Model parameter fits

cue	coh	ISI	η_1	η_2	$x_{0,1}$	$\sigma_{x_{0,1}}$	$x_{0,2}$	$\sigma_{x_{0,2}}$	$T_{0,1}$	$T_{0,2}$
0.50	0.65	4.00	-0.04	0.37	-0.14	0.53	0.13	0.39	0.54	0.75
		6.00	0.02	0.44	-0.02	0.28	-0.07	0.07	0.45	0.57
		8.00	0.12	0.27	-0.06	0.33	0.12	0.15	0.48	0.68
	0.85	4.00	0.01	0.73	0.03	0.32	0.06	0.19	0.91	0.68
		6.00	0.09	0.44	0.24	0.72	-0.34	0.14	0.46	0.36
		8.00	0.03	1.04	0.15	0.24	-0.21	0.16	0.80	0.61
0.60	0.65	4.00	-0.02	0.44	0.10	0.43	-0.07	0.25	0.52	0.61
		6.00	0.08	0.39	0.43	0.72	-0.07	0.10	0.41	0.59
		8.00	-0.07	0.28	0.07	0.58	0.18	0.24	0.59	0.72
	0.85	4.00	0.25	1.19	0.09	1.58	-0.36	0.32	0.47	0.36
		6.00	0.33	0.56	0.19	1.33	-0.17	0.30	0.48	0.55
		8.00	0.04	0.84	0.13	0.64	-0.10	0.25	0.54	0.64
0.70	0.65	4.00	0.00	0.54	0.18	0.21	0.02	0.29	0.59	0.77
		6.00	0.16	0.46	0.24	0.85	-0.35	0.22	0.53	0.43
		8.00	-0.05	0.59	0.13	0.26	-0.01	0.26	0.57	0.77
	0.85	4.00	0.71	1.29	-0.07	0.82	-0.34	0.38	0.44	0.34
		6.00	0.25	0.72	0.03	1.02	-0.27	0.21	0.50	0.32
		8.00	0.16	0.85	0.16	0.61	0.10	0.16	0.59	0.60
0.80	0.65	4.00	0.73	0.63	0.17	0.79	-0.20	0.57	0.42	0.47
		6.00	0.42	0.44	0.32	0.74	-0.21	0.29	0.41	0.34
		8.00	0.02	0.29	0.25	0.45	0.17	0.20	0.57	0.71
	0.85	4.00	1.13	1.15	-0.05	0.71	-0.32	0.31	0.57	0.31
		6.00	0.41	0.65	0.30	0.80	-0.34	0.43	0.50	0.39
		8.00	0.54	1.26	0.03	0.27	-0.07	0.16	0.45	0.55

Table S1: **Experiment 2: Parameter fits for two, unconnected DDMs (2DDM)**.

cue	coh	ISI	η_1	η_2	$x_{0,1}$	$\sigma_{x_{0,1}}$	$T_{0,1}$
0.50	0.65	4.00	0.09	0.46	0.20	1.80	0.59
		6.00	0.12	0.27	0.39	1.77	0.65
		8.00	-0.00	0.40	-0.38	1.64	0.60
	0.85	4.00	0.40	0.69	0.06	1.91	0.90
		6.00	0.08	0.64	0.29	2.35	0.83
		8.00	-0.08	1.12	0.98	2.40	0.76
0.60	0.65	4.00	0.34	0.28	0.25	1.78	0.62
		6.00	0.11	0.32	0.34	1.83	0.65
		8.00	-0.04	0.47	0.35	1.71	0.64
	0.85	4.00	-0.08	1.06	-0.98	2.51	0.64
		6.00	0.66	0.82	0.16	2.45	0.77
		8.00	0.25	0.76	0.10	3.42	0.59
0.70	0.65	4.00	-0.04	0.60	0.90	2.08	0.91
		6.00	0.45	0.49	0.44	2.12	0.88
		8.00	-0.06	0.74	0.96	1.71	0.99
	0.85	4.00	0.37	1.39	0.23	1.93	0.61
		6.00	0.56	1.38	-0.35	1.52	0.62
		8.00	0.22	1.04	0.53	1.55	0.62
0.80	0.65	4.00	1.07	0.37	0.68	1.87	0.61
		6.00	0.49	0.43	0.41	1.90	0.61
		8.00	0.18	0.43	0.92	1.24	0.63
	0.85	4.00	0.77	1.26	0.42	0.99	0.62
		6.00	0.63	0.84	0.10	1.10	0.61
		8.00	0.92	0.94	0.32	1.34	0.62

Table S2: **Experiment 2: Parameter fits for the Multi-Stage DDM (MS-DDM)**.

cue	coh	ISI	η_1	η_2	$x_{0,1}$	$\sigma_{x_{0,1}}$	$x_{0,2}$	$\sigma_{x_{0,2}}$	$T_{0,1}$	$T_{0,2}$
0.50	0.65	4.00	0.17	0.43	0.09	0.51	-0.33	1.14	0.39	0.42
		6.00	0.11	0.47	0.24	0.66	-0.21	0.18	0.46	0.35
		8.00	0.13	0.35	-0.05	0.30	0.15	0.23	0.50	0.59
	0.85	4.00	0.03	0.66	-0.01	0.32	0.13	0.28	0.77	0.66
		6.00	0.12	0.79	0.20	1.05	-0.34	0.16	0.50	0.39
		8.00	-0.03	1.17	0.16	0.23	-0.14	0.17	0.81	0.67
0.60	0.65	4.00	0.53	0.43	0.22	1.27	-0.34	1.00	0.54	0.64
		6.00	0.22	0.37	0.27	0.76	-0.34	0.12	0.42	0.38
		8.00	-0.07	0.27	0.07	0.58	0.12	0.15	0.59	0.75
	0.85	4.00	0.26	1.20	0.28	1.65	-0.26	0.45	0.44	0.48
		6.00	0.20	0.55	0.03	0.94	-0.11	0.20	0.44	0.41
		8.00	-0.03	0.82	0.11	0.45	0.05	0.19	0.50	0.60
0.70	0.65	4.00	0.09	0.74	0.17	1.32	-0.38	1.01	0.59	0.47
		6.00	0.25	0.58	0.23	1.06	-0.35	0.45	0.68	0.55
		8.00	0.11	0.15	0.40	1.59	-0.30	0.33	0.51	0.31
	0.85	4.00	0.69	1.31	0.03	0.99	-0.32	0.36	0.48	0.37
		6.00	0.25	0.72	0.03	1.02	-0.27	0.21	0.50	0.32
		8.00	0.16	0.85	0.16	0.61	0.10	0.16	0.59	0.60
0.80	0.65	4.00	0.60	0.40	0.18	0.48	0.16	0.17	0.55	0.62
		6.00	0.47	0.28	0.21	0.91	-0.21	0.76	0.41	0.68
		8.00	0.20	0.29	0.14	0.26	0.10	0.12	0.47	0.67
	0.85	4.00	0.82	0.97	0.11	0.57	0.09	0.21	0.60	0.51
		6.00	0.46	0.84	0.10	0.49	-0.04	0.20	0.56	0.55
		8.00	0.67	0.96	-0.03	0.21	-0.02	0.17	0.45	0.49

Table S3: **Combined dataset: Parameter fits for two, unconnected DDMs (2DDM)**.

cue	coh	ISI	η_1	η_2	$x_{0,1}$	$\sigma_{x_{0,1}}$	$T_{0,1}$
0.50	0.65	4.00	-0.07	0.41	-0.86	1.65	0.63
		6.00	0.26	0.41	0.39	1.81	0.63
		8.00	0.14	0.50	-0.84	1.82	0.63
	0.85	4.00	0.00	0.85	0.88	2.62	0.68
		6.00	0.25	0.67	0.08	2.55	0.75
		8.00	-0.09	1.29	0.95	2.89	0.81
0.60	0.65	4.00	0.18	0.35	0.25	1.78	0.68
		6.00	0.23	0.43	0.28	1.83	0.73
		8.00	-0.05	0.38	0.77	1.81	0.67
	0.85	4.00	-0.10	0.97	-0.99	3.38	0.54
		6.00	0.47	0.93	0.06	3.43	0.64
		8.00	0.25	0.82	0.13	3.22	0.60
0.70	0.65	4.00	-0.04	0.67	0.97	1.78	0.89
		6.00	0.18	0.56	0.15	3.11	0.64
		8.00	0.01	0.41	0.99	2.49	0.86
	0.85	4.00	0.36	1.23	0.21	2.11	0.64
		6.00	0.56	1.38	-0.35	1.52	0.62
		8.00	0.22	1.04	0.53	1.55	0.62
0.80	0.65	4.00	0.92	0.44	0.45	1.85	0.63
		6.00	0.71	0.41	0.23	1.94	0.58
		8.00	0.39	0.35	0.29	1.59	0.64
	0.85	4.00	0.98	0.81	-0.23	1.32	0.58
		6.00	0.59	0.97	0.04	1.52	0.62
		8.00	0.71	0.96	0.76	1.37	0.66

Table S4: **Combined dataset: Parameter fits for the Multi-Stage DDM (MS-DDM).**

Parameter-setting correlations

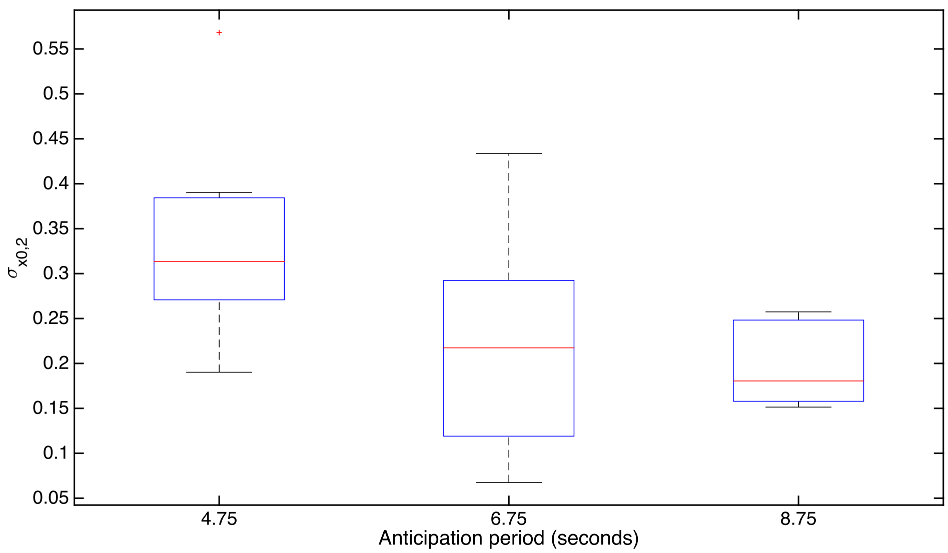

Figure S1: **Second-stage starting point variability decreases with ISI**.

Second-stage starting point variability decreases with ISI

Although the *2DDM* was disfavored in our model comparison, the fitted parameter values can provide insight as to the nature of the second-stage inference process. Specifically, we investigated whether the variability in second-stage starting point decreased with the length of the preceding anticipation period. If, during this anticipation period, successive samples were drawn, and the resulting estimate carried forward to the second stage, then the resulting random variable should be of lower variance. If, on the other hand, first-stage evidence were due to one or a fixed number of samples, or not carried forward to the second stage, then the resulting second-stage starting point variance should be unaffected by the length of the anticipation period. Consistent with the hypothesis that second-stage starting point is set by a process that continues to add samples during the entire anticipation period, we found that, across all combinations of cue level and coherence, second-stage starting point variability decreased with ISI (Experiment 1: $R = -.XXX$, $P = .XXX$; Experiment 2: $R = -.512$, $P = .011$; Combined dataset: $R = -.530$, $P = .008$).

Coherence gates the effect of cue on starting point, not drift rate

First-stage starting point was set by cue probability *only* when upcoming sensory evidence would be weak (low-coherence: $R = 0.699$, $P = 0.011$; high-coherence: $R = -0.028$, $P = 0.931$).