Abstract

The current study investigated the neural activity patterns associated with numerical sensitivity in adults. Event-related potentials (ERPs) were recorded while adults observed sequentially presented display arrays (S1 and S2) of non-symbolic numerical stimuli (dots) and made same/different judgments of these stimuli by pressing a button only when numerosities were the same (target trials). The main goals were to contrast the effects of numerical distance (close, medium, and far) and change direction (increasing, decreasing) between S1 and S2, both in terms of behavior and brain activity, and to examine the influence of individual differences in numeracy on the effects of these manipulations. Neural effects of distance were found to be significant between 360–600 ms after the onset of S2 (greater negativity-wave activity for closer numerical distances), while direction effects were found between 320–440 ms (greater negativity for decreasing direction). ERP change-direction effects did not interact with numerical distance, suggesting that the two types of information are processed independently. Importantly, subjects’ behavioral Weber fractions ($w$) for the same/different discrimination task correlated with distance-related ERP-activity amplitudes. Moreover, $w$ also correlated with a separate objective measure of mathematical ability. Results thus draw a clear link between brain and behavior measures of number discrimination, while also providing support for the relationship between non-verbal magnitude discrimination and symbolic numerical processing.

Introduction

Ratio dependency is a hallmark of numerical discrimination and can be decomposed into distance effects and magnitude effects (e.g., Moyer & Landauer, 1967; Dehaene & Changeux, 1993). The distance effect refers to an increase in accuracy and speed of response in discriminating two numerical values as the distance between them increases. The magnitude effect refers to a decrease in accuracy and slowing in speed at discriminating pairs of values with constant numerical distances but increasing numerical magnitude. Collectively, the implication of magnitude and distance effects is that the ease with which two numerosities are distinguished is a function of their ratio.

Neural correlates of the numerical distance effect have been examined using both functional MRI (fMRI) and event-related potentials (ERPs). Brain imaging studies using fMRI have consistently identified the involvement of posterior parietal areas in numerical processing.
including right and left intraparietal sulci (IPS; Cantlon, Brannon, Carter, & Pelphrey, 2006; Pinel, Dehaene, Riviere, & LeBihan, 2001; Pinel, Piazza, Le Bihan, & Dehaene, 2004). These regions were found to exhibit greater signal change for small than for large differences in number (Pinel, Dehaene, Riviere, & LeBihan, 2001; Pinel, Piazza, Le Bihan, & Dehaene, 2004). Early-latency neural signatures of the numerical distance effect have been found with ERPs in the form of increased positive-polarity activity at around 200 ms post-stimulus being associated with closer relative to farther numerical distances (Dehaene, 1996; Libertus, Woldorff, & Brannon, 2007; Temple & Posner, 1998). Although early latency ERP effects can be easily influenced by physical differences in stimuli, Libertus, Woldorff, & Brannon (2007) found such effects even when controlling for such differences, supporting the view that at least some numerical comparisons may be carried out quickly and relatively automatically. Close numerical distances have also been associated with greater negative-polarity components at longer latencies (230–540 ms), an effect which has been compared to the semantically related N400 component (Niedeggen & Rösler, 1999; Niedeggen, Rösler, & Jost, 1999; Paulsen & Neville, 2008; Szücs & Csépe, 2004; Szücs, & Csépe, 2005).

Psychophysical research has shown that people vary widely in their numerical sensitivity. A recent study by Halberda, Mazzocco, & Feigenson (2008) found that adolescents differ in their ability to discriminate non-symbolic numerosities, and that their performance in the numerical comparison task correlated with standardized math achievement scores from kindergarten through sixth grade. An interesting question is whether these individual differences in numerical sensitivity in behavioral tasks would be reflected in the neural signatures of numerosity discrimination, and whether such effects would be of a qualitative or quantitative nature. In fact, previous studies have found specific ERP components that differ as a function of group level performance during number judgment and discrimination tasks (Soltész, Szücs, Dékány, Márkus, & Csépe, 2007; Paulsen & Neville, 2008). For example, Soltész, Szücs, Dékány, Márkus, & Csépe (2007), found that during digit comparisons, dyscalculic adolescents failed to show the late ERP potential (P300) observed in an age-matched control group. Similarly, Paulsen and Neville (2008) found that individuals who performed well in a numerical discrimination task elicited larger N400-like negativities to numerically close stimuli than to numerically distant stimuli during discrimination, while no such distance effects were found in the ERPs of those who performed poorly.

**Numerosity Change Direction**

Distance and magnitude effects are well-known, but there is burgeoning evidence for another factor that affects discriminability: change direction, by which a relative increase in quantity is identified more rapidly and with greater accuracy than a relative decrease in quantity (Conson et al., 2008; Kaan, 2005; Paulsen & Neville, 2008). Effects of change direction can occur even for the same pairs of absolute numerosity values (Conson et al., 2008; Paulsen & Neville, 2008).

Change direction effects have been found in time judgments, symbolic number judgments, and non-symbolic numerosity judgments. For example, Conson and colleagues (2008) required participants to judge whether the first or the second of a pair of sequentially presented tones was longer or shorter. Tone durations varied between 1000–3000 ms with differences of 200–1000 ms by 200 ms increments. As expected, the authors found a significant temporal distance effect: accuracy increased with the difference between tone lengths. However, the authors also found that participants were more accurate when the second tone was longer compared to shorter in duration than the the first tone. Similarly, Kaan (2005) had participants compare sequentially presented number-word stimuli and found that participants responded more quickly and accurately when the second numerical stimulus was larger relative to when the second stimulus was smaller than the first stimulus.
Finally, Paulsen and Neville (2008) found a behavioral interaction between magnitude and numerical distance of dot stimuli that led to the examination of change direction. In their study, participants were required to judge whether the second stimulus (S2) of a pair of sequentially presented dot-array stimuli contained the same or a different number of dots than the first stimulus (S1) of the pair. Typical behavioral distance effects were observed, with small numerical distances eliciting a larger negative component between 350–450 ms than stimuli with greater numerical distances. Unexpectedly, a direction effect was also observed, in that accuracy was greater and RT was faster when the second numerosity was larger in magnitude than the first (e.g., 16:24), compared to when it was smaller (e.g., 24:16). S1:S2 pairs that decreased in magnitude also elicited a greater negativity around 400 ms compared to pairs that increased in magnitude. However, since the experiment was not designed to evaluate change direction, change direction effects were confounded with magnitude.

Present Experiment

The present experiment was conducted with three objectives. First, we wanted to test the prediction that performance in non-symbolic number discrimination would correlate with distance-related ERP processing effects. (Paulsen & Neville, 2008). Our second objective was to determine whether change direction effects would be observed when the magnitude of the S2 numerosity was carefully controlled. We predicted that S1 < S2 pairs would be more accurately detected as different than S1 > S2 pairs and that the increasing pairs would also elicit lesser ERP negativities around 400 ms. A third objective was to explore the relationship between behavioral and electrophysiological indices of numerosity discrimination and explicit assays of symbolic numeracy capabilities. The correlations that have been found between numerosity approximation and standardized math scores (Halberda, Mazzocco, & Feigenson, 2008) supports the idea that formal mathematical reasoning is based upon the primal ability to process number non-symbolically (Spelke & Dehaene, 1999). We tested this hypothesis here by administering two different measures of adult numeracy: one measure reflected subjective sense of numerical ability, and one measure reflected adults’ ability to reason about probabilities. Based on the recent findings by Halberda and colleagues (2008), we expected to find relationships between our measures of probabilistic reasoning and both the behavioral and neural measures of numerosity discrimination.

Methods

Participants

Twenty-one right-handed participants (12 female) between the ages of 18 and 35 (mean = 21.6 yrs) reporting normal or corrected-to-normal vision and with no known neurological impairments were recruited from Duke University and the surrounding community. All participants signed IRB-approved consent forms and were either paid for their time or received course credit. One subject was excluded for having too few artifact-free ERP trials, leaving a total of 20 participants in the final sample.

Procedure

Participants sat comfortably in a dimly lit, electrically shielded, acoustic chamber approximately 110 cm away from a CRT monitor, on which stimuli were presented within 5° of visual angle. Instructions were given to attend to the number of dots in the sequential pairs of stimulus arrays (S1 followed by S2), ignoring dot size, and to manually respond to any S2 that contained the same number of dots as the S1 that it followed. Participants were not informed

---

1Piazza, Izard, Pinel, Le Bihan, & Dehaene (2004) found a similar discrepancy between smaller and larger set sizes in a same-different task with respect the habituation displays: smaller S2 set sizes were identified as same more often than larger S2 set sizes. However, this bias was only found to occur with triangle stimuli and not dot stimuli.
that these S2 targets would occur on only 10 percent of the total trials. This design, with only a small fraction of the stimulus trials being targets and requiring a button press, allowed us to focus our ERP analyses on the discrimination processes on the nontarget stimulus trials, thereby minimizing the overlap of P300 response-associated effects on the ERP component of interest here. Responses were made by pressing the down-arrow key on a standard keyboard with either the left or right index finger, counterbalanced across participants. Trials were presented in four blocks of 150 trials with breaks between blocks. A practice block using similarly constructed stimuli was given to familiarize participants with the task. The practice trials contained only paired numerosities of 1:1 and 2:1 ratios and continued until the participant executed seven correct responses within ten trials.

Trials began with a 1000 ms display of a white fixation cross that was presented on a dark grey background and remained on the screen throughout the entire trial. After the initial fixation display, S1 was displayed for 500 ms, followed by a 500 ms inter-stimulus interval and a 500 ms display of S2. The trial ended with a 700 ms display of the fixation followed by an inter-trial interval that varied between 1300–1500 milliseconds. Trials were pseudo-randomized so that no more than three identical S1:S2 number ratios or S2 magnitudes appeared in a row. All participants received the same trial order. To reduce movement-related artifacts, participants were encouraged to keep their eyes focused on the fixation and refrain from blinking except during the initial fixation period.

**Dot Stimuli**

Sixty S1 and thirty S2 dot-array stimuli (white dots on a dark grey background) were generated using a Matlab script designed to vary item size among S1 and to control for total luminance among S2, following Piazza et al (2004). All S2 were equated for surface area, while the average surface area of S1 for each set size was equal to the surface area used for all S2. The dot radius for S1 varied within each set size by a maximum of ±15% of the radius determined by the mean surface area for that set size. Five exemplars of each of the numerosities 10, 12, 14, 15, 18, 21, 29, 33, 40, 43, 50, and 60 were used as S1, and nine exemplars of numerosities 20 and 30 were used as S2; additionally, one each of the S1 numerosities was generated as an S2 numerosity so that a 1:1 S1/S2 ratio could be presented on 10% of the total trials (Same trials). Stimulus images were randomly mirrored and rotated by a multiple of 90° upon presentation to prevent participants from categorizing and identifying stimuli based on their unique patterns of dot arrangement. S1 and S2 were paired to form approximate ratios of 1, 0.7 (Close), 0.6 (Medium), and 0.5 (Far). Half of these ratios were increasing (S1 < S2) and half were decreasing (S1 > S2). Each S1 and S2 pair was presented only once to create a total of 600 trials: 60 Same trials, and 180 each of Close, Medium, and Far trials.

**ERP acquisition**

The electroencephalogram (EEG) was recorded from 64 tin electrodes embedded in custom elastic caps (Electro-Cap International, Eaton, OH), using a bandpass filter of 0.01 to 100 Hz at a sampling rate of 500 Hz (SynAmps, Neuroscan). Additional electrodes were placed at the outer canthi and just beneath the right and left eyes to record electrooculograms and to monitor eye movement. Scalp electrodes were referenced to the right mastoid electrode during recording and re-referenced to the average of the right and left mastoids off-line. Impedances were kept below 5KΩ for scalp and eye electrodes, and below 2KΩ for mastoid electrodes.

**Numeracy Assessments**

Two numeracy assessments were used. The first was a *subjective measure* of numeracy (Fagerlin et al., 2007) that probed participants’ preferences for quantitative versus qualitative information, as well as their own sense of ability in working with numbers. The subjective measure consisted of 8 questions – e.g., “How good are you at calculating a 15% tip?” –
requiring a rating from 1–6. The second assessment was a more objective measure that contained 10 word problems requiring subjects to make calculations and convert numerical probabilities (e.g., “if person A’s risk of getting a disease is 1% in ten years, and person B’s risk is double that of A’s, what is B’s risk?”) (Lipkus, Samsa, & Rimer, 2001). These assessments were administered immediately following the ERP session.

Data analysis

**Behavioral data**—Accuracy was calculated for each participant at all S1:S2 ratios and for all S2 magnitudes. To examine effects of numerical distance and change direction, correct rejections (CR) on change trials were submitted to 2×2×3 repeated measures ANOVA (RM-ANOVA) using S2 Set Size (2 levels: 20 & 30), Distance (3 levels: Close, Medium, Far), and Change Direction (2 levels: Increase, Decrease) as factors. Significant interactions were followed up with simple pair-wise specific comparisons.

Two different models were used to estimate each subject’s numerical sensitivity (Appendix A). The first model was drawn from Piazza et al. (2004) and is based on the idea that a compressed number line resembles a Gaussian distribution laid out on a logarithmic scale (Pica, Lemer, Izard, & Dehaene, 2004). This model has two free parameters: \( w \), which determines the width of the function and represents the precision of the numerical estimates, and \( \delta \), which determines the depth of the function and represents a bias to respond “same”. The second model is drawn from Pearson, Roitman, Brannon, Platt, and Raghavachari (2010), and uses a Gompertz function rather than a Gaussian function. This second model also contains two free parameters, \( w \) and \( \kappa \), which correspond to the Gaussian model parameters \( w \) and \( \delta \), respectively. The Gompertz function is similar in form to the Gaussian function, but because it is a double exponential, the Gompertz function is asymmetric on a logarithmic scale, thus providing differential predictions depending on the order of presentation (i.e., large or small first). Two measures that can be used to assess the goodness of fit of a statistical model are Akaike and Bayesian information criteria (AIC and BIC respectively). We used AIC and BIC to compare the ability of the Gaussian and Gompertz models to estimate individual numerical sensitivity, as reflected in \( w \), and group level variability in \( w \) (Anderson, Burnham, & Thompson, 2000). Non-linear mixed effects regressions were also used to estimate the variance accounted for by each of the two models (Pineiro, Bates, DebRoy, Sarkar, & the_R_Core_team, 2008).

To examine the relationship between numerical sensitivity (as estimated by \( w \) in each model) and ERP effects we conducted two different types of analyses. First, we grouped subjects based on a median split in \( w \) to create high- and low-proficiency groups for comparison. Second, we performed analyses that treated \( w \) as a continuous variable across all participants.

**ERP data**—ERP averages were calculated for each participant for close, medium, and far distances, for both increasing-direction and decreasing-directions, using a 200 ms pre-stimulus baseline. Each participant’s data for every S2 presentation in the first 13 minutes of recording time were examined visually to individually set parameters to exclude ocular and motion-related artifacts in a standard artifact rejection routine. One participant was rejected from the ERP analysis for having fewer than 10 artifact-free trials in one of the conditions. Only trials in which the S2 was correctly rejected as different in number were included in the analysis, for two reasons: first, because false alarms responses would include button presses and other potential target-related artifacts of the ERPs, and second because a false alarm would indicate that a different S2 numerosity was not actually perceived as different. The resulting average number of trials for the decreasing direction were: Close, 27; Medium, 41; and Far, 57. For the increasing directions trial counts were: Close, 51; Medium, 67; Far, 70. A 4 × 8 grid of 32 channels over the entire scalp was used for analysis (Fig. 1): left and right Hemisphere, Lateral...
and Medial sites, and eight rows from anterior to posterior. Data shown in the figures and used
for analyses were low-pass filtered at ~50 Hz using a 9-point running average.

Although the primary effects of interest were not expected to occur until between 300–700 ms,
local peak latency and amplitude measures were also taken for earlier-latency components.
Because the morphology of waveforms differ over anterior and posterior locations (Fig. 4),
separate analyses were run using the measurements taken from the 16 anterior and 16 posterior
channels. With the exception of the anterior and posterior measures of the somewhat longer-
latency N3 response (latency around 350–400 ms), mean amplitude measurements for early
components were taken from time windows spanning 20 ms on either side of the mean local
peak latency for each distance-by-change-direction condition. For each of the anterior 16 and
posterior 16 channel sets, repeated-measures analyses of variance (RM-ANOVAs) were
conducted using Distance (3 levels), Change Direction (2 levels), and three electrode-location
factors: Hemisphere (2 levels), Laterality (2 levels), and Caudality (4 levels).

The time window used for analysis of late (> 300 ms) distance and change-direction effects
was identified by the following procedure. Onset and offset analyses were performed for
distance and change direction effects separately. Mean amplitude measurements were taken
from the same 32 channels described above between 280–700 ms for successive 20 ms time
windows that overlapped by 10 ms. Effect onsets were identified as the first of five successive
time windows that all returned a significant main effect of either distance or change direction
at \( p < 0.05 \) of a RM-ANOVA using amplitude measurements taken from all 32 channels. Offsets
were similarly identified as the last of five successive time windows that all had a significant
main effect. Finally, the time window used to take mean amplitude measurements for the
analysis of distance and change direction effects simultaneously was determined by the overlap
of onset-offset windows of each. These means were submitted to a RM-ANOVA with Distance,
Change Direction, and Electrode Location as within-subjects factors, and Proficiency as a
between-subjects factor. Significant main effects of Distance were followed up with pairwise
comparisons using RM-ANOVA and the electrode-location factors. All reports of significance
for ERP data used Greenhouse-Geisser epsilon corrections for violation of sphericity. Main
effects and interactions are reported only for the factors of Distance, Set Size, or Change
Direction.

To examine the relationship between behavioral performance and ERP effects of numerical
distance, behavioral data (individual estimates of the numeracy sensitivity, \( w \), derived as
described above) were correlated with the mean of a difference wave formed by subtracting
the ERPs for the Far condition from the ERPs for the Close condition, collapsing across
Increasing and Decreasing conditions. The mean amplitude used for this statistic was taken
from seven fronto-central electrode sites identified by the topography of the Distance effect,
within the same time window used for the analysis of the late Change Direction and Distance
effects.

Results

Behavioral data

Both the Gaussian and Gompertz models that were tested accounted for a large percentage
of the variance in identifying the S2 numerosity as different from S1 (Fig. 2): \( R^2_{\text{Gaussian}} = 0.74 \),
\( R^2_{\text{Gompertz}} = 0.76 \), suggesting similar fits that were both very good. Akaike and Bayesian
information criteria (AIC, BIC) give estimates of a model’s goodness of fit, and can be used
as tools for model selection, whereby lower values indicate better fit. Information statistics for
the Gompertz model (AIC = −208.51, BIC = −183.47) were lower than those for the Gaussian
model (AIC = −180.26, BIC = −155.21), demonstrating that the Gompertz model yielded a
better fit. Accordingly, individual parameter estimates from this model were used in subsequent
analyses. Lower Proficiency (LP; n = 10) and Higher Proficiency (HP; n = 10) participants were sorted by a median split of the calculated \( w \)'s, which ranged from 0.27 to 0.39 (Mdn = 0.33). Lower and Higher Proficiency groups had mean \( w \)'s of 0.36 and 0.31, respectively. These values are a little higher than the 0.19–0.20 range of Weber fractions obtained in some studies of numerical discrimination (Barth, La Mont, Lipton, Dehaene, Kanwisher, & Spelke, 2005; Cantlon & Brannon, 2007; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008), but within the range obtained in others (Halberda, Mazzocco, & Feigenson, 2008); differences may be due to both the method of estimation\(^2\) and the task.

Participants exhibited classic behavioral effects of numerical Distance whereby accuracy (correct rejections) increased monotonically with distance: Close (54%), Medium (73%), Far (87%). Participants were also accurate at correctly identifying the Same Targets (87%). A RM-ANOVA on the correct-rejection rates returned a significant main effect of Distance, \( F(2,38) = 222.83, p < 10^{-20} \). A main effect of proficiency as a between-subjects factor was also significant, \( F(1,18) = 27.46, p < 10^{-5} \), which interacted only with effects of Distance, \( F(2,36) = 7.36, p < 0.01 \) (all other \( p > 0.17 \)), indicating that effects of Set Size and Change Direction did not differentially affect HP and LP groups. Eight of the nine pairwise t-tests performed to follow up on the Proficiency by Distance interaction were significant test at a Bonferroni adjusted alpha of 0.005. The one exception was the comparison of proficiency groups at Far distances, \( p > 0.008 \), suggesting that as the numerical distance between S1 and S2 increased, group differences decreased.

As predicted, participants were less accurate on Decreasing trials (58%) than they were on Increasing trials (85%), \( F(1,19) = 43.38, p < 10^{-5} \) (Fig. 3). A significant interaction between Change Direction and Distance was also found, \( F(2,38) = 11.33, p = 0.001 \), indicating that the difference between correct rejection rates for Increasing and Decreasing directions was greater at Close (33%) and Medium (32%) distances than it was at Far distances (16%). Paired sample t-tests for effects of Distance at each level of Change Direction, and for effects of Change Direction at each level of Distance, were all significant at a Bonferroni adjusted alpha of 0.005 (.05/9). Collectively these results show the robustness of Distance and Direction effects, and also indicate that as numerical Distance increases, the effects of Change Direction grow weaker. That is, at larger ratios, a ceiling in accuracy limits the observation of Change Direction effects.

A main effect of Set Size was also found, \( F(1,19) = 105.32, p < 10^{-9} \); S2 stimuli of Set Size 20 had a higher correct rejection rate (78%) than Set Size 30 (65%), indicating that smaller S2 numerosities were easier to discriminate (Fig. 3B). A Set Size x Change Direction interaction, \( F(1,19) = 134.1, p < 10^{-10} \), was followed up with pairwise t-tests which revealed that for Set Size 30, Increasing directions were responded to with greater accuracy (86%) than Decreasing directions (44%), \( t(19) = 9.85, p < 10^{-9} \), but that for Set Size 20, the difference between Increasing (84%) and Decreasing directions (72%) was only marginally significant (with a Bonferroni adjusted alpha of 0.0125 (0.05/4), \( p = 0.013 \)). There was no significant difference between Set Size 20 and Set Size 30 for Increasing directions \( (p > 0.15) \), but there was a marked difference between Set Size 20 and Set Size of 30 for Decreasing directions, \( t(19) = 13.46, p < 10^{-11} \), suggesting that the main effect of Set Size found here was driven by the very low accuracy for Decreasing direction Set Size 30.

Finally, to explore the relationship between number discrimination and numeracy, correlation analyses between individual \( w \)'s and scores on the numeracy assessments were performed. One participant was excluded from this analysis because he/she responded identically to 7 of the 8 subjective assessment questions and to 5 of the 10 objective measure questions, suggesting

\(^2\)Slightly lower estimates of \( w \) were obtained from the Gaussian model: \( mean = 0.30, range 0.15–0.39 \).
that the questions were not answered truthfully. Mean scores for the subjective and objective numeracy assessments were 36.4 (SD = 4.85) out of 40 and 8.95 (SD = 1.73; range 3–10) out of 10, respectively, suggesting that as a whole, participants had a preference for, and felt relatively comfortable with, their ability to use numerical information, and that participants were fairly capable of working through problems concerning probabilities. Although w was not found to correlate significantly with the subjective measure of numeracy ($p > 0.29$), it did correlate significantly with the objective measure of numeracy, $r = -0.47, p < 0.05$. Thus, individuals who performed better at the magnitude discrimination task were also better at translating and reasoning with probabilities presented symbolically, consistent with recent findings by Halberda, Mazzocco, & Feigenson (2008).

**ERP data**—As shown in Figure 4, ERPs to the S2 stimuli elicited a series of negative and positive deflections that varied as a function of electrode site. Over anterior sites, negative peaks occurred at around 100 ms, 200 ms, and 410 ms, and positive peaks at 180 ms and 290 ms. Over posterior sites, negative peaks occurred at around 150 ms and 360 ms, with a positive peak appearing at around 270 ms. Analysis of the local peak latencies of components occurring earlier than 300 ms and their corresponding mean amplitudes across the 40 ms surrounding those peaks did not return any significant effects of Distance or Change Direction, all $p > 0.08$. A third negative deflection at around 400 ms (N3) was observed over all sites that appeared to be superimposed on a larger positive going wave. Analysis of N3 peak latencies over anterior sites returned a significant effect of numerical Distance, $F(2,38) = 4.12, p < 0.05$, and a significant linear trend for Distance, $F(1,19) = 6.92, p < 0.05$, showing that as numerical Distance increased, N3 peak latency decreased: Close (423 ms), Medium (410 ms), and Far (399 ms). Over posterior sites, a similar effect of Distance was observed on the N3 in this latency range, $F(2,38) = 13.41, p < 10^{-3}$: Close (368 ms), Medium (365 ms), and Far (347 ms). A significant effect of Change Direction on N3 latency, $F(1,19) = 4.83, p < 0.05$, showed that Increasing direction trials had earlier N3 latencies (355 ms) than did Decreasing direction trials (365 ms). There was no interaction between Distance and Change Direction ($p > 0.76$) on these latencies, indicating that the two effects are relatively independent with respect to neural response.

As seen in Figure 4, Close distances elicited the greatest negativities in the 300–700 ms range, followed by Medium, and then Far. Onset and offset analyses were performed for effects of numerical Distance and of Change Direction on the means of ERP amplitudes for successive 20 ms time-windows that overlapped by 10 ms, starting at 280 ms and ending at 700 ms. These analyses revealed that numerical Distance effects reached significance from 360–600 ms, while effects of Change Direction began and ended earlier, 320–440 ms. To rule out a significant contribution of Set Size on the ERP components of interest, an onset/offset analysis of Set Size was performed as was done for effects of Distance and Change Direction. Effects of Set Size did not reach significance for more than two of the 20-ms time-windows between 280–700 ms, suggesting that any ERP effects of Set Size in the current data set are negligible. We thus used a 360–440 ms time-window for all the 32 channels analyzed for effects of distance and change direction.

A main effect of Distance in the data from the 32 channels was found between 360–440 ms, $F(2,38) = 11.76, p < 0.001$. Pair-wise tests showed that Close (3.39 $\mu$V) was marginally more negative than Medium (3.96 $\mu$V), $F(1,19) = 3.75, p < 0.07$, Medium was more negative than Far, (4.80 $\mu$V), $F(1,19) = 9.79, p < 0.01$, and that Close was more negative than Far, $F(1,19) = 20.58, p < 0.001$. Change Direction was also significant, $F(1,19) = 15.91, p < 0.001$, with

---

3This larger positive going component was also found by Paulsen & Neville (2008). In that study, it was analyzed by using a 450–600 ms time window for mean ERP amplitudes and was suggested to be similar to a P300 effect unrelated to numerical discrimination. We focus here on the earlier negative going component.
decreasing direction (3.51 \mu V) eliciting a greater negativity than Increasing direction (4.60 \mu V). There was no interaction between Distance and Change Direction (p > 0.22), again supporting their relative independence.

**Numerical Proficiency and ERP Numerosity Effects**—Including Proficiency as a between-subjects factor in the RM-ANOVA on the ERP N3 amplitudes returned a significant interaction between Distance and Proficiency, F(2,36) = 5.31, p < 0.05, but not an interaction between Direction and Proficiency, p > 0.76. There were no effects of proficiency on N3 latencies (all p > 0.20).

To examine the interaction between distance and proficiency on the N3 ERP amplitudes, we analyzed the high-performing (HP) and low-performing (LP) groups separately. Both groups showed ERP distance effects between 360 and 440 ms, F(1,18) = 12.47, p < 0.001, and F(1,18) = 4.50, p < 0.05, respectively. Pairwise comparisons of distance for the ERP activity during this latency range found that for the HP group, Close (2.7 \mu V) was more negative than Medium (4.03 \mu V), F(1,9) = 15.88, p < 0.01, but that the Medium vs. Far (4.84 \mu V) comparison only approached significance, F(1,9) = 3.19, p < 0.11. The LP group showed a somewhat different pattern: Close (4.08 \mu V) was not significantly different from Medium (3.90 \mu V), p > 0.63, and Medium was significantly more negative than Far (4.75 \mu V), F(1,9) = 7.79, p < 0.05. As illustrated in Figures 5 & 6, the amplitudes for medium and far trials were similar in the HP and LP groups suggesting that the close condition was driving the interaction between distance and proficiency in the ERPs. Pairwise comparisons including Proficiency as a between-subjects factor were run to test this hypothesis and revealed that Proficiency modulated the distance effects between Close and Medium, F(1,9) = 9.26, p < 0.01, and between Close and Far, F(1,18) = 7.43, p < 0.05, but not between Medium and Far, p > 0.92. Consequently, further analyses of the relationship between the relationship between distance related ERPs and proficiency were done using the Close – Far difference waves.

To assess the relationship between nonverbal numerical acuity and distance related ERP effects, we examined the relationship between \( w \) and the Close-minus-Far difference waves (360–440) at the seven fronto-central electrode sites where the effects of Distance appeared most prominent (Figs. 4 & 7). A significant correlation between \( w \) and the Close-Far difference wave was found (\( r = 0.54, p < 0.05 \)). Tests for correlations between subjective and objective numeracy measures and distance-related ERPs did not reach significance, although a trend was found for the relationship between the objective numeracy measure and ERP amplitude, (\( r = -0.371, p = 0.12 \)): individuals with greater scores of probabilistic reasoning tended to have greater distance-related ERP effects.

**Discussion**

Our main finding was that nonverbal numerical acuity as quantified by \( w \) was correlated with a long-latency (360–600 ms) negative-polarity ERP wave (N3). Specifically participants with better discrimination in the numerosity task (smaller \( w \)) tended to display greater amplitude differences between ERPs to close and far numerical distances. This longer-latency distance-related negativity may be related to the N400 family of effects that occur for semantically meaningful stimuli (Kutas & Federmeier, 2000). Such negativities are typically interpreted as indicating the relative effort involved with integrating a target into the context of the preceding prime (McPherson & Holcomb, 1999; Misra & Holcomb, 2003; Paulsen & Neville, 2008).

It is important to note that the effort expended in making numerical comparisons or attention allocated to close versus far stimuli may have contributed to the distance-related ERPs obtained. However, the same distance-related ERP numerosity effects we report here were found in an earlier study in which subjects engaged in a non-numerical color discrimination...
task (Paulsen & Neville, 2008). Given that the color task used by Paulsen and Neville (2008) did not require any explicit numerical processing and the same distance-related ERPs during implicit numerical processing were found, this suggests that attention is unlikely to be the sole driving force in this effect. Nevertheless, it remains possible that the allocation of attentional resources to numerical comparisons acts to amplify the ERP effects of numerical distance and change direction. More generally, differences in attentiveness could contribute to between-individual variability in proficiency and/or within-individual ERP differences between close and far stimuli. Additionally, one could argue that our numerical distance ERP effects might be better characterized as discrimination ERP effects, and could be elicited using other non-numerical stimuli. However, Paulsen and Neville (2008) reported markedly different ERP effects for color distance among a similar range of accuracies in discrimination. Thus, the possible generalization of our numerical distance ERP effects as discrimination effects may be restricted to certain types of stimuli. Future studies are needed to further address the possibility that attention, discrimination ability, and item difficulty interact.

The latencies of ERP effects over anterior and posterior sites were found to be similar, indicating several possibilities. One possible simple explanation is that there is a single deep generator, or perhaps bilateral pair of deep generators, leading to a widespread distribution at the scalp. Greater activation in anterior cingulate for numerical comparisons of small distance relative to large distance is consistent with such an explanation (Ansari & Dhital, 2006). Another possibility is that the source of activity is widespread, with multiple generators acting in parallel. Based on the neuroimaging literature, during number comparison activation in prefrontal regions may occur in conjunction with activation in posterior regions such as IPS (Ansari & Dhital, 2006; Ansari, Lyons, van Eimeren, & Xu, 2007; Chochon, Cohen, van de Moortele, & Dehaene, 1999). On the other hand, parietally distributed ERPs associated with numerical distance tend to occur at earlier latencies than the effects we report here (Libertus, Woldorff, & Brannon, 2007; Pinel, Dehaene, Riviere, & LeBihan, 2001). Given the distribution and timing of the current findings, we can only speculate that our distance ERP effect is likely to be arising mainly from frontal brain regions.

A second finding from this study was that non-symbolic numerical acuity was positively correlated with explicit probabilistic reasoning. Recent studies demonstrate a positive correlation between symbolic reasoning about number and the ability to represent non-symbolic numerical magnitudes in early childhood (Halberda, Mazzocco, & Feigenson, 2008; Booth & Siegler, 2008). It seems likely that the correlation we observe between numerical acuity and probabilistic reasoning reflects the same cognitive scaffolding whereby nonverbal magnitude representations support acquisition of a broad range of symbolic math. Although ERP amplitudes and probabilistic reasoning scores were not significantly correlated, there was a trend in this direction ($p = 0.12$). The lack of significance could be due to insufficiently high signal-to-noise ratio in the ERPs, or the fact that the numeracy questionnaire showed relatively small amount of variability with scores approaching ceiling, or both. A numeracy measure providing greater variability may be necessary to more definitively elucidate this relationship.

Our third main finding was that after controlling for the effects of the numerosity of the S2, effects of change direction were confirmed and were observed both behaviorally and in the ERPs. Participants were better at correctly rejecting S1-S2 pairs as different when S2 had a larger number of dots than did S1 compared to when S2 had a smaller number of dots than S1. There was also an interaction between change direction and numerical distance for accuracy, such that as numerical distance increased, accuracy differences between increasing and decreasing directions decreased. Thus, change direction mattered relatively little for large S1:S2 ratios, possibly because performance approached ceiling levels. Robust effects of change direction were also found in the ERP data. Decreasing S1:S2 pairs elicited a greater negativity
than increasing S1:S2 pairs. This ERP change direction effect was not found to interact with the ERP distance effect, suggesting that the effects are independent with respect to brain activity.

The behavioral effects of change direction are particularly noteworthy because it is commonly assumed that it is only the ratio of the comparators that determine the discriminability of their difference (Barth, Kanwisher, & Spelke, 2003; Barth et al., 2005; Brannon, Canton, & Terrace, 2006; Cantlon & Brannon, 2006, 2007; Halberda et al., 2008; Jordan & Brannon, 2006a, 2006b; Moyer & Landauer, 1967; Xu & Spelke, 2000). Orthogonal to the ratio effects obtained, we found robust effects of change direction in both behavioral and ERP data, showing that decreasing numerosities are less discriminable and elicit greater ERP negativities around 400 ms than do increasing numerosities. This novel finding further underscores Weber’s law in its application to numerical discrimination, affirming the importance of the initial intensity or magnitude for noticing a change. One possible explanation for why change direction effects have rarely been found is that a majority of studies on numerical discrimination present comparator stimuli simultaneously.

Our data was well-fit by a Gompertz based model of numerical comparison (Pearson et al., 2010). A key aspect of this model as opposed to other models of numerical comparison (Piazza et al., 2004; Pica et al., 2004) is that it captures the change direction effect. The Gompertz model works by having S1 serve as a numerical anchor, or reference point, against which S2 is compared. Thus, the model differentiates between S1:S2 pairs that decrease or increase despite their common ratio (e.g., 1:2 or 2:1) while maintaining constant predictions for identical ratios of different set sizes (e.g., 2:1 and 4:2). Although some of the asymmetry of the change direction effect is captured by the Gompertz model used here, it appears to be not quite flexible enough to capture all of the asymmetry. Both the Gompertz and Gaussian model functions can accept parameters that adjust for the degree of asymmetry in the discrimination curve (Fig. 2), though it has not yet been worked out where this parameter might be best placed (e.g., to modify $\omega$ or $n$ on one half of the curve) to yield a maximum fit with a clear interpretation, or whether a different model altogether would be more parsimonious.

Conclusions

Electrophysiological measures of brain activity elicited in response to a nonverbal numerical same-different decision showed robust variation as a function of the numerical distance between the two stimuli, the ordinal direction of the two stimuli, and their overall magnitude. Moreover, and particularly noteworthy, was that there was a positive relationship between numerical acuity abilities and distance-related ERP effects. In particular, for the most difficult numerical comparisons, subjects with higher nonverbal numerical acuity as measured by the parameter $\omega$ showed higher amplitude N3 ERPs during the latency range of 360–600 ms. In addition, $\omega$ also correlated with the amplitude of the N3 Close-minus-Far difference wave, whereby individuals with higher numerical acuity showed larger distance-related N3 ERP amplitudes. These findings indicate that individual differences in numerical acuity can be identified by robust and covarying measures of brain activity and behavior. Our results also provide additional evidence for the relatively little-studied phenomenon of change direction, whereby sequential numerical pairs of increasing magnitude are more easily discriminated than pairs of decreasing magnitude.

Acknowledgments

We would like to thank Sridhar Raghavachari and John Pearson for their assistance with modeling the behavioral data, and to members of the Infant Cognition Center at Duke University for their help in preparing this manuscript. This research was supported by a McDonnell Scholar Award to EMB and NINDS R01-051048 to MGW.
References


Barth H, La Mont K, Lipton J, Dehaene S, Kanwisher N, Spelke E. Non-symbolic arithmetic in adults and young children. Cognition. 2005


Appendix A

**Gaussian function**

\[ P_{\text{different}}(n_2, n_1) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{\delta + \log \left( \frac{n_2}{n_1} \right)}{w \sqrt{2}} \right) + \text{erf} \left( \frac{\delta - \log \left( \frac{n_2}{n_1} \right)}{w \sqrt{2}} \right) \right] \tag{1} \]

Equation 1 is a slightly modified version of the equation given in the supplementary data pages of Piazza et al. (2004) and defines the probability of responding “different” \((P_{\text{different}})\) in a Same/Different task to a numerosity stimulus with \(n_2\) items given a preceding stimulus with \(n_1\) items. The form of the internal representation of numerosities \(n_1\) and \(n_2\), i.e. the tuning curves for \(n_1\) and \(n_2\), are taken to be Gaussian on a logarithmic scale; the error function \((\text{erf})\) is used here to integrate both halves of the Gaussian curve. The \(w\) term in this equation represents the width of the tuning curves and therefore the precision of the internal representation, while the \(\delta\) term represents the depth of the curve and therefore the tendency to respond with a bias towards either “different” or “same”. As the ratio \(n_1/n_2\) approaches 1, \(\log(n_1/n_2)\) approaches 0, and \(P_{\text{different}}\) approaches 0; unlikely to respond “different”. As \(w\) approaches 0, \(P_{\text{different}}\) approaches 1 when \(n_1\) does not equal \(n_2\).

**Gompertz function**

\[ P_{\text{different}}(n_2, n_1) = 1 - k \left( e^{-\log(2)e^{\frac{n_2-n_1}{n_1}}} \right) \left( 1 - e^{-\log(2)e^{\frac{n_2-n_1}{n_1}}} \right) \tag{2} \]

Equation 2 is a double exponential equation based on the biophysical model of numerical classification presented by Pearson et al. (2010). As in Equation 1, \(w\) determines curve width and representational precision, while \(k\) adjusts the depth of the curve and therefore a bias in responding “same” or “different,” with a zero probability of responding “different” when \(n_1 = n_2\) and \(k = 4\). Although the shape of the functions defined by Equations 1 and 2 are similar, Equation 2 differs in that it is fundamentally asymmetric: there is a greater probability of responding “different” when \(n_2 > n_1\) than when \(n_2 < n_1\).
Figure 1. Electrode locations and electrode factors used for ERP analysis
Figure 2. Overlays of accuracy predictions and fits for Gaussian and Gompertz models as a function of S1/S2 ratio

Data points include standard errors and are presented on both linear and logarithmic scales. Note that the Gompertz model provides a closer fit to the observed values on the right side of the distribution, reflecting the asymmetry in performance between increasing and decreasing directions.
Figure 3. Accuracy in discrimination task as a function of S1/S2 ratio
Figure 3A. Distance effects are seen in the accuracy measures for both increasing and decreasing ratios. Accuracy is much lower for decreasing ratios. Figure 3B. Accuracy is roughly equal for S2 set sizes 20 and 30 for increasing directions, but lower for set size 30 than set size 20 for decreasing directions.
Figure 4. ERP averaged waveforms to S2 stimuli as a function of distance and change direction. Displayed are electrodes from lateral and medial sites of rows 1, 3, 5, and 7 (Fig. 1). ERP effects for numerical distance (top) were larger over fronto-central and central sites from 360–600 ms, with Close more negative than Medium, and Medium more negative than Far. Decreasing-direction ERPs were more negative than Increasing direction ERPs from 320–440 ms.
Figure 5. ERP difference waves comparing distance and direction effects by proficiency group
HP, Higher Proficiency; LP, Lower Proficiency. Distance difference waves show that Medium
distances elicited greater negativity than Far, and that this difference was similar for both HP
and LP groups. Although Close elicited greater negativities than Far for both groups, this
difference was much larger for the HP than for the LP group. Direction effects – Decreasing
more negative than Increasing – were similar for HP and LP groups. Displayed waveforms are
from electrode Cza, visible as the center emboldened electrode in Figure 7.
Figure 6. Mean ERP amplitudes to S2 between 360–440 ms as a function of distance and proficiency group. Note similar amplitude for Higher Proficiency and Lower Proficiency groups at Far and Medium distances, but different amplitude for Close distances. Mean amplitude is for all 32 electrodes used in analysis; error bars indicate standard error.
Figure 7. Topography and correlation of ERP distance waves (Close – Far) between 360–440 ms with numerical discrimination
Figure 7A: topography for all 20 participants. Distance effects appear strongest over anterior-central medial sites. Figure 7B: group differences for Close-Far amplitude in the topographies for Higher Proficiency (HP) and Lower Proficiency (LP) groups. Differences appear much larger for HP than for LP, with similar distributions. Figure 7C: correlation between ω and amplitude of Close-Far distance waves averaged across the seven emboldened electrode sites in Figs. 7A & 7B. Greater discriminability (smaller ω) is associated with larger Close-Far amplitude differences.