Fast saccades toward numbers: Simple number comparisons can be made in as little as 230 ms

Milica Milosavljevic
Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA

Erik Madsen
Division of Humanities and Social Sciences, Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA

Christof Koch
Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA, & Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea

Antonio Rangel
Division of Humanities and Social Sciences, Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA

Visual psychophysicists have recently developed tools to measure the maximal speed at which the brain can accurately carry out different types of computations (H. Kirchner & S. J. Thorpe, 2006). We use this methodology to measure the maximal speed with which individuals can make magnitude comparisons between two single-digit numbers. We find that individuals make such comparisons with high accuracy in 306 ms on average and are able to perform above chance in as little as 230 ms. We also find that maximal speeds are similar for "larger than" and "smaller than" number comparisons and in a control task that simply requires subjects to identify the number in a number–letter pair. The results suggest that the brain contains dedicated processes involved in implementing basic number comparisons that can be deployed in parallel with processes involved in low-level visual processing.

Keywords: minimum reaction time, decision making, mental number line, number comparison, eye tracking

Citation: Milosavljevic, M., Madsen, E., Koch, C., & Rangel, A. (2011). Fast saccades toward numbers: Simple number comparisons can be made in as little as 230 ms. *Journal of Vision, 11*(4):4, 1–12, http://www.journalofvision.org/content/11/4/4, doi:10.1167/11.4.4.

Introduction

A basic question in cognitive psychology and neuroscience is how quickly different types of computations can be performed, while maintaining a target level of accuracy. This question is important because its answer provides useful hints as to the type and amount of processing required in different tasks, as well as the extent to which computations are performed by specialized low-level systems versus more general high-level systems. Here, we are interested in the speed with which individuals are able to accurately make basic number comparisons, such as identifying which of two single-digit numbers is larger. Number comparisons are important because they are linked to decision making. For example, it has been shown that the precision of one's intuitive number sense is correlated with the ability to save (Peters, Slovic, Vastfjall, & Mertz, 2008).

A sizable number of studies have investigated the processing and comparison of numerical quantities. In a classic study, Moyer and Landauer (1967) established the *distance effect* (Dehaene, 1997; Dehaene & Changeux, 1993) in binary numerical comparisons, which states that accuracy increases and response times decrease with the absolute distance between the two numbers being compared. This effect has been replicated in a number of studies (Dehaene, 1989, 1996) and has motivated the "mental number line hypothesis," which states that numeric magnitude is spatially encoded in the brain (Dehaene, Bossini, & Giraux, 1993; Moyer & Landauer, 1967). Song and Nakayama (2008) investigated this

doi: 10.1167/11.4.4 Received September 17, 2010; published April 6, 2011 ISSN 1534-7362 © ARVO

hypothesis using a visually guided manual reaching task in which subjects were asked to compare a single digit number displayed at the center of the screen with the number "5" by reaching with their index finger toward the appropriate location on the screen (left, center, or right of the displayed number to signify smaller, equal, or larger, respectively). They found that the greater the numeric deviation from 5, the greater the deviation of the pointing trajectory.

More recent work building on these ideas has shown that random walk models of binary responses, such as the Drift Diffusion Model, account well for the pattern of response and reaction times in such tasks (Bogacz, 2007; Busemeyer & Johnson, 2004; Dehaene, 2007; Gold & Shadlen, 2007; Krajbich, Armel, & Rangel, 2010; Link & Heath, 1975; Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010; Ratcliff, 1978; Ratcliff & McKoon, 2008; Sigman & Dehaene, 2005; Usher & McClelland, 2001). See especially Sigman and Dehaene's (2005) effort to dissect the comparison task into subcomponents (perceptive, decision, motor) and evaluate the timing of each of these steps.

This previous literature has shown that number comparisons can be made quickly. Moyer and Landauer (1967) found reaction times that varied from about 500 ms for a numerical distance of 8 to about 620 ms for a numerical distance of 1. Unlike Moyer and Landauer (1967), most of the literature on number comparisons does not display two digits simultaneously but rather displays a single digit and asks subjects to identify whether the number is larger or smaller than a fixed comparison number that does not vary across trials (Dehaene, 1996; Libertus, Woldorff, & Brannon, 2007). Using this alternative paradigm, Dehaene (1996) reported behavioral reaction times of about 370 ms and 390 ms for large and small numerical distances, respectively. In a similar task, Libertus et al. (2007) reported mean reaction times ranging from about 380 ms for large numerical distances to about 430 ms for small numerical distances. Using the manual reaching task described above, Song and Nakayama (2008) found mean reaction times of 338 ± 42 ms at a numerical distance of 1, which decreased to 313 ± 33 ms at a numerical distance of 4.

However, the existing literature does not provide measures of the minimum response times (MRTs) at which reliable comparisons can be made. This is an important shortcoming because MRTs are thought to provide critical clues about the level of processing (e.g., bottom-up vs. top-down) involved in different computations. Further, there is reason to believe that mean RTs might be significantly larger than MRTs in numerical comparisons. For instance, event-related brain potentials (ERPs) related to the numerical distance effect have been found in parietal sensors about 200 ms after stimulus onset (Pinel, Dehaene, Riviere, & LeBihan, 2001; Temple & Posner, 1998).

Here, we apply a novel methodology from visual psychophysics to provide a better behavioral measurement of the fastest speed at which number comparisons can be made. Over the last decade, visual psychophysicists have developed a useful set of tools to measure a lower threshold for the time required for successful task performance (Bannerman, Milders, de Gelder, & Sahraie, 2009; Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001; Kirchner & Thorpe, 2006; VanRullen & Thorpe, 2001a, 2001b). The application of these methods to perceptual decision making has shown that individuals can make high-accuracy perceptual discriminations extremely rapidly. For example, Kirchner and Thorpe (2006) found that, on average, subjects could identify which of two natural scenes flashed to the left and right hemifields contained an animal in 228 ms. Utilizing a similar paradigm, Bannerman et al. (2009) showed that subjects could distinguish a fearful facial expression or body posture from a neutral one in 350 ms on average. Most recently, it has been shown that forced-choice saccades to identify human faces can be performed above chance and initiated with a mean reaction time of just 154 ms (Crouzet, Kirchner, & Thorpe, 2010).

In our experiments, we flash pairs of single-digit numbers briefly (20 ms) to the left and right hemifields and ask subjects to saccade to the location of the larger number either as fast as possible or after careful deliberation (Experiment 1). The use of brief presentations and saccadic responses is useful because it speeds up motor responses, thus reducing the total fraction of reaction times that are unrelated to the computation of interest. In Experiment 2, we ask subjects to identify the *smaller* of the two digits under speed instructions in order to investigate if it is equally easy for the brain to make rapid "larger than" and "smaller than" number comparisons. Finally, in Experiment 3, we compare these results to a number identification control task in which subjects are shown a letter–number pair and asked to identify the location of the number.

Experiment 1: "Larger than" number comparisons in speed vs. accuracy conditions

Experiment 1 examined the basic psychometrics of single-digit number comparisons. The design was based on Kirchner and Thorpe (2006; henceforth "KT 2006"). On each trial, two single-digit numbers were flashed on the screen for 20 ms, in the left and right hemifields. Subjects were asked to make a saccade to the location of the larger number. The task was performed in a speed condition, in which subjects were instructed to respond as quickly as possible, and in an accuracy condition, in which they were asked to respond only after they were sure of the correct choice.

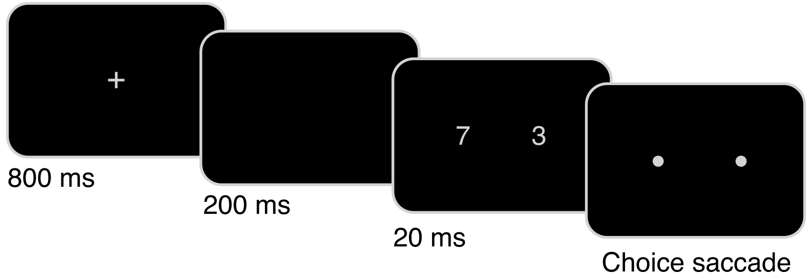

Figure 1. Sample trial from Experiment 1.

Methods

Participants

Twelve Caltech students with normal or corrected-to-normal vision participated in two 1-h sessions and were paid $20 per session. All participants gave written informed consent to participate. The study and experimental procedures were approved by the California Institute of Technology Institutional Review Board.

Stimuli

Subjects were shown two single-digit numbers (ranging from 0 to 9) in each trial (Figure 1). Stimuli were positioned at 5° to the left and right of the center of the screen. Each character had dimensions of 1.3° × 0.7°.

Apparatus

Participants were seated in a dimly lit room with head position controlled by forehead and chin rests. Eye-position data were acquired from the right eye at 1000 Hz using the Eyelink 1000 eye tracker (SR Research, Osgoode, Canada). The distance between the computer screen and subject was 80 cm. Images were presented on a CRT screen, using the MATLAB Psychophysics toolbox and Eyelink toolbox extensions (Brainard, 1997; Cornelissen, Peters, & Palmer, 2002). Responses (saccades toward the left or right option) were recorded when a saccade was initiated and crossed a threshold of 2.2° horizontal distance from the center of the screen. Saccadic reaction time was determined by computing the time difference between stimulus onset and saccade initiation.

Task

Each trial began by requiring subjects to fixate continuously on a faint gray cross in the center of the screen for at least 800 ms (Figure 1). Afterward, the fixation cross disappeared, leaving the screen blank for 200 ms. This gap period served to accelerate saccade initiation (Fischer & Weber, 1993). Two single-digit numbers (ranging from 0 to 9) were then shown simultaneously for 20 ms, each positioned at 5° from center in the left and right hemifields. Two faint dots were next displayed at the previous locations of the two numbers to indicate the corresponding choice options. The trial ended when subjects made a saccade to one side of the screen to indicate their choice. To minimize learning, pairs were constructed so as not to repeat the same unordered combination of numbers in the same 45-trial set. For example, if "2–3" was included in a set, "3–2" was not included in the same set.

Subjects completed 900 trials in each of two experimental conditions: a speed condition in which they were asked to respond as quickly as possible and an accuracy condition in which they were asked to indicate their response only after they were sure of the correct choice. The data for each condition were collected on different days and task order was counterbalanced across subjects so that six subjects completed the speed condition first.

Minimum reaction times

It has become standard practice when analyzing the results of ultra-rapid decision tasks to compute a measure of the minimum reaction time (MRT) at which subjects began responding with above-chance accuracy (Crouzet et al., 2010; Kirchner & Thorpe, 2006). The typical MRT analysis sorts reaction times in increasing order, divides them into discrete bins, and then tests the average accuracy of observations in each bin against a null hypothesis of chance performance, with the null hypothesis rejected only when above-chance results are observed in a certain number of consecutive bins. This procedure has raised some controversy because the resulting MRT measures are highly sensitive to bin width and placement, as well as to the number of consecutive above-chance bins required for significance.

In order to avoid this problem, we computed MRTs using a more robust method taken from the statistical quality control literature (Chandra, 2001; Roberts, 1959). The method has also been used to filter "fast guesses" from perceptual data when fitting the Ratcliff Drift Diffusion Model (Ratcliff & McKoon, 2008; Vandekerckhove & Tuerlinckx, 2007). The basic idea is given as follows.

Observations are ordered from low to high response times. Let X_i denote the accuracy of the ith ordered response (1 = correct, 0 = incorrect). An exponentially weighted moving average (EWMA) measure of accuracy is then computed using the following formula:

$$\mathrm{EWMA}_i = \lambda X_i + (1-\lambda)\mathrm{EWMA}_{i-1}, \qquad (1)$$

where λ is a parameter indicating how much weight to give to past (ordered) observations in the moving average. Note that when $\lambda = 1$, the EWMA statistic is based only on the most recent observation. In contrast, as λ approaches 0 previous observations are given increasing weight relative to the latest observation. EWMA_0 was set to 0.5 (i.e., chance performance).

Intuitively, the EWMA measure provides an estimate of how accuracy changes with increasing reaction times. This measure can then be compared against the null hypothesis of chance performance on all trials. This null hypothesis generates a confidence interval for the EWMA after i observations given by

$$\mu \pm N\sigma\sqrt{\frac{\lambda}{2-\lambda}\left(1-\left(1-\lambda\right)^{2i}\right)}. \qquad (2)$$

The first term in this expression is the mean EWMA statistic under the null hypothesis, which in our experiment is given by $\mu = E(X_i) = 0.5$. The second term provides an expression for the width of the confidence interval: N is the number of standard deviations included in the confidence interval, and $\sigma = \mathrm{Std}(X_i) = 0.5$ is the standard deviation of each observation X_i under the null hypothesis. Note that under the null hypothesis performance in every trial depends on the independent flip of a fair coin.

The MRT can then be defined as the smallest ordered reaction time at which the EWMA measure permanently exceeds this confidence interval. For our analyses, we chose conservative parameters to reduce the possibility of false positives: $\lambda = 0.01$ and $N = 3$. The EWMA analysis was run separately for each subject and condition.

Results and discussion

Figure 2 depicts the reaction time distributions for both conditions (speed and accuracy) for correct and error trials. Table 1 provides individual information on accuracy and response times. To screen for anomalous trials, the fastest and slowest 0.5% of trials pooled across all subjects were discarded. Average response times were 306 ± 15 ms (mean ± *SEM*) for the speed and 345 ± 15 ms for the accuracy condition. The average per-subject difference in reaction times between the two conditions (accuracy–speed) was 39 ± 11 ms (two-sided paired t-test, $p = 0.0036$). Average accuracies were 91.2 ± 1.0% in the speed condition and 95.4 ± 1.0% in the accuracy condition. The average per-subject accuracy difference between the two conditions (accuracy–speed) was 4.3 ± 1.1% (two-sided paired t-test, $p = 0.0030$). As expected, subjects were faster and less accurate in the speed condition.

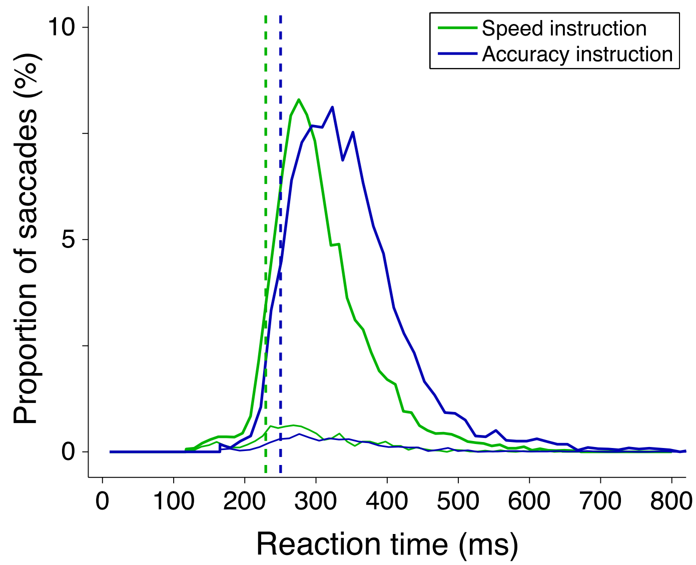

Figure 2. Reaction time distributions for the speed and accuracy conditions in Experiment 1. Correct trials = thick line. Error trials = thin line. Vertical lines depict mean MRTs.

Mean MRT was 230 ± 11 ms in the speed condition and 250 ± 11 ms in the accuracy condition. The increase in MRT from the speed to the accuracy condition was small but statistically significant (20.8 ± 7.1 ms; two-sided paired t-test, $p = 0.014$). For better comparison with the previous literature, we also computed the binned MRT measure of KT 2006, which is computed using the pooled data. We chose 10-ms bins centered at 10 ms, 20 ms, etc., and a threshold of 10 consecutive bins at a 5% confidence level. The resulting MRT statistic was 170 ms (i.e., the 165- to 175-ms bin) for the speed task and 190 ms for the accuracy task. This result shows that our MRT measures are more conservative than the KT measures and give more confidence that the average subject is indeed performing at above-chance levels at the reported MRT values.

Next, we tested for the presence of a numerical distance effect, as shown in Figure 3. For reaction times, we fitted a linear regression of reaction times on the absolute number difference for each subject. For accuracy, we fitted a logistic regression of an indicator variable of correct responses on the absolute number difference for each subject. The results are shown in Table 2. The mean fitted RT slope across subjects was −3.1 ± 0.8 ms per unit (two-sided t-test, $p = 0.0021$) in the speed task and −5.6 ± 0.7 ms per unit (two-sided t-test, $p = 5.9 \times 10^{-6}$) in the accuracy task. The mean fitted accuracy slope was 0.53 ±

S	*N*	Accuracy (%)	Mean RT (ms) ± *SEM*	MRT
Experiment 1: Speed condition				
1*	882	86.8	409 ± 2.16	237
2*	897	91.8	310 ± 1.74	248
3*	896	90.7	265 ± 1.51	200
4*	886	94.7	277 ± 1.72	200
5*	898	89.4	272 ± 1.58	164
6*	891	90.7	261 ± 1.39	205
7	878	86.4	285 ± 2.65	220
8	889	97.6	401 ± 2.42	303
9	896	93.8	297 ± 1.19	244
10	887	94.5	295 ± 1.38	239
11	894	86.8	268 ± 1.26	221
12	898	90.5	332 ± 1.39	273
All	10692	91.2 ± 1.0	306 ± 15	230 ± 11
Experiment 1: Accuracy condition				
1*	897	97.4	405 ± 2.30	310
2*	900	97.9	366 ± 2.50	244
3*	895	97.9	311 ± 2.35	235
4*	891	96.7	292 ± 1.90	221
5*	893	95.5	290 ± 1.50	220
6*	900	95.3	271 ± 1.53	202
7	870	97.0	420 ± 3.11	209
8	893	97.2	413 ± 3.16	312
9	897	96.0	317 ± 1.57	255
10	896	97.5	357 ± 2.32	258
11	892	86.7	318 ± 1.82	247
12	865	90.1	385 ± 3.21	290
All	10689	95.4 ± 1.0	345 ± 15	250 ± 11

Table 1. Individual mean accuracy, mean RT, and minimum RT (MRT) in Experiment 1. *Note:* *Denotes that subject completed the speed condition first.

0.08 (two-sided *t*-test, $p = 2.6 \times 10^{-5}$) in the speed task and 0.60 ± 0.12 (two-sided *t*-test, $p = 4.0 \times 10^{-4}$) in the accuracy task. Thus, in both cases, responses became faster and more accurate as a function of numerical distance, confirming the presence of a numerical distance effect in our experimental setup.

Finally, we looked for an effect of the order in which the two tasks were performed on accuracy and RTs. We found no significant order effects either in the speed condition results (speed condition first: accuracy = 90.7 ± 1.1% and RT = 299 ± 23 ms; accuracy condition first: accuracy = 91.6 ± 1.8% and RT = 313 ± 20 ms; difference: accuracy = −0.92 ± 2.1%, $p = 0.67$ and RT = −14 ± 30 ms, $p = 0.65$) or the accuracy condition results (speed condition first: accuracy = 96.8 ± 0.5% and RT = 323 ± 21 ms; accuracy condition first: accuracy = 94.1 ± 1.9% and RT = 368 ± 18 ms; difference: accuracy = 2.7 ± 1.9%, $p = 0.19$ and RT = −46 ± 28 ms, $p = 0.13$).

The results of Experiment 1 are consistent with the hypothesis that the brain is able to carry out automatic, rapid, and accurate number comparisons between single digit numbers with little top-down control. In particular, we found that deliberate efforts to increase accuracy slowed responses very little and provided relatively small accuracy improvements over speeded responses.

Experiment 2: "Smaller than" number comparisons

Experiment 2 is nearly identical to the speed condition of Experiment 1, except that subjects were asked to saccade as quickly as possible to the location of the *smaller* number. The goal was to investigate if the brain could execute "larger than" and "smaller than" number comparisons at similar maximal speeds.

Methods

Methods are almost identical to those for the speed condition of Experiment 1 and, thus, are omitted. Twelve new subjects participated in the task.

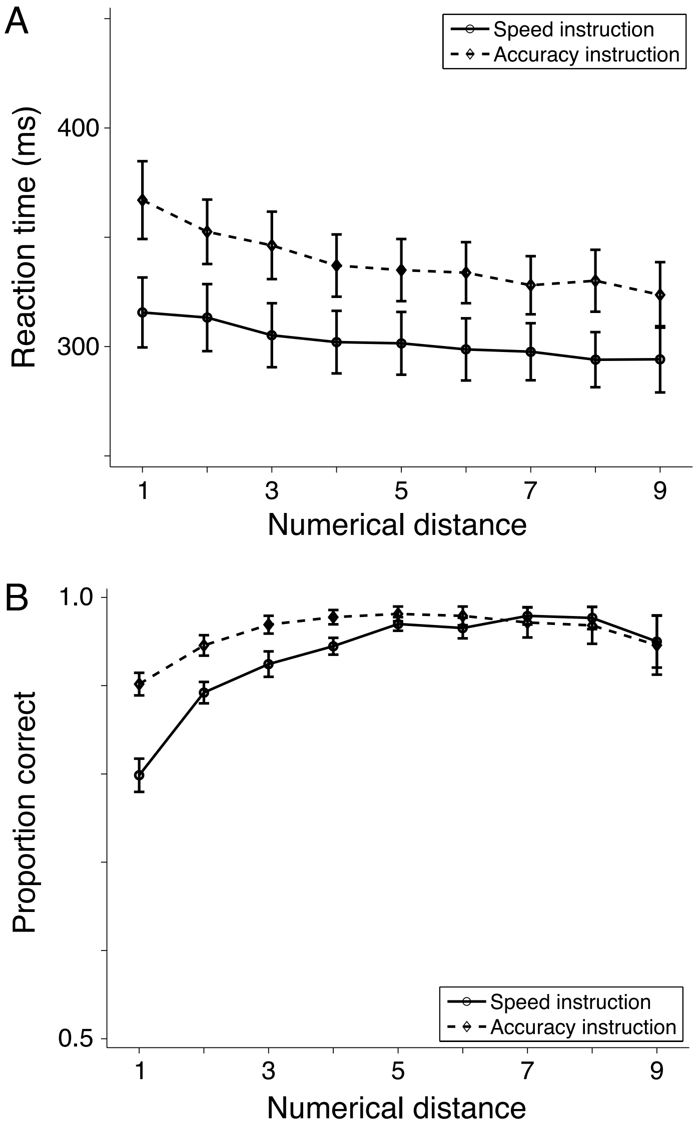

Figure 3. Mean reaction time and accuracy as a function of numerical distance for Experiment 1. Bars denote *SEM* across subjects.

Results and discussion

We compared the results of Experiment 2 with the speed condition of Experiment 1. Figure 4 depicts the reaction time distribution for correct and error trials in Experiments 2 and Experiment 1 (speed condition), and Table 3 provides individual information on accuracy, mean RT, and minimum RT. The mean RT for the task across all subjects was 322 ± 10 ms. The difference in mean RT between experiments was not significant (Experiment 2 − Experiment 1 speed: 15.7 ± 17.4 ms; two-sided *t*-test, $p = 0.38$). The mean accuracy for the task across all subjects was 84.6 ± 1.9%. The difference in mean accuracy was statistically significant (Experiment 2 − Experiment 1 speed: −6.6 ± 2.2%; two-sided *t*-test, $p = 0.0060$).

The mean per-subject estimated MRT was 247 ± 11 ms (Table 3; MRT). Comparing this result with the average per-subject MRT from the speed condition of Experiment 1, we found that the difference in average MRT was 17.5 ± 15.0 ms (Experiment 2 − Experiment 1 speed; two-sided *t*-test, $p = 0.255$).

As before, we estimated regression models of the impact of numerical distance on reaction time and accuracy, as shown in Figure 5 and Table 4. We compared these estimates with those for the speed condition of Experiment 1 to investigate the differences between the two tasks. We found no difference in the average constant of either regression (two-sided *t*-test; RT: $p = 0.61$, accuracy: $p = 0.65$), even though at each numerical distance the estimated accuracy was higher in the "larger than" condition than in the "smaller than" condition. There was a significant difference in the slope of the accuracy curve (Experiment 2 − Experiment 1 speed: −0.25 ± 0.08; two-sided *t*-test, $p = 0.0064$), though not in the slope of the RT curve (Experiment 2 − Experiment 1 speed: 1.66 ± 0.89 ms; two-sided *t*-test, $p = 0.08$).

A two-way ANOVA reveals no dependency on numerical distance ($p = 0.92$) but a significant difference between the two conditions ($p = 0.0029$) with an estimated set size of 17.7 ms with no interaction between the two ($p = 0.99$).

The results show that "larger than" and "smaller than" number comparisons have similar psychometric performance. The difference in MRT between Experiment 2 and the speed condition of Experiment 1 was not significant as measured by RT and MRT. However, a two-way ANOVA on the aggregated data was significant for the difference in reaction times between the two experiments. Note that the size of the effect in all three comparisons is similar, a bit under 20 ms. This is quantitatively similar to the difference between the accuracy and speed conditions for Experiment 1. In fact, the mean estimates for Experiment 2 and the accuracy condition for Experiment 1 are nearly indistinguishable. Thus, our results suggest that the MRTs for both types of comparison are very similar.

Experiment 3: Number identification control

One concern with our experiments is that they might overestimate the amount of time required to make number comparisons because, in addition to making a comparison, they require subjects to identify the two stimuli and initiate a motor response. We address this concern by comparing performance in the speed condition of Experiment 1 with performance in a closely related control task

S	RT constant (ms/unit) ± *SEM*	RT slope (ms/unit) ± *SEM*	Accuracy constant ± *SEM*	Accuracy slope ± *SEM*
Experiment 1: Speed condition				
1*	414 ± 18	−1.3 ± 1.0	0.82 ± 0.03	0.340 ± 0.003
2*	321 ± 11	−3.3 ± 0.6	0.59 ± 0.06	0.69 ± 0.01
3*	275 ± 8.4	−2.8 ± 0.5	0.52 ± 0.05	0.659 ± 0.009
4*	289 ± 11	−3.3 ± 0.6	0.93 ± 0.09	0.79 ± 0.02
5*	271 ± 9.3	0.3 ± 0.5	0.85 ± 0.04	0.435 ± 0.005
6*	273 ± 7.1	−3.2 ± 0.4	0.81 ± 0.05	0.514 ± 0.007
7	295 ± 26	−2.7 ± 1.4	0.97 ± 0.03	0.273 ± 0.003
8	441 ± 20	−10.9 ± 1.1	1.8 ± 0.2	0.76 ± 0.04
9	310 ± 5.1	−3.34 ± 0.3	0.88 ± 0.07	0.72 ± 0.02
10	306 ± 6.9	−3.0 ± 0.4	0.79 ± 0.09	0.84 ± 0.02
11	274 ± 5.9	−1.5 ± 0.3	0.60 ± 0.03	0.428 ± 0.004
12	339 ± 7.1	−2.0 ± 0.4	2.55 ± 0.05	−0.076 ± 0.003
All	317 ± 16	−3.1 ± 0.8	1.0 ± 0.2	0.53 ± 0.08
Experiment 1: Accuracy condition				
1*	442 ± 18	−10.0 ± 1.0	1.6 ± 0.2	0.85 ± 0.05
2*	382 ± 23	−4.1 ± 1.3	2.3 ± 0.2	0.57 ± 0.03
3*	330 ± 20	−5.3 ± 1.1	1.4 ± 0.25	1.12 ± 0.09
4*	311 ± 13	−5.2 ± 0.7	1.8 ± 0.1	0.61 ± 0.02
5*	306 ± 8.1	−4.2 ± 0.4	1.82 ± 0.09	0.43 ± 0.01
6*	286 ± 8.5	−4.0 ± 0.5	0.4 ± 0.1	1.24 ± 0.05
7	449 ± 35	−8.1 ± 1.9	2.4 ± 0.1	0.36 ± 0.02
8	450 ± 35	−10.1 ± 1.9	2.3 ± 0.1	0.46 ± 0.02
9	331 ± 9.0	−3.6 ± 0.5	0.9 ± 0.1	0.98 ± 0.04
10	374 ± 20	−4.7 ± 1.1	1.9 ± 0.2	0.70 ± 0.04
11	331 ± 12	−3.7 ± 0.7	1.58 ± 0.03	0.084 ± 0.002
12	401 ± 39	−4.3 ± 2.1	2.99 ± 0.06	−0.192 ± 0.002
All	366 ± 17	−5.6 ± 0.7	1.8 ± 0.2	0.6 ± 0.1

Table 2. Individual parameter fits for the reaction time and accuracy curves in Experiment 1. *Note:* *Denotes that subject completed the speed condition first.

(Experiment 3) that involves number identification and motor response but not number comparison. This provides an additional estimate of the MRTs associated exclusively with the number comparison component of the task.

Methods

Methods are nearly identical to those for Experiment 1, and only the differences are highlighted here. Twelve new subjects participated in the task by completing 600 trials each. Stimuli were randomly selected pairs consisting of one single-digit number (ranging from 0 to 9) and one capital letter from a set of nine. The letters used in the experiment were chosen to be easily distinguishable from single-digit numbers: X, V, F, T, J, K, L, U, Y. Subjects were instructed to saccade to the location of the number as quickly as possible.

Results and discussion

Figure 6 depicts reaction time distributions for Experiment 3 and for the speed condition of Experiment 1, for

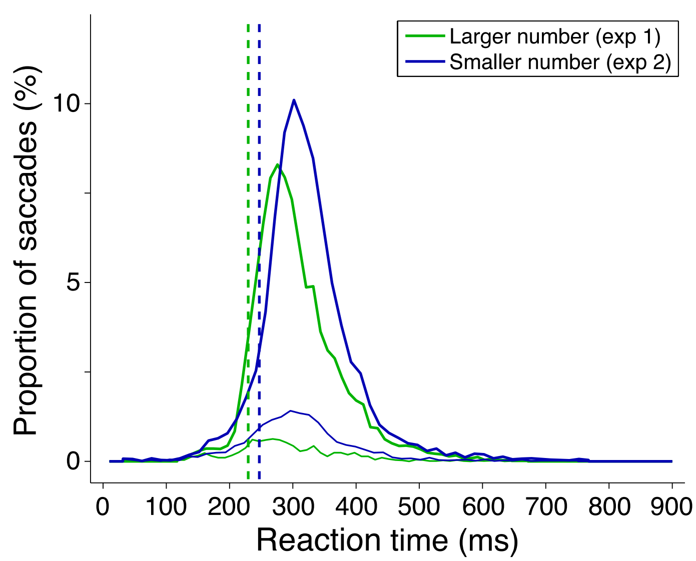

Figure 4. Reaction time distributions for Experiment 2 and Experiment 1 (speed condition). Correct trials = thick line. Error trials = thin line. Vertical lines depict mean MRTs.

Experiment 2: "Smaller than" comparisons

S	*N*	Accuracy (%)	Mean RT (ms) ± *SEM*	MRT
13	900	86.2	301 ± 1.4	213
14	885	85.3	334 ± 3.0	184
15	893	86.7	389 ± 2.5	295
16	847	70.7	289 ± 4.2	275
17	899	91.7	336 ± 2.1	225
18	898	87.6	314 ± 1.9	259
19	899	95.1	314 ± 1.7	250
20	896	83.7	315 ± 1.9	250
21	883	86.5	296 ± 2.7	203
22	898	77.4	282 ± 1.9	242
23	893	77.5	378 ± 2.9	309
24	900	86.4	310 ± 1.7	259
All	5344	84.6 ± 1.9	322 ± 10	247 ± 11

Table 3. Individual mean accuracy, mean RT, and minimum RT (MRT) in Experiment 2.

correct and error trials. Table 5 provides individual information on accuracy, mean RT, and MRT. The mean RT across subjects was 262 ± 5 ms, which was significantly faster than in the "larger than" number comparison task (Experiment 3 − Experiment 1 speed: −43.6 ± 15.4 ms; two-sided *t*-test, $p = 0.0096$). The mean accuracy for the task across subjects was 89.1 ± 1.8%, which was not significantly different from the number comparison task (Experiment 3 − Experiment 1 speed: −2.1 ± 2.1%; two-sided *t*-test, $p = 0.33$).

The mean per-subject MRT was 210 ± 5 ms (Table 5), which was not significantly different from the number comparison task at conventional confidence levels (Experiment 3 − Experiment 1 speed: −19.2 ± 11.6 ms; two-sided *t*-test, $p = 0.11$). In other words, there is no evidence that comparing two numbers and choosing the larger one takes longer than identifying a number over a letter.

As before, these results show that the estimated difference in MRTs between two tasks with similar perceptual and motor demands, which only differ in the need to make number comparisons, is on the order of 20 ms. This finding suggests that number comparisons can be made rapidly and in parallel with low-level visual processing and identification, as any other model is inconsistent with our finding that the additional computational burden of number comparisons adds very little additional processing time to a basic identification and response initiation task.

General discussion

Our results show that individuals make basic binary number comparisons with high accuracy in 306 ms on average and are able to perform above chance in as little as 230 ms, that maximal speeds are similar for "larger than" and "smaller than" numerical comparisons, and that they are also similar in a control task that simply requires subjects to identify the number in a number–letter pair.

The results suggest that the brain contains dedicated processes involved in implementing basic number comparisons that can be deployed in parallel with processes involved in low-level visual processing. Such ultra-rapid responses are generally believed to rely on parallel feed-forward processing of objects in different regions of the visual field by the early visual pathways (VanRullen, 2007).

It is useful to compare the psychometric properties of number comparisons to those of other processes that have been investigated in the literature using similar methods. Thorpe, Fize, and Marlot (1996) showed that natural scenes could be rapidly categorized according to whether

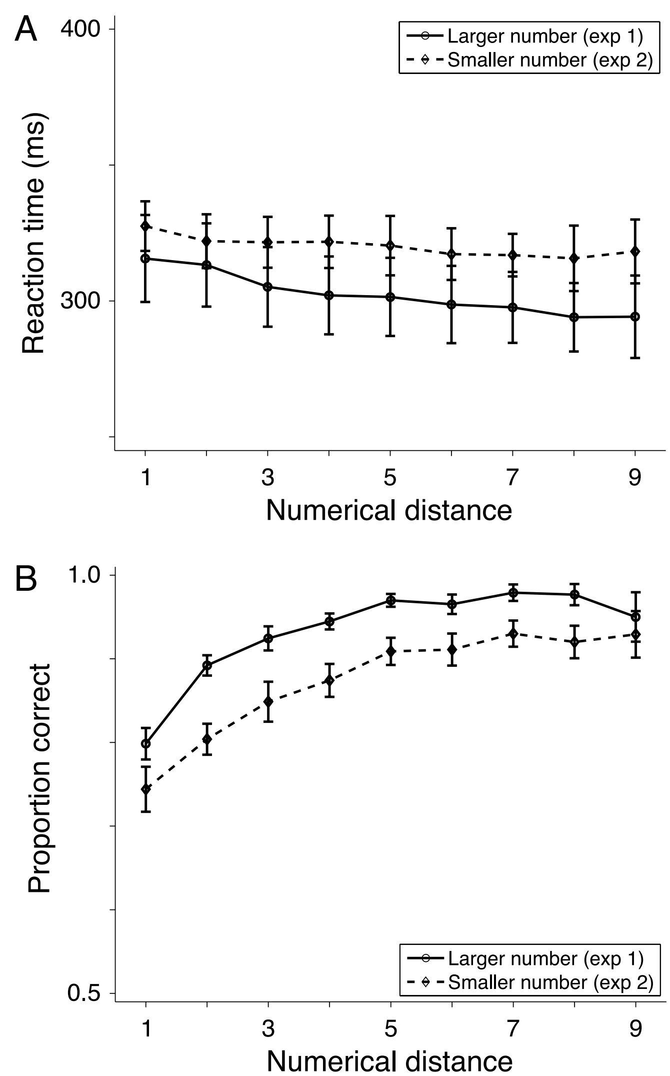

Figure 5. Mean reaction time and accuracy across subjects for Experiment 2 and Experiment 1 (speed condition). Bars denote *SEM* across subjects.

Experiment 2: "Smaller than" comparisons

S	RT constant (ms/unit) ± *SEM*	RT slope (ms/unit) ± *SEM*	Accuracy constant ± *SEM*	Accuracy slope ± *SEM*
13	302 ± 7.8	−0.4 ± 0.4	0.58 ± 0.03	0.417 ± 0.004
14	334 ± 32	0.06 ± 1.7	0.89 ± 0.03	0.269 ± 0.003
15	398 ± 23	−2.2 ± 1.3	1.43 ± 0.03	0.127 ± 0.002
16	298 ± 67	−2.2 ± 3.6	0.28 ± 0.02	0.174 ± 0.001
17	330 ± 17	1.6 ± 0.9	0.94 ± 0.05	0.518 ± 0.008
18	322 ± 13	−2.1 ± 0.7	0.78 ± 0.04	0.387 ± 0.004
19	329 ± 11	−3.9 ± 0.6	2.00 ± 0.08	0.317 ± 0.008
20	326 ± 14	−3.0 ± 0.8	0.94 ± 0.03	0.209 ± 0.002
21	299 ± 27	−0.7 ± 1.5	0.98 ± 0.03	0.275 ± 0.003
22	292 ± 14	−2.6 ± 0.7	0.34 ± 0.02	0.268 ± 0.002
23	379 ± 33	−0.3 ± 1.8	0.52 ± 0.02	0.210 ± 0.002
24	316 ± 10	−1.5 ± 0.6	1.28 ± 0.03	0.170 ± 0.002
All	327 ± 9.3	−1.4 ± 0.4	0.9 ± 0.1	0.28 ± 0.03

Table 4. Individual parameter fits for the reaction time and accuracy curves in Experiment 2.

or not they contain an animal in a go/no go task using button-press responses (median RT of 445 ms on "go" trials and differential ERP activity in 150 ms). VanRullen and Thorpe (2001a, 2001b) found that animals and vehicles could be categorized in similar time frames in a go/no go task, indicating no preference for biologically relevant stimuli (mean RT of 364 ms for animals, 376 ms for vehicles; minimum RT of 225 ms for animals, 245 ms for vehicles as measured by earliest above-chance responses; differential ERP activity detected in 150 ms for both tasks). Training provided no advantage: natural scenes to which subjects had been previously exposed over a 3-week period could be categorized according to whether they contained an animal as quickly as an entirely novel set of stimuli (while mean RTs were 424 ms for familiar and 444 ms for novel stimuli, this discrepancy was due to elimination of very long reaction times for familiar stimuli: differential ERP activity was detected within 150 ms for both types of stimuli; Fabre-Thorpe et al., 2001). Further, Kirchner and Thorpe (2006) argued, using a saccadic choice paradigm, that the processing required for ultra-rapid perceptual decision making may be even faster than previously believed. They showed that a pair of natural scenes flashed in the left and right hemifields could be compared for the presence of an animal with a median RT of 228 ms and a minimum RT of 120 ms. Utilizing a similar 2-AFC paradigm, Bannerman et al. (2009) showed that subjects could distinguish a fearful facial expression or body posture from a neutral one in under 350 ms (mean reaction times). Similarly, forced-choice saccades to identify human faces can be performed above chance and initiated with a mean reaction time of 154 ms and a minimum RT of just 100 ms (Crouzet et al., 2010).

In contrast, the number comparisons reported in the current paper are slower than the pure perceptual discriminations studied in this previous literature. One possible explanation for this difference is based on our choice of the minimum RT measure, which produces more conservative MRTs than the KT 2006 measure. Note, however, that even accounting for this difference, our task produced MRTs 20–40 ms slower than the ERP activity reported in several of the studies cited above. Alternatively, the difference might be due to important distinctions between our task and many of the previous paradigms. For example, in the KT 2006 task, subjects had to decide which of two natural scenes contained an animal. However, as only information from a single scene is necessary to make a decision, the two images provide redundant information, which might increase the efficiency

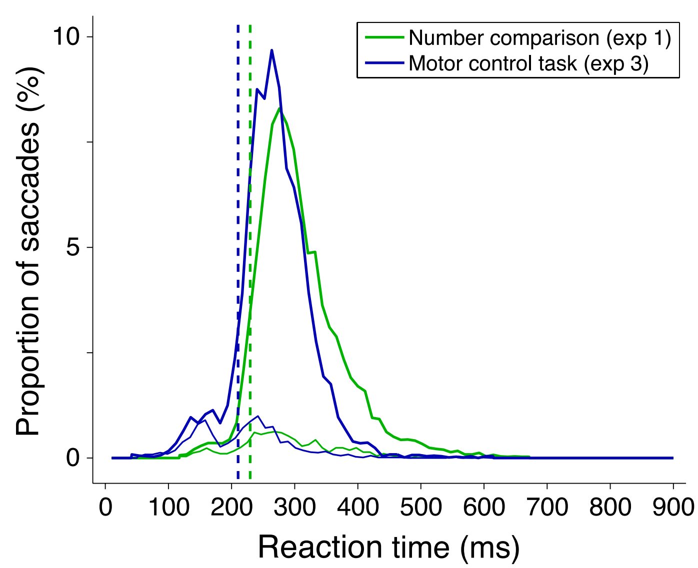

Figure 6. Reaction time distributions for Experiment 3 and Experiment 1 (speed condition). Correct trials = thick line. Error trials = thin line. Vertical lines depict mean MRTs.

Experiment 3: Number identification control

S	*N*	Accuracy (%)	Mean RT (ms) ± *SEM*	MRT
25	600	92.0	260 ± 1.8	204
26	600	96.0	292 ± 1.8	228
27	600	94.0	260 ± 1.9	185
28	576	90.5	288 ± 2.7	220
29	598	89.0	256 ± 2.8	188
30	599	93.5	271 ± 2.1	226
31	590	92.9	270 ± 2.2	215
32	591	94.8	251 ± 2.1	185
33	598	84.3	268 ± 2.7	217
34	593	85.8	253 ± 1.9	212
35	592	76.9	231 ± 3	212
36	590	79.7	249 ± 3.7	232
All	3564	89.1 ± 1.8	262 ± 5	210 ± 5

Table 5. Individual mean accuracy, mean RT, and minimum RT (MRT) in Experiment 3.

of choices in accordance with signal detection theory. Furthermore, in some of these non-numerical tasks it is not even necessary to process a whole image, since identifying an eye or a feather is sufficient to categorize an image as containing an animal. In contrast, in our Experiments 1 and 2, the two information sources were not redundant, implying that subjects had to process information from both stimuli in order to solve the task. This introduces an additional level of difficulty to the task that might account for some of the differences between our reaction times and those found by KT 2006 and others.

The methodology used in this paper to measure MRTs has also been applied by our group to study the speed at which subjects can make simple subjective value-based choices (i.e., choose which of two food items to eat; Milosavljevic, Koch, & Rangel, 2010). We found mean reaction times of 403 ± 21 ms and a mean MRT of 313 ± 17 ms, but with lower accuracies than those obtained here (73.3 ± 1.6%). Such rapid reaction times suggest that the computation and comparison of values in everyday decision making may recruit a set of cognitive processes similar to those involved in basic number comparisons (King & Janiszewski, 2011; Peters et al., 2008; Valenzuela & Raghubir, 2010). Needless to say, this possibility is highly speculative, and further investigation of the similarities and differences between these mechanisms is necessary to evaluate its validity.

Our experimental design has two important limitations that should be addressed in future studies. First, our stimuli always represented quantities using Arabic numerals, which are overlearned stimuli, in particular for our subject pool (members of a university community that places a premium on mathematical ability). It will be important to investigate if the psychometric properties of the basic number comparison process change when numerical information is represented in other ways (e.g., verbal vs. analog vs. auditory) which have been extensively studied in related domains (Barth, Kanwisher, & Spelke, 2003; Brannon, 2003; Cantlon, Platt, & Brannon, 2009; Dehaene, 1992; Dehaene & Cohen, 1995; Piazza, Pinel, Le Bihan, & Dehaene, 2007), and if the results hold for a more representative sample of the general population. Second, we considered only non-negative single-digit numbers. It will be important to investigate if our findings extend to multi-digit and negative numbers (Dehaene, Dupoux, & Mehler, 1990; Fischer, 2003; Fischer & Rottmann, 2005; Ganor-Stern, Tzelgov, & Ellenbogen, 2007; Hinrichs, Yurko, & Hu, 1981).

Acknowledgments

We would like to thank Alex Huth, Rufin VanRullen, two reviewers, and the editor for their helpful comments. This work was funded by the Tamagawa Global Center of Excellence (MM), the Mathers Foundation, ONR, DARPA, NSF, and the WCU Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology, R31-2008-000-10008-0 (CK).

Commercial relationships: none.
Corresponding author: Milica Milosavljevic.
Email: mmilosav@hss.caltech.edu.
Address: 1200 E. California Blvd., Pasadena, CA 91125, USA.

References

Bannerman, R. L., Milders, M., de Gelder, B., & Sahraie, A. (2009). Orienting to threat: Faster localization of fearful facial expressions and body postures revealed by saccadic eye movements. *Proceedings of the Royal Society B, 276,* 1635–1641.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition, 86,* 201–221.

Bogacz, R. (2007). Optimal decision-making theories: Linking neurobiology with behaviour. *Trends in Cognitive Sciences, 11,* 118–125.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brannon, E. M. (2003). Number knows no bounds. *Trends in Cognitive Sciences, 7,* 279–281.

Busemeyer, J. R., & Johnson, J. G. (2004). *Computational models of decision making*. Cambridge, MA: Blackwell.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences, 13,* 83–91.

Chandra, M. (2001). *Statistical quality control*. New York: CRC Press.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavioral Research Methods, Instruments, & Computers, 34,* 613–617.

Crouzet, S., Kirchner, H., & Thorpe, S. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision, 10*(4):16, 1–17, http://www.journalofvision.org/content/10/4/16, doi:10.1167/10.4.16. [PubMed] [Article]

Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception & Psychophysics, 45,* 557–566.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1–42.

Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors methods. *Journal of Cognitive Neuroscience, 8,* 47–68.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

Dehaene, S. (2007). A few steps toward a science of mental life. *Mind, Brain, and Education, 1,* 28–47.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and numerical magnitude. *Journal of Experimental Psychology: General, 122,* 371–396.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5,* 390–407.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1,* 83–120.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 626–641.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13,* 171–180.

Fischer, B., & Weber, H. (1993). Modes of saccade generation and their attentional control. *Behavioral & Brain Sciences, 16,* 595–610.

Fischer, M. H. (2003). Cognitive representation of negative numbers. *Psychological Science, 14,* 278–282.

Fischer, M. H., & Rottmann, J. (2005). Do negative numbers have a place on the mental number line? *Psychological Science, 47,* 22–32.

Ganor-Stern, D., Tzelgov, J., & Ellenbogen, R. (2007). Automaticity and two-digit numbers. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 483–496.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535–574.

Hinrichs, J. V., Yurko, D. S., & Hu, J. M. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 890–901.

King, D., & Janiszewski, C. (2011). The sources and consequences of the fluent processing of numbers. *Journal of Marketing Research, 48,* 327–341.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46,* 1762–1776.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in goal-directed choice. *Nature Neuroscience, 13,* 1292–1298.

Libertus, M., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behavioral and Brain Functions, 3,* 1–15.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40,* 77–105.

Milosavljevic, M., Koch, C., & Rangel, A. (2010). *Consumers can make above chance decisions in as little as 300 ms*. Working Paper, California Institute of Technology.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making, 5,* 437–449.

Moyer, R. S., & Landauer, T. K. (1967). The time required for judgments of numerical inequality. *Nature, 215,* 1519–1520.

Peters, E., Slovic, P., Vastfjall, D., & Mertz, C. K. (2008). Intuitive numbers guide decisions. *Judgment & Decision Making, 3*, 619–635.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron, 53,* 293–305.

Pinel, P., Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage, 14,* 1013–1026.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922.

Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics, 1,* 239–250.

Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. *PLoS Biology, 3,* e37.

Song, J.-H., & Nakayama, K. (2008). Numeric comparison in a visually-guided manual reaching task. *Cognition, 106,* 994–1003.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences, 95,* 7836–7841.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Usher, M., & McClelland, J. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592.

Valenzuela, A., & Raghubir, P. (2010). *Are consumers aware of top–bottom but not of left–right inferences? Implications for shelf space positions*. Working Paper, Baruch College, City University of New York.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review, 14,* 1011–1026.

VanRullen, R. (2007). The power of the feed-forward sweep. *Advances in Cognitive Psychology, 3,* 167–176.

VanRullen, R., & Thorpe, S. J. (2001a). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception, 30,* 655–668.

VanRullen, R., & Thorpe, S. J. (2001b). The time course of visual processing: From early perception to decision making. *Journal of Cognitive Neuroscience, 13,* 454–461.