy and reaction time.

For many two-alternative choice tasks in nonhuman primates (NHP), in which response times often extend to a full second or longer, it is usually found that response times are inversely correlated with stimulus strength for both dynamic (Roitman & Shadlen 2002) and static stimuli (Ratcliff et al., 2003). These response times also correspond to the temporal evolution of neural activity in the parietal (Roitman & Shadlen 2002), prefrontal cortex (Kim & Shadlen 1999, Kiani et al., 2014), and in the superior colliculus (Jun et al., 2021). The relationship between behavioral and neural dynamics has been central to the development of computational models of perceptual decision-making.

There is a growing recognition that decision processes occur on multiple time scales (Murray et al., 2014, Soltani et al., 2021), including continuous decisions (Yoo et al.,

2021). Rapid decisions are of particular interest due to their importance in natural environments where decisions are made under considerable time pressure and survival is often at stake. Some studies have reported behavioral evidence of oculomotor decisions with average response times in the range of 200-300 msec and neural activity in FEF differentiating target location in that time frame (Cohen et al., 2010; Stanford et al., 2010; Shankar et al 2011)

Transitive inference (TI) tests the ability to learn and generalize about relationships that obey the law of transitivity (that is, if A>B and B>C, then it follows that A>C). In the current study, monkeys performed a TI task in which, in each behavioral session, they learned the implicit order in a novel set of images during pairwise presentations. While accuracy improved with learning, choice reaction times remained constant and short, even though monkeys were under little time pressure. This behavior is distinct from other perceptual tasks where reaction times tend to decrease with increasing accuracy. TI also differs from other perceptual tasks in that variations in accuracy are not directly related to the physical properties of the stimulus but to variables implicit to the task demands. For example, "symbolic distance" is the difference between the implicit ranks of items (so, given an order ABC, the pair AB has a symbolic distance of 1, whereas AC has a distance of 2), and is a strong predictor of response accuracy. The differences between the TI task and other perceptual tasks led us to ask if the DDM framework was compatible with behavior (accuracy and reaction times) measured during TI learning, and thus whether rapid decisions share common characteristics and mechanisms with more deliberative decision processes.

We tested this by fitting the data with a generalized drift diffusion model using PyDDM (Shinn et al. 2020). The results suggest that both the rate of evidence accumulation (drift rate) and the amount of accumulated evidence needed to initiate a response (boundary position) vary with learning and with symbolic distance. These results contrast with most other tasks which can be fit by varying drift rate while accumulating evidence to a constant boundary.

Methods

Subjects

Subjects were three male rhesus macaques (F, H, and L), 13-14 years old and weighing 8-10 kg at time of experiment. The research was conducted in accordance with U.S. Department of Health and Human Services (National Institutes of Health) guidelines for the care and use of laboratory animals and was approved by the Institutional Animal Care and Use Committee at Columbia University and the New York State Psychiatric Institute. Monkeys were prepared for experiments by surgical implantation of a post for head restraint and a recording chamber to give access to prefrontal cortex. All surgery was performed under general anesthesia (isoflurane 1-3%) and aseptic conditions. Subjects had prior experience performing transitive inference tasks (Jensen et al. 2015).

Visual Stimuli and Eye Movements

The subjects were seated in an upright primate chair during the experiment and responded to visual stimuli by making saccadic eye movements. Stimuli were generated and

controlled by a CRS VSG2/3F video frame buffer and displayed on a CRT monitor with a resolution of 1280 x 1024 pixels at 60 Hz. The viewing distance was 60 cm.

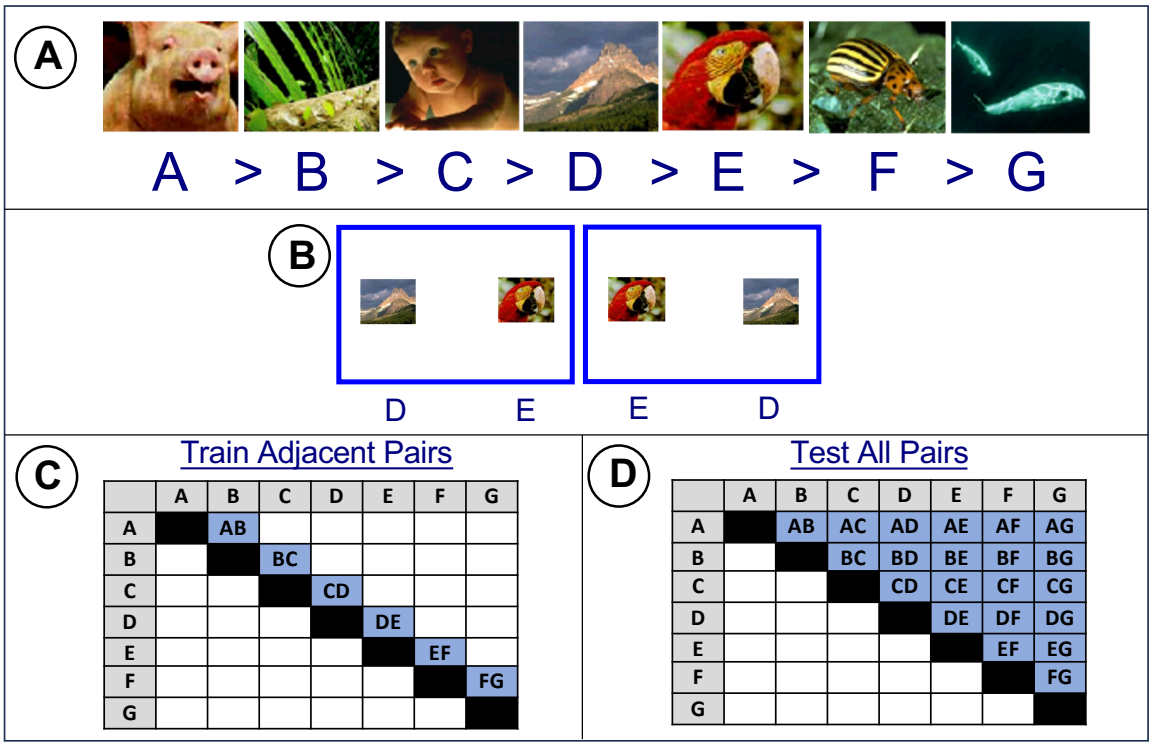
Visual stimuli consisted of 140-by-130 pixel color photographs (7° by 8° visual angle) and small squares that served as fixation and eye movement targets. To control the retinal position of visual stimuli, monkeys were required to maintain fixation while the stimuli were presented. The fixation point was a small red square (0.5° visual angle).

Eye position was recorded using a monocular scleral search coil (CNC Engineering, Seattle WA) and digitized at a rate of 1 kHz (Judge et al. 1980). Subjects expressed choices by making eye movements from the central fixation point to peripheral target stimuli. Eye velocity was computed offline by convolving eye position with a digital filter. The filter was constructed by taking the first derivative of a temporal Gaussian, $G(t)$, such that $\frac{dG}{dt} = -k \cdot t \cdot exp\left(\frac{t^2}{\tau^2}\right)$, where $\tau = 8$ msec and k is a constant that sets the filter gain to 1.0. This filter does not introduce a time shift between the position input and velocity output, but adds temporal uncertainty to the velocity estimates. Horizontal and vertical eye velocities ($h'(t)$ and $v'(t)$, respectively) were combined to estimate radial eye speed $r'(t)$ using the formula $r'(t) = \sqrt{h'(t)^2 + v'(t)^2}$.

Transitive Inference Task

Prior to the beginning of each behavioral session, a set of 7 photographs never before presented to the subjects was selected from a database of over 2500 images. A different set of images was used for each behavioral session. These stimuli were randomly assigned a unique

rank, indicated by the letters A thru G, to create an ordered list (**Fig. 1A**). No explicit information about stimulus rank was presented to the subjects. That is, the rank assigned to



	A	B	C	D	E	F	G
A		AB					
B			BC				
C				CD			
D					DE		
E						EF	
F							FG
G							

	A	B	C	D	E	F	G
A		AB	AC	AD	AE	AF	AG
B			BC	BD	BE	BF	BG
C				CD	CE	CF	CG
D					DE	DF	DG
E						EF	EG
F							FG
G							

Figure 1. **Transitive Inference Transfer Task.** This task tests subjects' ability to learn the implied serial order of a set of images from pairwise presentations. **A.** Set of 7 randomly selected images arranged in an ordered list (A>B>C>D>E>F>G). A novel set of images was used for each behavioral session. **B.** Choice trials with spatial position counterbalanced. **C.** Adjacent pairs used for the training phase of the session **D.** All pairs used for the transfer phase.

each stimulus was never displayed, nor was there any information about serial order conveyed by the spatial or temporal presentation of the stimuli.

On each trial (**Fig. 1B**), the subject first fixated a small target in the center of the screen. Following a random delay (0.4 sec to 1.2 sec, positively skewed with a mean of 0.5 sec), the fixation point was extinguished, and two stimuli appeared equidistant from the center of the screen in opposite directions. To receive a reward, subjects had to select the stimulus with the

lower rank by making a saccadic eye movement to the selected stimulus within 1.5 sec of stimulus onset and fixating for a random interval (0.4 to 0.6 sec, uniformly distributed). When this interval elapsed, auditory feedback was delivered indicating that the response was either correct (high tone, 880 Hz) and would be rewarded, or was incorrect (low tone, 440 Hz). To receive the reward on correct trials, subjects had to maintain fixation for another interval (0.35 to 0.65 sec, uniformly distributed), after which the screen went dark and fluid rewards (juice or water drops) were delivered through a sipper tube.

In the training phase of the task, adjacent pairs of stimuli were presented. Each pair comprised two stimuli whose ordinal positions were adjacent in the list (A/B, B/C, C/D, etc, **Fig. 1C**). To receive a reward, the subject was required to make a saccadic eye movement to the stimulus with the lower rank (e.g. A from A/B). For a 7-item list, there are 6 adjacent pairs. With positional counterbalancing, this resulted in training blocks that were 12 trials in length. This phase lasted up to 20 blocks (i.e. 240 trials in total).

In the testing phase of the task, all possible combinations of the 7 items were presented (**Fig. 1D**). The subjects were still required to saccade to the lower ranking stimulus. This phase tested the subjects' ability to transfer knowledge about the list order gained during the adjacent pairs phase to novel, non-adjacent pairs. For a 7-item list, there are 21 possible pairs (6 adjacent, 15 non-adjacent). With counterbalancing of the stimulus positions, this resulted in blocks of 42 trials. This phase generally lasted for 5-10 blocks (240-420 trials).

Each pair of stimuli could be identified in terms of the ranks of the individual stimuli they included. Pairs could be also described in terms of "symbolic distance," a metric that is

calculated by taking the difference between the individual stimulus ranks (D'Amato & Colombo 1990). For example, since the pair B/C consists of items that are adjacent in the list (i.e. ranks

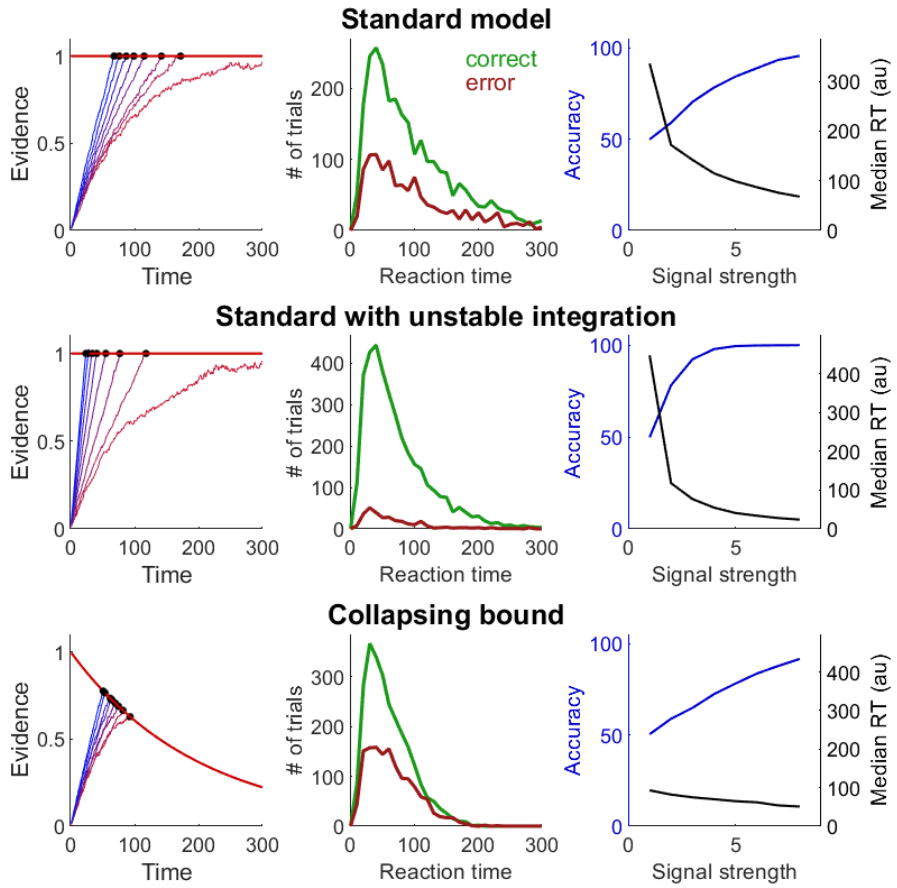


Figure 2. DDM Simulations. Top: Standard model with constant bound and no leak. Left: red traces represent weaker signal strengths, blue is stronger signal. Center: reaction time histograms for correct (green) and error (red) trials for a single, moderate signal strength. Right: Accuracy (percent correct) and reaction time as a function of signal strength. **Middle:** Standard model with leak. **Bottom:** Collapsing bound model.

2 and 3, respectively), the symbolic distance is 1, while the pair A/G (ranks 1 and 7) has a symbolic distance of 6.

Generalized Drift Diffusion Modeling (GDDM)

Behavioral data were fit to a Generalized DDM using the PyDDM toolbox (Shinn et al., 2020). We used several different approaches to understand the effect of each variable on the dynamics of the GDDM model in relation to serial learning. In all approaches, the data to be fit were reaction time and accuracy segregated by symbolic distance or trial block. The two conditions (symbolic distance or trial block) were fit separately. Parameters were estimated for each NHP subject individually.

Standard model. In the first, "standard," approach, a minimal version of the GDDM was implemented with four parameters: drift rate (rate of evidence accumulation), boundary level (amount of accumulated evidence required for a decision), noise (instantaneous variability in drift rate), and non-decision time. The time evolution of the decision variable, x (accumulated evidence), was given by:

Eqn 1. $$dx = [drift * condition] * dt + dW$$

"*Drift*" is the drift rate parameter. "*Condition*" is either symbolic distance or trial block, depending on how the data are segregated, and is coded as an integer; [1…,6] for symbolic distance or [1,…,8] for trial block. "*W*" is the momentary noise represented as a Weiner process with standard deviation = 1. We fixed the noise parameters to $mu = 0$ and $sigma = 1.0$ to avoid redundancy with the boundary height parameter as scaling noise variability is equivalent to scaling the boundary height. The time resolution (dt) was 10 msec and the overall duration was 1.0 sec.

For this approach, the fitting algorithm searched for a single set of parameters (drift rate, bound, and non-decision time) that fit the data within each condition. All parameters except for noise were allowed to vary freely within a wide range (**Table 1**). The drift rate was the product of the scalar drift parameter and the condition index, and therefore varied linearly with condition. The boundary was constant both in time, and across conditions. The same parameter ranges were used for all NHP and all conditions.

Leaky integration. In the standard approach, there was no loss of evidence over time. However, evidence integration can be imperfect. Thus, in a second, "leaky integration" approach, imperfect integration was modeled by adding a leak term to the standard model:

Eqn 2. $$dx = [drift * condition + leak * x] * dt + dW$$

The *leak* is a constant that multiplies accumulated evidence (x) at each moment. If the leak term is negative, the evidence tends to decay exponentially toward zero. If the leak is positive, it acts like a gain factor that causes the evidence to increase exponentially (unstable integration).

Collapsing bound. In both the standard and leaky integration approaches, the decision boundary remained constant while evidence was being accumulated. To understand the effect of allowing the boundary to change over time, a "collapsing bound" approach was tested. In this approach, the bound collapsed exponentially, thus introducing a sixth parameter, *tau*, the collapse rate. In this approach, the equation for momentary evidence was the same as **Eqn 1**, but the time-dependence of the boundary was given by:

Eqn 3. $$bound = B_0 * exp(-t/tau)$$

Where B_0 is the initial boundary level at $t = 0$. This model was tested both without a leak parameter (using **Eqn 1** for evidence accumulation) and with a leak (using **Eqn 2**).

Simulations. To illustrate how each model parameter affected dynamics and performance, simulations of the standard, leaky integration, and collapsing bound models are shown in **Fig 2**. The left panel illustrates the time evolution of accumulated evidence for a range of drift rates (signal strengths). The center panel shows the reaction time histograms for correct and error trials. The right panel shows accuracy and reaction time. Non-decision time was set to zero in all simulations.

For the standard model (**Fig 2**, top row), low signal strength was associated with chance performance accuracy and long reaction times. The addition of a positive leak term (**Fig 2**, middle row) restricted the range of reaction times by speeding up evidence accumulation, but also reduced accuracy for moderate signal strengths. The collapsing bound model (**Fig 2**, bottom row) performed similarly to the leaky integration model but was more effective in reducing long reaction times and produced more symmetric reaction time distributions.

Table 1. PyDDM parameter ranges**.**

Range	**Drift**	**Boundary**	**Noise**	**Leak (if present)**	**Tau**	**Non-decision**
Standard model						
min	0.0	0.1	1.0	0	-	50 msec
max	10.0	10.0	1.0	100	-	400 msec
Collapsing bound model						
min	0.0	0.3	1.0	-5.0	0.01	50 msec
max	5.0	3.0	1.0	5.0	10	400 msec

Results

Behavioral Performance

Behavioral performance was quantified in terms of accuracy (percent correct) and response time (time from stimulus onset until choice saccade onset). A total of 273 behavioral sessions were recorded (F: 85, H: 98, L: 90 sessions). Each session started with a novel set of pictures that were arranged into a list by assigning a rank order to each picture (**Fig. 1A**). The ranking was initially unknown to the subject but was learned over the course of the session.

Reaction times (RT) as a function of trial number for all sessions completed by NHP H are shown in **Fig. 3A**. The three phases of the task are color-coded (single target trials in gray, adjacent pairs in blue, all pairs in magenta). There was an abrupt shift in the RT distribution (**Fig. 3B**) at the transition from single stimuli (median RT 192 msec) to stimulus pairs (median RT 270 to 275 msec).

Adjacent-pair (training phase) and all-pair (testing phase) trials were grouped into blocks of 50 consecutive trials to obtain robust estimates of RT (**Fig. 3A**). The shape and median value of the RT distributions remained constant across trial blocks. The RT distributions were similar in shape for correct (**Fig 3B** solid lines) and incorrect (dashed lines) trials, and for adjacent (**Fig. 3B,** blue lines) and non-adjacent pairs (magenta lines).

The effect of learning on accuracy and RT for this NHP is shown in **Fig. 3C**. During adjacent pair training, logistic regression showed a significant effect of trial number on accuracy for all NHPs (F: $p = 0.03$, df = 15,206; H: $p < 0.0001$, df = 16,012; L: $p < 0.0001$, df = 6658).

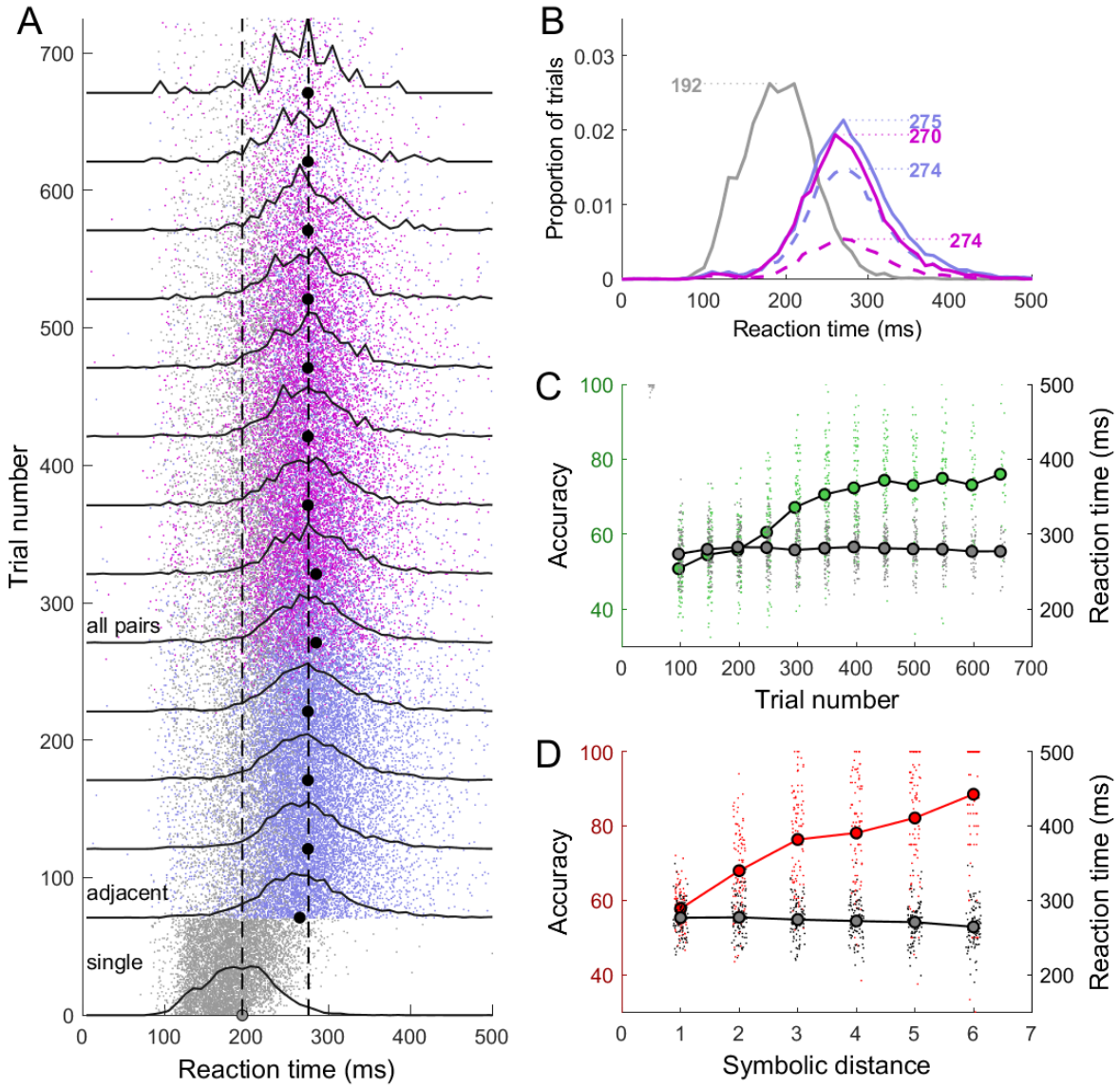


Figure 3. **Behavioral Performance on TI transfer task. A.** Reaction time vs trial number for NHP H. The colors represent the three task phases (blue – presentation of single items, cyan – adjacent pairs, red – all pairs). **B.** Reaction time histograms. Colors as in A. For choice trials, solid lines are correct responses and dashed lines are incorrect. **C.** Accuracy (green) and reaction time (gray/black) during learning. **D.** Accuracy (red) and reaction time (grey/black) as a function of symbolic distance.

Logistic regression coefficients for accuracy versus trial number were similar for all NHP (0.096 for NHP F, 0.07 for NHP H, 0.11 for NHP L). The effect of trial number on RT, as assessed by linear regression, was also significant ($p < 0.0001$ for all NHPs). However, the slopes of the regression lines were shallow, amounting to an increase of 5 – 7 msec for every 100 trials.

During all-pairs testing, the relationship between trial number and accuracy remained significant ($p < 0.0001$ for all NHP) and the RT trend turned slightly negative (-0.5 msec per 100 trials) and was significant for 2 NHP ($p < 0.002$). However, the relationship was inconsistent - negative for one NHP (H: decrease of 1.6 ms/400 trials), and positive for the other (L: increase of 4 ms/400 trials).

Symbolic distance (SD, the difference in rank between the two stimuli in each pair) had a strong effect on accuracy (**Fig. 3D**). Logistic regression coefficients for SD were consistent across NHP (0.35 for NHP F, 0.28 NHP H, 0.22 for NHP L). For RT, there was a small but significant reduction with SD for both NHPs (F: slope = -1.5 msec, $p < 0.0001$; H: -2.9 msec, $p < 0.0001$; L: -0.33 msec, $p = 0.02$). The results overall indicate a strong effect of trial number and SD on performance accuracy, but little effect on response time.

Speed-accuracy trade-off

Although there was not a clear relationship between accuracy and choice reaction time when data were combined over sessions, a more fine-grained analysis might reveal subtler trends. We therefore calculated the residual accuracy and average reaction time in each block of 50 consecutive trials during the adjacent-pair training and all-pairs testing phases of the task.

"Residual" is here defined as the accuracy or average reaction time within each block of 50 trials normalized by subtracting the averages of each of those measures calculated over all choice trials for the entire session. By accounting for session-to-session variability, the residuals proided estimates of the change in accuracy per unit change in reaction time that could be combined across sessions.

Figs. 4A,B,C show an improvement in accuracy with lengthening reaction times during the training phase of the task (first 250 choice trials of each session) for all three NHP. Each data point represents a single 50-trial block from one session. There was an average increase in performance ranging from 0.24 to 0.43 percent correct for every additional millisecond of reaction time. The correlation between accuracy and RT was significant for each NHP. During

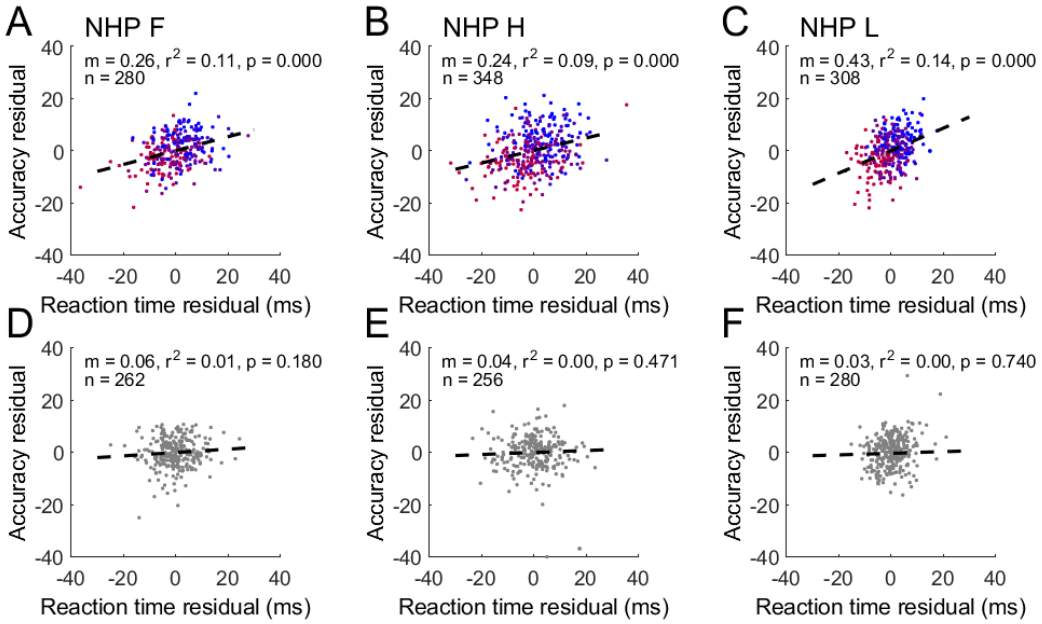


Figure 4. Speed-accuracy trade-off. A,B,C. Improvement in accuracy v. increase in RT during training phase for each NHP. 'm' is the linear regression slope, 'r^2' is the coefficient of determination, 'p' is the significance of the correlation (Pearson), and 'n' is the number of 50-trial blocks per session times the number of sessions. The color gradient indicates trial blocks that were earlier in training (red) vs. later (blue). **D,E,F.** All-pairs testing phase.

the subsequent testing phase (the first 250 trials after end of training)(**Fig. 4 D,E,F**), no significant relationship was found for any NHP.

During the all-pairs testing phase, it is also possible to sort trials into six groups by symbolic distance (SD), rather than by trial block. When this was done, a significant negative relationship was found between the change in accuracy and increasing reaction time (**Fig. 5 A,B,C**); trials with larger SD had higher accuracy and shorter RTs.

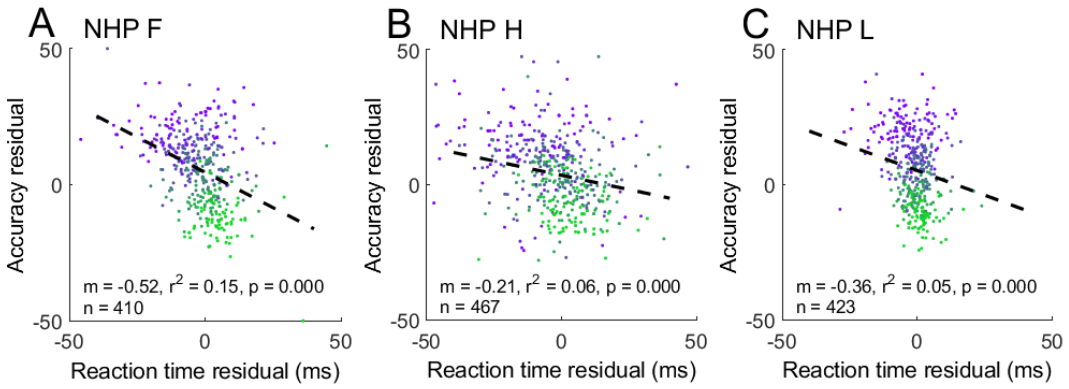


Figure 5. Speed-accuracy trade-off for trials grouped by symbolic distance. A,B,C. Decrement in accuracy v. increase in RT during testing phase for each NHP. 'm' is the linear regression slope, 'r^2' is the coefficient of determination, 'p' is the significance of the correlation (Pearson), and 'n' is the number of symbolic distances times the number of sessions. The color gradient indicates smaller (green) vs. larger (magenta) symbolic distances.

PyDDM Results

Choice reaction times during the TI transfer task were generally under 300 msec even though subjects were allowed up to 1500 msec after stimulus onset to respond by making a

saccadic eye movement to the target or distractor. Choice RTs averaged only 35 msec longer than simple RTs for NHP F (215 vs. 180 msec median RT), 85 msec for NHP H (275 vs. 190 msec), and 30 msec for NHP L (211 vs. 181 msec). RTs remained nearly constant despite learning-related improvements in accuracy. RTs were also constant across symbolic distance despite large differences in accuracy. Drift diffusion models (DDMs) with constant bounds and variable drift rate typically show increases in accuracy that are accompanied by shorter reaction times. As the current results did not show this pattern, it raises the question of whether DDMs can fit the behavior exhibited by NHP during the transitive inference task and thus shed light on how decisions are made during TI learning.

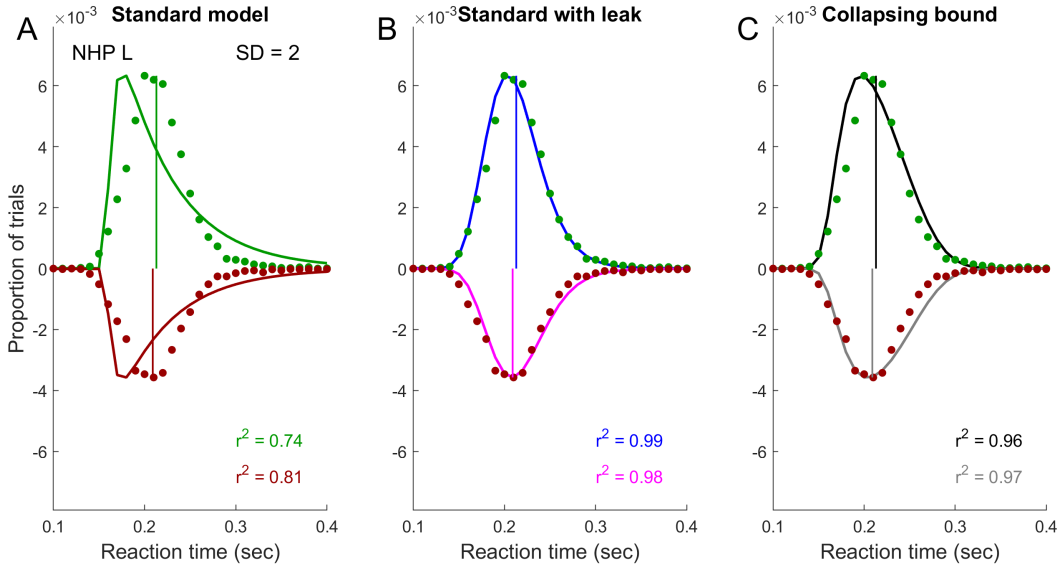


Figure 6. Drift Diffusion model fits. Data from all sessions for NHP L with symbolic distance (SD) equal to 2. Green dots are reaction time distributions for correct trials, red dots for error trials. Vertical red and green lines indicate the median RT for each distribution. Green and red curves are fits for correct and error trials, respectively, using the standard model. The correlation between the fits and the data are shown in red and green text. Blue and magenta curves and text are fits obtained with the leaky integration model. Black and grey curves and text are fits obtained with the collapsing bound model.

To test this, we first fit a generalized DDM with four parameters (drift rate, noise, boundary, and non-decision time, see Methods) to the current behavioral data using the PyDDM package (Shinn et al. 2020). An example of the PyDDM fits for data segregated by outcome (green = correct, red = error) for NHP L is shown in **Fig. 6** (the fits for all NHP are shown in supplemental **Figs. S1** and **S2**). The fits for the standard model are shown as the green and red curves. These fits were relatively poor. This was confirmed by the variance

Table 2. PyDDM fitted parameters for each model, NHP, and condition. For "condition," SD is symbolic distance, TB is trial block. Model 1 is the standard model with constant bound and no leak. Model 2 is the standard model with leak. Model 3 is the collapsing bound model with no leak. Model 4 is the collapsing bound with leak. BIC is Bayesian information criterion. VAC is variance accounted for (r^2).

NHP - Condition	Model	Drift	Bound	Noise	Leak	Tau	Non-decision	BIC	VAC
NHP F - SD	1	0.68	0.30	1.0	-	-	152 msec	-73232	0.78
NHP F - SD	2	0.99	9.0	1.0	43.1	-	105 msec	-85582	0.96
NHP F - SD	3	0.55	0.58	1.0	-	0.15	122 msec	-83621	0.93
NHP F - SD	4	0.50	0.58	1.0	-4.9	0.14	118 msec	-83706	0.93
NHP H - SD	1	0.61	0.35	1.0	-	-	184 msec	-39867	0.75
NHP H - SD	2	0.83	8.6	1.0	29.8	-	116 msec	-52508	0.96
NHP H - SD	3	0.50	0.88	1.0	-	0.15	126 msec	-48531	0.91
NHP H - SD	4	0.45	0.92	1.0	-4.7	0.15	110 msec	-48737	0.92
NHP L - SD	1	0.57	0.30	1.0	-	-	154 msec	-103034	0.73
NHP L - SD	2	0.93	8.2	1.0	69.4	-	133 msec	-120970	0.97
NHP L - SD	3	0.47	0.46	1.0	-	0.15	137 msec	-118035	0.92
NHP L - SD	4	0.47	0.47	1.0	-3.9	0.10	140 msec	-118598	0.93
NHP F - TB	1	0.29	0.30	1.0	-	-	151 msec	-54357	0.79
NHP F - TB	2	0.41	9.9	1.0	44.5	-	100 msec	-64415	0.96
NHP F - TB	3	0.25	0.65	1.0	-	0.15	116 msec	-57016	0.95
NHP F - TB	4	0.50	0.58	1.0	-4.9	0.14	118 msec	-83706	0.95
NHP H - TB	1	0.27	0.35	1.0	-	-	187 msec	-34058	0.74
NHP H - TB	2	0.36	9.5	1.0	28.6	-	105 msec	-45301	0.96
NHP H - TB	3	0.23	0.99	1.0	-	0.15	125 msec	-38285	0.91
NHP H - TB	4	0.21	0.92	1.0	-4.8	0.15	107 msec	-38316	0.91
NHP L - TB	1	0.28	0.25	1.0	-	-	156 msec	-76581	0.73
NHP L - TB	2	0.46	8.2	1.0	71.2	-	133 msec	-90833	0.96
NHP L - TB	3	0.22	0.86	1.0	-	0.10	116 msec	-80666	0.93
NHP L - TB	4	0.24	0.54	1.0	-3.7	0.10	132 msec	-80857	0.93

accounted for by the standard model (VAC, measured by r^2), which was generally less than 80%. The fitted parameter values are shown in **Table 2**.

To estimate the effect of within-session learning on model fits, the data were grouped into 8 blocks of 50 consecutive trials spanning the first 400 trials of paired presentations. During these trials, accuracy improved from chance levels to roughly 80% correct. The variance accounted for (r^2) was generally $< 80\%$ for each trial block condition for all NHP (**Fig S2**). These results indicate that a DDM with constant bounds and no leak term provided a poor fit to data in which accuracy improves across conditions while RTs remain relatively constant. We next explore the role of each model parameter by testing variants of the standard DDM.

Effect of Drift Rate. The drift term captures whether the rate of evidence accumulation is sensitive to the stimulus conditions, as indexed by symbolic distance, or stage of learning, as indexed by trial block. The fitted parameters indicate an increased rate of evidence accumulation (drift rate) as a function of both SD and trial block (TB). This is expected as drift rate has a strong effect on accuracy. The drift rate estimates were similar across NHP. However, there was a factor of 2 difference between conditions, most likely because performance accuracy reached higher asymptotic levels as a function of symbolic distance (**Fig. 3D**) than it did for trial block (**Fig. 3C**).

Effect of Leak. The leak term modulates the rate of evidence accumulation. It can be thought of as a gain control that dampens or enhances the effect of drift rate and has a more powerful effect as more evidence is accumulated. Adding a leak term to the standard model improved the fits, as shown by the blue and magenta curves (**Fig. 6**). These fits were

substantially better than the standard non-leaky model, resulting in VACs in the range of 88% to 97%. The Bayesian Information Criterion (BIC) scores were also improved (**Table 2**).

The leak term can either slow or accelerate the rate of evidence accumulation depending on its sign. For the standard model with a leak, the fitted leak term was positive, indicating unstable integration. The magnitude of the leak term was quite large such that it played a dominant role in the model's performance, minimizing the effect of drift rate and boundary.

Effect of Boundary. The initial boundary level sets the amount of accumulated evidence needed to render a decision, whereas the collapse rate determines if that amount of evidence is stable over time. The collapsing bound model with no leak provided fits that were as good as the standard (constant bound) model with leaky integration, as measured by VAC and BIC. Adding a leak term to the collapsing bound model further improved the fits but only by a small amount (**Table 2**).

When a leak term was added to the collapsing bound model, the fitted values were negative, indicating true leaky integration (loss of evidence over time). This contrasts with the function of the leak parameter in the constant bound model, which was positive, indicative of unstable integration.

Effect of Non-decision Time. All NHP had non-decision times that were similar across conditions. However, there were differences among the NHP. Non-decision times for NHP F and L were on average 30-32 msec shorter than NHP H. It is of interest that the simple RTs for NHP F and L had the same average (181 msec), while H had the longest simple RTs at 192 msec suggesting that simple RTs may be related to non-decision time.

335 Speed-accuracy trade-off. The modeling approaches considered so far have been
336 evaluated on the basis of their ability to fit average accuracy and reaction time data. We have
337 not yet considered the relationship between residual accuracy and reaction time (i.e. the speed-
338 accuracy trade-off) shown in **Figs. 4** and **5**. The striking feature of these trade-offs is that they
339 have positive slope as a function of trial block during training and negative slope as a function
340 of symbolic distance during testing.

341 The standard DDM tends to produce a negative relationship between accuracy and
342 reaction time as indicated in **Fig. 2.** One way to turn this into a positive relationship is to
343 allow the boundary to vary. If the boundary height increases with signal strength, then
344 accuracy and reaction time can increase together. The relationship between boundary position
345 and signal strength that leads to increasing reaction times can be calculated analytically

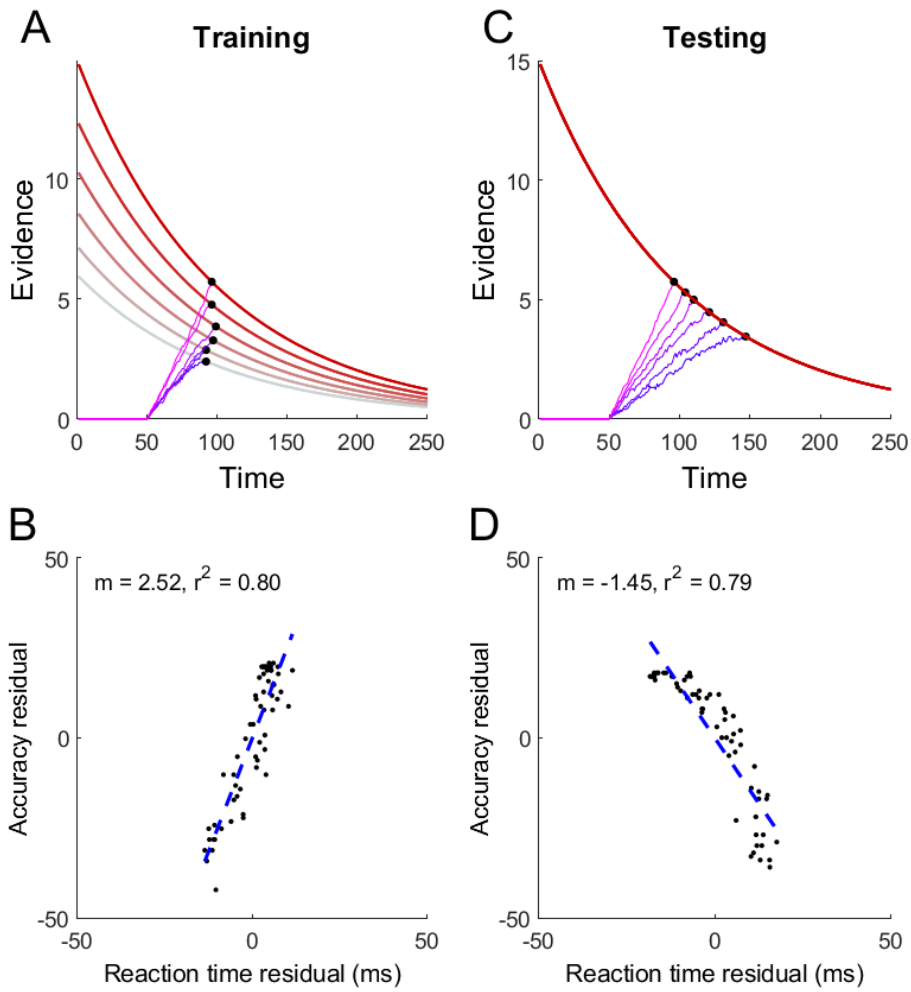


Figure 7. Variable collapsing bound model. A. Initial boundary position increases with signal strength during training. **B.** Positive correlation between residual reaction time and accuracy. **C.** Initial boundary position stabilizes after training. **D.** Negative relationship between reaction time and accuracy during testing.

(Bogacz et al 2006, 2008). These calculations reveal that accuracy is related to the product of drift rate and boundary position, while reaction time is more closely related to their ratio.

Fig. 7A illustrates a "variable collapsing bound model" in which an increasing initial boundary position leads to reaction times that increase with signal strength (**Fig. 7B**). When

the initial boundary position is fixed (**Fig. 7C**), reaction time decreases with signal strength leading to a negative relationship between accuracy and reaction time (**Fig. 7D**).

In the collapsing bound model, initial boundary position was constant across trial blocks. We tested a version of the collapsing bound model in which initial boundary position was free to vary in each trial block. In this case, the fitted initial boundary position tended to increase

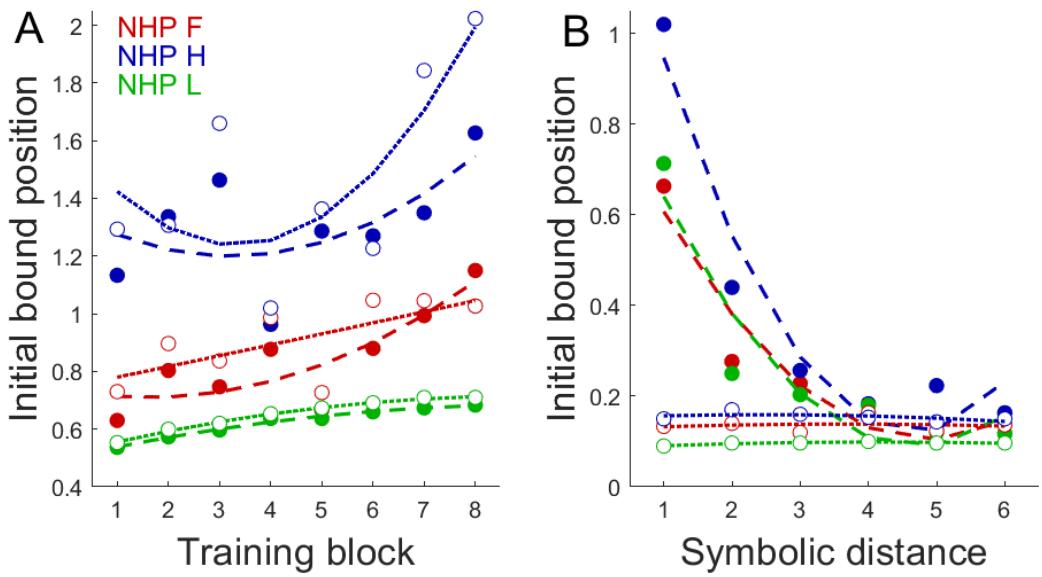


Figure 8. Variable collapsing bound fitted to training and testing data. A. Initial boundary position fit independently for each training block for all 3 NHP. Solid symbols are parameters for a model with a leak term. Dashed lines are quadratic fits to these data. Open symbols and dotted lines are parameters and quadratic fits to a model without a leak term. **B.** Initial boundary position fit to each symbolic distance. Same conventions as **A**.

during training for all 3 NHP (**Fig. 8A**). This was true whether there was a leak term included in the model or not. When each symbolic distance was fit separately (**Fig. 8B**), the initial bound position decreased when the leak term was included. Without the leak term, the bound position was constant across symbolic distance.

Discussion

The transitive inference (TI) paradigm tests subjects' ability to learn the implied serial order of a set of pictorial stimuli without any explicit spatial or temporal cues. This is accomplished by presenting pairs of images drawn from a rank-ordered set of images and rewarding subjects for choosing the image that has the lower rank. When presented with a novel set of images, NHP subjects initially respond at chance levels, with performance improving over the first ~300 trials of training. Little is known about the latent variables that underlie decision-making during TI learning. One way to approach this is by analyzing performance accuracy and reaction time in the context of evidence accumulation models.

In the current study, choice reaction times were generally under 300 msec, and only 30 to 85 msec longer than simple reaction times, suggesting that decisions were made on a compressed time scale. Nevertheless, these decisions supported learning, as evidenced by an increase in response accuracy, within each behavioral session. The improvement in performance with learning was accompanied by very modest reductions in response time.

Response accuracy also depended on the symbolic distance between each pair of images. Larger symbolic distances were associated with higher accuracy. However, there was only a small effect of symbolic distance on reaction time, consistent with a previous report from our lab (Munoz et al., 2020).

This pattern of results contrasts with some other studies of decision making (Roitman & Shadlen 2002, Ratcliff et al., 2003, Kim & Shadlen 1999, Kiani et al., 2014, Jun et al., 2021) in which choice RTs tend to be much longer (500-1000 msec), suggestive of a more deliberative

decision process. Choice RTs in other studies have also been found to be inversely related to accuracy. This pattern of anti-covarying accuracy and RT has been taken as strong evidence for gradual evidence accumulation as predicted by sequential sampling frameworks such as the drift diffusion model, which accumulates evidence over time until the amount of evidence supports a decision with a given level of accuracy.

The drift diffusion model can incorporate a number of variables whose interplay changes the model dynamics and allows for distinct modes of operation. The constant bound mode is illustrated in **Fig. 2,** top row. This is perhaps the most common mode reported in studies that provide support for evidence accumulation in decision-making. Unfortunately, the dynamics of this model can result in catastrophically long reaction times when evidence is weak and accumulation is mostly stochastic. We call this the "zero-drift catastrophe."

A time-dependent boundary can avoid the zero-drift catastrophe. When evidence is accumulating slowly, there is little advantage to lengthening the accumulation period as there is no gain in performance over time. A collapsing boundary can hasten the decision in this case. On the other hand, when evidence is accumulating rapidly, increasing the boundary by a small amount may increase performance dramatically for a very slight additional time investment.

Adding leaky integration by incorporating a term that is proportional to momentary evidence can dramatically modulate the rate of evidence accumulation (**Fig. 2,** middle row). When the leak term is positive, it acts like a gain factor, resulting in unstable integration that rapidly races toward the boundary. To fit the current data with a fixed boundary, PyDDM found that a large gain factor was needed. Thus, evidence accumulation was dominated by the

gain factor rather than by the drift rate. The need to use such a large gain factor to fit the current data calls into question the plausibility of models with unstable integration.

An alternative to unstable integration is to allow the evidence boundary to collapse exponentially over time (**Fig. 2**, bottom row). This approach also reduces long reaction times and makes the reaction time histograms more symmetrical, which is in accord with the data. The collapsing bound approach provided fits that were comparable to the constant bound with leak model. When the collapsing bound was accompanied by a leak term, the fitted leak was negative, indicating true leaky integration. The model fits were improved slightly over the collapsing, no leak condition. Thus, the collapsing bound model can fit the current data with or without a leak term. The collapsing bound model provided excellent fits to the current data and seems the most plausible of the various approaches compared in this study.

A closer look at the relationship between accuracy and reaction time (**Figs. 4 and 5**) revealed a weak but significant positive correlation during the training phase of the task. In the DDM, speed-accuracy trade-offs reflect the interplay of drift rate and boundary position. When changes in accuracy are driven by changes in drift rate, increasing drift rate leads to higher accuracy and shorter reaction times. When changes in accuracy are driven by changes in the boundary, the opposite pattern can be seen – increasing the bound height leads to higher accuracy and longer reaction times. It is therefore possible that, during learning, increases in drift rate are accompanied by adjustments of the boundary height that allow subjects to gain accuracy by slightly delaying their responses.

These observations led us to consider a hybrid "variable collapsing bound" model (**Fig. 7**). In the hybrid model, the collapsing boundary gradually increases during the training phase of the task (**Fig. 7A**), leading to a positive correlation between accuracy and reaction time (**Fig. 7B**). The boundary stabilizes during the testing phase (after learning) (**Fig. 7C**), giving rise to a negative relationship between accuracy and reaction time (**Fig. 7D**), as found in the behavioral data when accuracy and reaction time are sorted by symbolic distance (**Fig. 5.**)

Conclusion

The drift diffusion model is often fit to data where accuracy and reaction time vary inversely. However, this class of model has the flexibility to characterize results that are less typical, thus shedding light on a broad range of decision-making tasks. Here, we show that it can be applied to a reasoning task where reaction times are nearly constant despite learning-related gains in accuracy. During learning, as accuracy starts to improve, an exponentially collapsing boundary can be slowly adjusted, in effect, gradually increasing the amount of evidence required as evidence begins to accumulate more rapidly. Accuracy in this task also depends on internally computed variables, such as symbolic distance. Here, the collapsing boundary minimizes changes in reaction time as accuracy improves. In both scenarios, decision-making is made more efficient, compared to a more conventional constant bound model, by taking advantage of an increasing rate of evidence accumulation during learning to increase accuracy with little cost in reaction time. On the other hand, when evidence is accumulating

slowly, a collapsing boundary reduces the total amount of evidence required for a decision as little accuracy is gained by increasing the integration time.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on request.

Competing interests

The authors declare they have no competing interests.

Acknowledgements

This study was supported by NIH R01-MH111703 (VPF and HST).

Author Contributions

The study was designed by FM, GJ, YA, VPF, and HST. Data were collected by FM and YA. Data were analyzed by FM, GJ, MS, JDM and VPF. The manuscript was written by FM, GJ, MS, and VPF.

Supplemental Figures

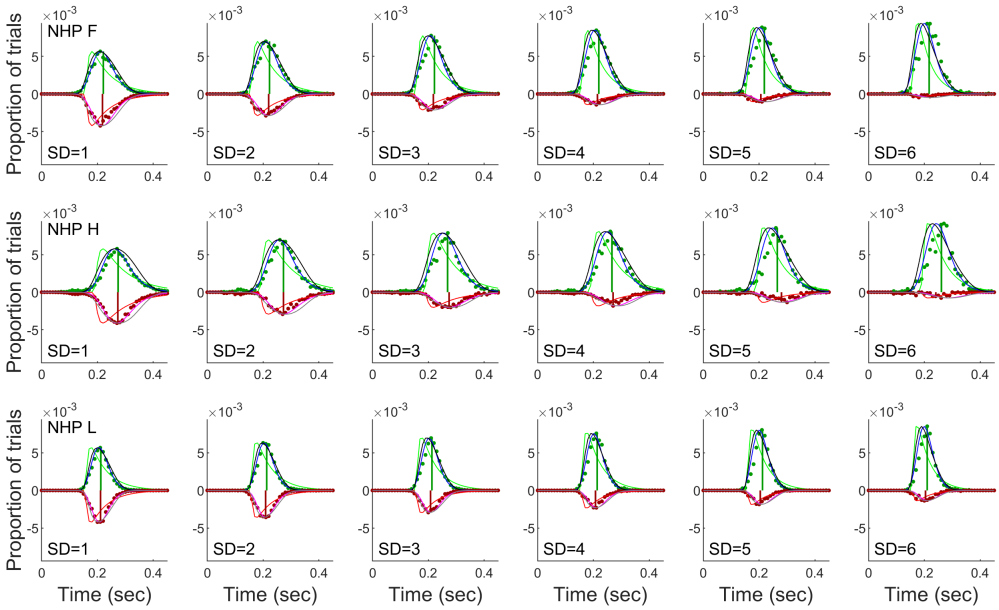


Figure S1. Drift Diffusion model fit for symbolic distance. Data from all sessions for all NHP segregated by trial block. Green dots are reaction time distributions for correct trials, red dots for error trials. Vertical lines indicate the median RT for each distribution. Green and red curves are PyDDM fits for the constant bound model. Blue and magenta curves are fits for the constant bound with leak model. Black and grey curves are fits for the collapsing bound model.

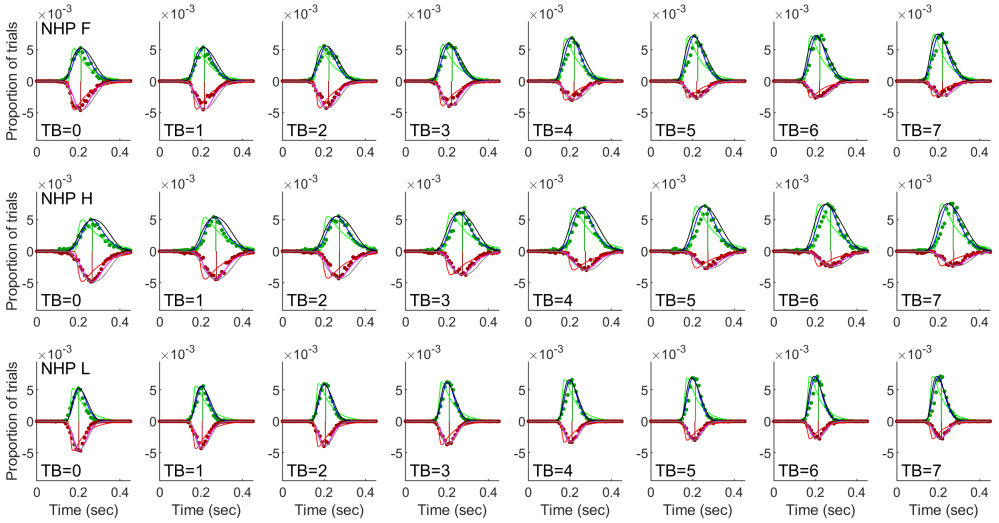


Figure S2. Drift Diffusion model fit for trial block. Data from all sessions for all NHP segregated by trial block. Green dots are reaction time distributions for correct trials, red dots for error trials. Vertical lines indicate the median RT for each distribution. Green and red curves are PyDDM fits for the constant bound model. Blue and magenta curves are fits for the constant bound with leak model. Black and grey curves are fits for the collapsing bound model.

References

1. Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol Rev. 2006 Oct;113(4):700-65. doi: 10.1037/0033-295X.113.4.700. PMID: 17014301.
2. Bogacz R, Wagenmakers E-J, Forstmann BU, Nieuwenhuis S (2008) The neural basis of the speed-accuracy trade-off. Trends in Neurosciences. 33(1):10-16
3. Cohen JY, Crowder EA, Heitz RP, Subraveti CR, Thompson KG, Woodman GF, Schall JD. Cooperation and competition among frontal eye field neurons during visual target selection. J Neurosci. 2010 Mar 3;30(9):3227-38. doi:

4. Jensen G, Muñoz F, Alkan Y, Ferrera VP, Terrace HS. Implicit value updating explains transitive inference performance: the betasort model. *PLOS Computational Biology*, *11*, e1004523. (2015).

5. Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Research*, *20*, 535-538. (1980).

6. Jun EJ, Bautista AR, Nunez MD, Allen DC, Tak JH, Alvarez E, Basso MA. Causal role for the primate superior colliculus in the computation of evidence for perceptual decisions. Nat Neurosci. 2021 Aug;24(8):1121-1131. doi: 10.1038/s41593-021-00878-6. Epub 2021 Jun 28. PMID: 34183869; PMCID: PMC8338902.

7. Kiani R, Cueva CJ, Reppas JB, Newsome WT. Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. Curr Biol. 2014 Jul 7;24(13):1542-7. doi: 10.1016/j.cub.2014.05.049. Epub 2014 Jun 19. PMID: 24954050; PMCID: PMC4191655.

8. Kim JN, Shadlen MN. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat Neurosci. 1999 Feb;2(2):176-85. doi: 10.1038/5739. PMID: 10195203.

9. Munoz, F., Jensen, G., Kennedy, B.C. *et al.* Learned Representation of Implied Serial Order in Posterior Parietal Cortex. *Sci Rep* **10**, 9386 (2020). https://doi.org/10.1038/s41598-020-65838-9

10. Murray, J., Bernacchia, A., Freedman, D., Romo R, Wallis JD, Cai X, Padoa-Schioppa C, Pasternak T, Seo H, Lee D, Wang X-J. A hierarchy of intrinsic timescales across primate cortex. Nat Neurosci 17, 1661–1663 (2014). https://doi.org/10.1038/nn.3862

11. Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

12. Ratcliff R, Cherian A, Segraves M. A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. J Neurophysiol. 2003 Sep;90(3):1392-407. doi: 10.1152/jn.01049.2002. Epub 2003 May 21. PMID: 12761282.
13. Roitman JD, Shadlen MN. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci. 2002 Nov 1;22(21):9475-89. doi: 10.1523/JNEUROSCI.22-21-09475.2002. PMID: 12417672; PMCID: PMC6758024.
14. Shankar S, Massoglia DP, Zhu D, Costello MG, Stanford TR, Salinas E. Tracking the temporal evolution of a perceptual judgment using a compelled-response task. J Neurosci. 2011 Jun 8;31(23):8406-21. doi: 10.1523/JNEUROSCI.1419-11.2011. PMID: 21653845; PMCID: PMC3134312.
15. Maxwell Shinn, Norman H Lam, John D Murray (2020) A flexible framework for simulating and fitting generalized drift-diffusion models eLife 9:e56938
16. Soltani A, Murray JD, Seo H, Lee D. Timescales of Cognition in the Brain. Curr Opin Behav Sci. 2021 Oct;41:30-37. doi: 10.1016/j.cobeha.2021.03.003. Epub 2021 Mar 31. PMID: 34026949; PMCID: PMC8136243.
17. Stanford, T., Shankar, S., Massoglia, D. *et al.* Perceptual decision making in less than 30 milliseconds. *Nat Neurosci* **13**, 379–385 (2010). https://doi.org/10.1038/nn.2485
18. Wald A (1945) Sequential tests of statistical hypotheses. Ann Math Statist. 16(2):117-186. Doi: 10.l1214/aoms/1177731118
19. Yoo SBM, Hayden BY, Pearson JM. Continuous decisions. Philos Trans R Soc Lond B Biol Sci. 2021 Mar;376(1819):20190664. doi: 10.1098/rstb.2019.0664. Epub 2021 Jan 11. PMID: 33423634; PMCID: PMC7815426.