Title
Evidence for Modality-Specific Processes in Approximate Numerical Comparison

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Tokita, Midori
Ishiguchi, Akira

Publication Date
2012

Peer reviewed

Evidence for Modality-Specific Processes in Approximate Numerical Comparison

Midori Tokita (tokita.midori@ocha.ac.jp)
Akira Ishiguchi (ishiguchi.akira@ocha.ac.jp)
Ochanomizu University, Otsuka Bunkyo-ku,
Tokyo, 112-8610 Japan

Abstract

It has been claimed that a genuinely abstract number representation exists and is capable of representing the numerosity of any set of discrete elements irrespective of whether they are presented in visual or auditory modality. To test whether adults can compare large numerosities cross-modally as accurately as intra-modally, we measured Weber fractions and a point of subjective equality of numerical discrimination in the visual, auditory, and cross-modal conditions with use of a carefully controlled experimental procedure. Results showed distinct differences between the performances of the visual and the auditory condition in such way that numerical discrimination of the auditory sequence is more precise than that of visual sequence. Moreover, the performance of cross-modal trials differed among participants, with the exception that they were all worse than the auditory condition and that the number of visual stimuli was overestimated. Taken together, our findings implied that numerical discrimination of the auditory and visual stimuli mediates the modality-specific processes, suggesting that the numerical representation process can be complex of multiple stages.

Keywords: numerical discrimination; sensory modality; cross-modal comparison

Introduction

Many studies supported the idea that humans possess innate neural mechanisms that generate approximate, not precise, numerical representations. Results from studies of numerical competence in infants, young children, and nonhuman animals have shown that the approximate numerical system is evolutionally old and is equipped early in human development (e.g., Cantlon & Brannon, 2006; Feigenson, Dehaene & Spelke, 2004; Hauser, Tsao, Garcia, & Spelke, 2003; Whalen, Gallistel & Gelman, 1999). Furthermore, converging empirical findings from several areas of cognitive neuroscience argue for biological determined mechanisms for approximate number representation (e.g., Nieder & Dehaene, 2009; Piazza, 2010). At the same time, certain researchers have prompted extensive investigation over the processes of number representation in the behavioral and neurophysical field (e.g.,Kadosh, Lammertyn, & Izard 2008; Kadosh & Walsh, 2009).

One of the claims made by the proponents of abstract numerical representation is that the processing of approximate numerical representation is independent of sensory modality. They argued that abstract numerical representation could genuinely be capable of representing the numerical of any set of discrete elements, whether they were presented in the visual or auditory condition (Barth, Kanwisher, & Spelke, 2003; Gallistel & Gelman, 1992; Jordan & Brannon, 2006; Piazza, 2010). In these studies, it has been demonstrated that there was no cost of comparing numerosities across versus within visual and auditory stimulus sets. They claimed that the comparison across presentation modality was not performed using modality-specific numerical representations but rather using the true abstract numerical representation system. Evidence for modality-independent numerical representation ability has also been claimed in infants (e.g., Jordan & Brannon, 2006; Kobayashi, Hiraki & Hasegawa, 2005) and animals (Jordan, Brannon, Logothetis & Ghazanfar, 2005).

It has, however, remained unclear whether these approximate numerical representations are truly modality-independent. Three primary reasons exist for doubting the modality-independence of the approximate numerical representation. First, some evidence has shown that there were significant differences in the performance of numerical judgments for visual, auditory, and tactile senses (e.g., Riggs, Ferrand, Lancelin, Fryziel, & Dumur, 2006; Lechelt, 1975; Philippia, van Erp, & Werkhoven, 2008). For example, in the rapid counting experiment, Lechelt (1975) compared adult performance in numerosity judgment of visual, auditory, and tactile stimuli and demonstrated that perceived numerosity differed among modalities. Philippi, van Erp, & Wekhoven (2008) demonstrated that the stimuli with a short interstimulus interval (ISI) are underestimated and the tendency is stronger for visual than for auditory stimuli. Second, it is known that the processing of temporal information is much more efficient in the auditory than in the visual modality (Penny, Gibbon, & Meck, 2000; Ivry, 2008). For example, in time related tasks such as duration discrimination and empty interval estimation, the performance in the auditory presentation is significantly better than that in the visual and the tactile presentations (e.g., Grondin, 2010). As it has suggested that the temporal information affects the numerical discrimination (Tokita & Ishiguchi, 2011), there is the possibility that numerical discrimination among modalities differed when the experimental conditions are rigorously controlled. Third, limitations may exist within experimental procedures of empirical studies that claimed modality-independence of numerical representation in terms of control of stimuli, precision in measurement, and numbers of items tested. For example, Barth et al (2003) used a cross-modal comparison task and found that accuracy on these tasks was comparable to those on intramodal tasks, suggesting that non-numerical cues did not play a substantial role even in intramodal tasks. Numerical contrasts in their studies were, however, quite large such as Weber fraction of .50 or greater. With this level of measurement precision, the difference in the performance of each task could remain undetected. More to it, in infant and animal studies, the number of items tested was smaller than four. Because it remains unclear whether a system for representing small numbers of objects is distinct from that for representing larger numbers of objects, it is necessary to test whether the effects of sensory modality differ among a variety of numerosities.

In this study, we tested whether and how the numerical comparison of visual, auditory, and cross-modal presentation would differ under the adequate control of the concerns discussed above. We measured Weber fractions of discrimination task to assess the difference in the precision. Many studies have shown that both behavioral and neuronal tuning functions obey the Weber law (i.e., discriminability depends on the ratio of the numerical to be compared) over a broad range of numerosities (e.g., Burgess & Barlow, 1983; Nieder & Dehaene, 2009; Tokita & Ishiguchi, 2011; Whalen, Gallistel, & Gelman, 1999). We also measured a point of subjective equality (PSE) to test the accuracy of numerical comparison. Importantly, we involved rigid stimuli controls so that other properties such as stimuli duration and interval duration would not be confounded with the number of elements.

In Experiment 1, we compared the performance of numerical discrimination between the visual and the auditory presentation. In Experiment 2, we compared the performance of the visual, auditory, and cross-modal numerical comparison to examine how the numerical information in the different modality may integrate.

Experiment 1

We examined the precision of approximate numerical comparison in two sensory modalities: visual and auditory. The schematic view of stimuli presentation is shown in Figure 1.

In a visual condition, elements in a set were consisted of sequences of flashes, while in an auditory condition, elements in a set were presented in a tone sequence. We employed two levels of standard event numbers (i.e., standard number), 10 and 20, to test whether and how precision across presentation conditions would differ among standard numbers. To examine the precision, we obtained Weber fractions that indicate the participant's variance of numerical comparison. In deriving the Weber fractions, we used the method of constant stimuli in which participants in each trial decided which stimuli—standard or comparison—had more events.

Method

Participants Five participants participated in the experiment. All had normal or corrected-to-normal hearing and vision. All participants had no prior experience in numerical comparison tasks.

Design Two independent variables were examined in the experiment: the sensory modality (i.e., visual and auditory) and standard umber (i.e., 10 and 20). The numbers in the comparison stimuli for the standard number of 10 and 20 were "8, 9, 11, 12" and "16, 18, 22, 24", respectively. Trials in the visual and auditory conditions were separated and each constituted trial blocks. Two experimental conditions were presented among participants in a counterbalanced order. Trials in all the standard number sets were intermixed within a block. Each condition had 320 trials (40 repetitions × 4 comparison levels × 2 standard numbers), resulting in 640 trials in total. Each block had 64 trials, with 10 blocks in total. Participants performed three or four blocks in each experimental session, which took three days in total. Intermissions of approximately three minutes were given between blocks. Sequence of the trials was completely randomized within a block. Standard stimuli came first in half the trials and second in the remaining ones. Participants were given 16 practice trials before the actual experiment began.

Stimuli In the visual condition, two sequences of light gray dots appeared in a dark gray display region. Luminance of the dot was approximately 8 cd/m^2. In the auditory condition, two sequences of tone were presented fwith the built-in-speaker of the desktop computer at the intensity of about 60 dB (Sound pressure level). Auditory stimuli were 700 Hz pure sinusoidal sounds generated by Machintosh's computer.

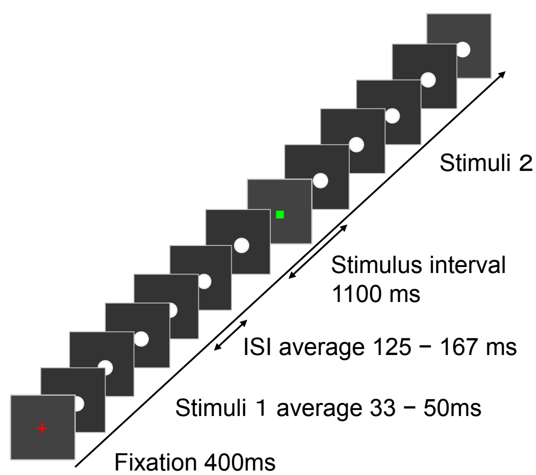

Figure 1: Schematic view of stimuli used in Experiment 1 and 2. The pair of events was sequentially presented in random orders.

In both conditions, we carefully controlled the stimulus duration and the inter-stimulus interval (ISI) so that the time for a sequence and presentation rate of stimuli would not be confounded with the number of elements. All element in a particular sequence had the same duration, but the durations varied from sequence to sequence between 33 to 50 ms. In half of the trials in a block, the average ISI was 125 ms in both standard and comparison sequences. In the remaining half, average ISI in the comparison sequences were carefully controlled so that average total interval for the standard sequence and that for the comparison sequence would be approximately equal. Thus, the number of events would be the only cue for numerical judgments. Many studies have provided evidence that the minimam ISI between two successive stimuli for correctly reporting their temporal order is about 40 ms and that this order threshold is invariant for auditory visual stimuli (e.g., Poppel, 1997; Kanabus, Szelag, & Poppel, 2002). Thus, sets of events in this experiment were perceived as successive independent of the sensory modality. To make the sequence aperiodic, we randomly added temporal jitter (−24, −17, −8, 0, 8, 17, or 24 ms) to each ISI so that the temporal rate would not constitute a rhythmic pattern.

Importantly, ISI were carefully determined so that the participants would not make judgments based on the verbal counting and/or temporal patterns. To make verbal counting impossible, the longest stimulus interval was set to be less than 250 ms, as previous studies have proved that participants could not rely on verbal or sub-verbal counting within that duration (e.g., Piazza, Mechelli, Price & Butterworth, 2006; Tokita & Ishiguchi, 2011).

Measurements The PSE and Weber fractions were measured using the method of constant stimuli. First, the number of events in comparison stimuli was plotted on the x axis and the proportion of "greater" response for each comparison stimulus was plotted on the y axis. The plotted data points constructed the psychometric function approximated by a cumulative Gaussian function, on which the difference threshold was obtained. This difference threshold was defined as the smallest amount of the element number change, for which a correct response rate of 75% was achieved. Weber fractions were obtained by dividing the difference thresholds by the standard numbers. The PSEs were obtained as the value of the location on the psychometric function at which the standard and comparative choice probabilities were equal to 50%. In this experiment, we obtained the standardized PSE, dividing the PSE by the standard number.

Procedure Participants sat in a darkened room at a distance of approximately 115 cm from the presentation screen. A numeric keypad was placed directly in front of the participants. The participants made responses by pressing the "1" or "3" key.

Each trial started with a red fixation cross for 400 ms followed by the first sequence. Pairs of sequences—standard and comparison sequences—were shown in succession in random order. The two sequences were separated by a stimulus interval of 1100 ms.

The participant's task was to choose which sequence, the first or second, contained more elements. Feedback with a short beep sound was given when participants made an incorrect choice. At the beginning of each session, the participants were explicitly instructed to attend to the number of elements presented and to discriminate on the basis of the numerical they felt, and not by verbal counting. They were also instructed to see the center of monitor in the auditory condition as well.

A Macintosh G4 computer was used to generate the display and the sound, and to record the data. Stimuli were presented on a color monitor at a refresh rate of 120 Hz (SONY Color Graphic Display Model GDM-F400).

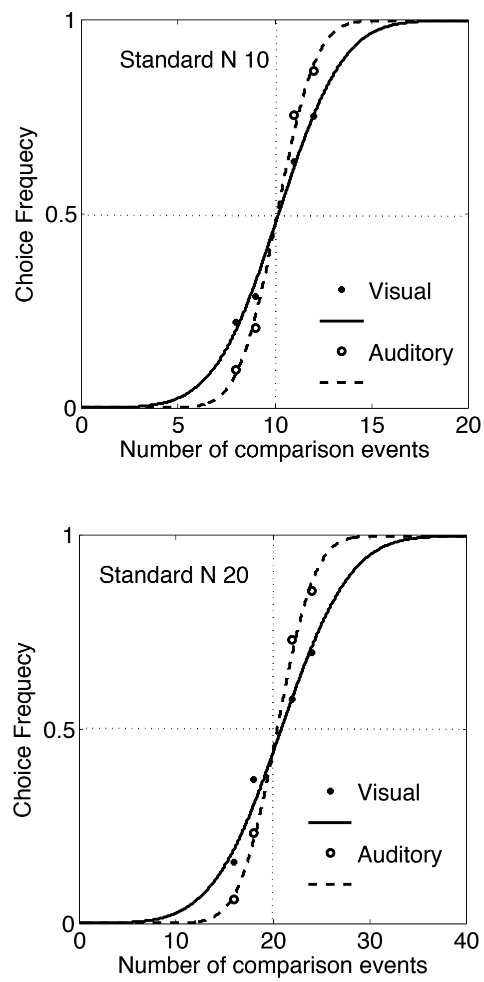

Figure 2: Average psychometric functions for the each presentation condition (a) standard number of 10 and (b) standard number of 20.

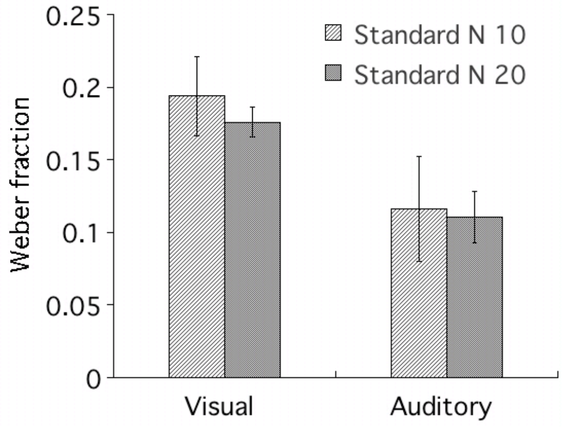

Figure 3: The means of Weber fractions in the visual and the auditory conditions at standard number of 10 and 20. Error bars represent standard deviations.

Results and discussion

Figure 2 shows the average psychometric functions for each standard number. Figure 3 shows the mean Weber fraction in each condition. The fits of data points to psychometric functions were generally good, and the Pearson moment correlation coefficient exceeded 0.9 in all cases with the exception of the visual condition at the standard number 10 and 20 for one participant. The data of the participant were excluded while those of the remaining participants were used for further analysis.

To test whether and how precision in numerical comparison differs between the visual and the auditory conditions, a 2 modality (visual and auditory) × 2 standard numbers (10 and 20) repeated measures analysis of variance (ANOVA) was conducted on individual Weber fractions. This yielded a significant main effect for presentation modality [F (1, 3) =90.38, p < .01]. Weber fractions in the auditory condition were significantly smaller than those in the visual condition, suggesting that the numerical judgment in auditory modality was more precise that that in visual modality. No significant effect of the standard numbers was observed [F (1, 3) = .48, p > .1], suggesting that the precision of numerical judgment was not affected by the number of elements within the numerical range tested in this experiment.

Experiment 2

We tested the precision of approximate numerical comparison in three presentation conditions (i.e., visual, auditory, and cross-format). Since no systematic difference was observed between the standard numbers, we only use one standard number 10 in this experiment. Stimuli presentations of the visual and the auditory conditions were the same as those in Experiment 1. In the cross-modal condition, elements in one set were presented in the visual sequence and those in the other set were presented in the auditory sequence. To examine the precision, we obtained Weber fractions that indicate the participant's variance of numerical comparison. To test the accuracy of the numerical comparison, we derived the point of subjective equality.

Method

Participants. Newly recruited five participants participated in the experiment. All participants had no prior experience in numerical comparison tasks. All had normal or corrected-to-normal hearing and vision.

Design We compared three presentation conditions: the visual, the auditory and the cross-modal condition. The cross-modal condition had two sub-conditions: the cross-modal condition 1 and the cross-modal condition 2. In cross-modal condition 1, standard stimuli were visual sequence and comparison stimuli were auditory sequence. In cross-modal condition 2, standard stimuli were auditory sequences and comparison stimuli were the visual sequences. The numbers in the comparison element for the standard number at 10 were "7, 8, 9, 11, 12, and 13".

Trials in the visual, auditory, and two cross-modal conditions were separated and each constituted trial blocks. Three experimental conditions were presented among participants in a pseudo-counterbalanced order. Each condition had 192 trials (32 repetitions × 6 comparison levels), resulting in 768 trials in total. Each block had 48 trials, with 16 blocks in total. Participants performed five to six blocks in each experimental session, which took three days in total. Intermissions of approximately three minutes were given between blocks. Sequence of the trials was completely randomized within a block. Standard stimuli came first in half the trials and second in the remaining ones. Participants were given 12 practice trials before the actual experiment began.

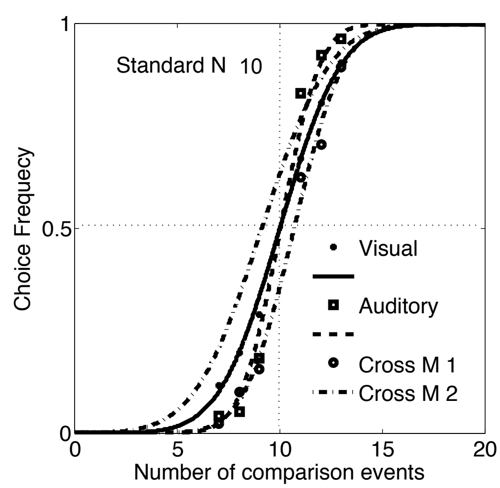

Figure 4: Average psychometric functions for the each presentation condition.

The stimuli, measurement, and procedures were the same as those in Experiment 1, with following exception. In the cross-modal condition, the auditory stimuli and the visual stimuli were shown in succession in random order.

Results and discussion

Figure 4 shows the average psychometric functions for each condition. The fits of data points to psychometric functions were generally good, and the Pearson moment correlation coefficient exceeded 0.9 in all cases. Figure 5 shows the mean of Weber fractions and that of standardized PSEs of all participants. We averaged over Weber fractions for two cross-modal conditions for all participants and used the data for further analysis.

In order to test whether and how precision in numerical comparison differs between the visual, the auditory, the cross-modal conditions, a 3 condition repeated measures ANOVA was conducted on individual Weber fractions. There was a significant main effect for presentation modality [F (2, 4) = 9.43, p < .01], and a Bonferroni post hoc analysis revealed that the Weber fractions in the visual and the cross-modal conditions were significantly larger than those in the auditory condition, indicating that precision in the visual and the cross-modal conditions was substantially worse than that in the auditory condition the same as the results in Experiment 1. The results suggested that the performance of the cross-modal trials would lie between that of the visual and auditory trials.

In order to test how cross-modal comparison affected the accuracy of numerosity comparison, we conducted a one-sample t test to compare the mean of the PSEs of the cross-modal condition 1 and that of the cross-modal condition 2 with the PSE of 0, respectively. The mean of PSEs in the cross-modal condition 1 was significantly larger than 0 [$t(4)=3.54$, p < .05] and the mean of PSEs in the cross-modal condition 2 was significantly smaller than 0 [$t(4)=-5.43$, p < .05]. The results showed that the number of visual stimuli was overestimated relative to that of auditory stimuli in both cross-modal conditions.

Discussion

We investigated whether and how precision in approximate numerical judgment between visual, auditory, and cross-modal presentations would differ. Our results demonstrated three significant findings. First, precision for numerical comparison of auditory sequence was significantly higher than that of visual sequence across two standard numbers. Second, precision in the visual and the cross-modal conditions was substantially worse than that in the auditory condition. Third, the number of visual elements was overestimated relative to that of auditory elements. Taken together, our results imply the existence of modality-specific processes in numerical comparison of the visual and auditory stimuli.

Our results are consistent with the previous studies that have shown the difference in counting precision across modalities (e.g., Lechelt, 1975). Lower precision in the visual presentation is also consistent with the results of those studies. It is noteworthy that the similar effects were observed between the counting task and numerical comparison task.

What is the source of difference in the precision in numerical representation between visually and auditory presented stimuli, and how does the discrepancy in precision occur? In any modality, or cross-modal condition, stimuli need to be successively enumerated across time when the items of a set are presented sequentially. In this condition, the cardinal value of stimuli can be represented by the last numerical quantity. Common aspects of those numerical judgment is that they are time related irrespective to the sensory modality. In other time related tasks such as duration discrimination and empty interval estimation, it is known that the performance in the auditory presentation is significantly better than that in the visual and the tactile presentations. Thus, it is predicted that the temporal resolution may cause the superiority of auditory modality in numerical judgments. Further investigations are necessary to explore the possibility.

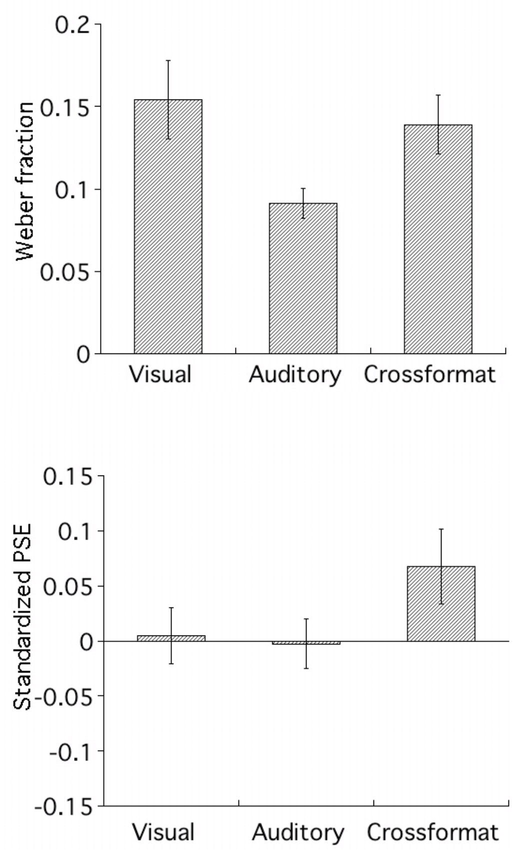

Figure 5: The means of Weber fractions and the means of PSEs of each modality condition. Error bars represent standard deviations.

As the performance of the cross-modal trials seemed to lie between that of the visual and auditory trials, it could be predicted that the convergent system could integrate the information from the auditory and visual numerical processing to form the higher abstract numerical presentations.

Another novel finding from this study is the overestimation of visual stimuli in the cross-modlaity comparison. Why did observer overestimate the number of visual stimuli relative to that of auditory stimuli? One possiblity it that observers may overestimate the number of events with greater uncertainty (i.e., visual stimuli) in the decisional process. Another possibility is that the observers may perceive the events at the faster rate more numerous; since the time estimation for visual stimuli is shorter than the auditory stimuli, the visual stimuli appear with faster rate when they are presented at the identical rate. To test this possibility, we need to examine how human compare numerosities across modality in further investigation.

In conclusion, this study provided evidence for modality-specific processes in approximate numerical representation in human adults. Although many studies support the idea that human adults as well as infants and non-human animals share the modality-independent numerical system, it remains unknown how numerical information from the modality-specific system is combined at the judgment stage. Our findings imply that the process of approximate numerical representation is complex and involves multiple stages.

References

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

Burgess, A., & Barlow, B. H. (1983). The precision of numerical discrimination in sequences of random dots. *Vision Research*, *23*, 811–820.

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for Ordering Small and Large Numbers in Monkeys and Humans. *Psychological Science, 17*, 401–406

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Gallistel, C.R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*, 43–74.

Groudin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72*, 561-582

Hauser, M. D., Tsao, F., Garcia, P., & Spelke, E. S. (2003). Evolutionary foundations of number: spontaneous representation of numerical magnitudes by cotton-top tamarins. *Proceedings for the Royal Society of London, 270*, 1441–1446.

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, *8*, 273–280.

Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *Proceedings for the national academy of sciences, 103*, 3486–3489.

Jordan, K. E., Brannon, E. M., Logothetis, N. K., & Ghazanfar, A. A. (2005). Monkeys match the number of voices they hear to the number of faces they see. *Current Biology*, 15, 1034–1038.

Kanabus, M., Szelag, E., Rojek, E., & Roppel, E. (2002). Temporal order judgment for auditory and visual stimuli. Acta Neurobiologiae Experimentalis, 62, 263-270.

Kadosh, R. C., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and brain Sciences*, *32*, 313–373.

Kadosh, R. C., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*, 132–147.

Kobayashi, T., Hiraki, K., & Hasegawa, T. (2005). Auditory-visual intermodal matching of small numerosities in 6-month-old infants. *Developmental Science, 8*, 409–419.

Lechelt, E. C. (1975). Temporal numerosity discrimination: Intermodal comparisons revisited. *British Journal of Psychology*, *66*, 101–108.

Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. The *Annual Review of Neuroscience, 32*, 185–208.

Penny, T. B., Gibbon, J., & Meck, W. H.(2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1770–1787.

Philippia, T. G., van Erp, J. B. F., & Werkhoven, P. J. (2008). Multisensory temporal numerosity judgment, *Brain Research*, *1242*, 116–125

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Science, 14*, 542–551.

Pöppel, E. (1997) A hierarchical model of temporal perception.*Trends in Cognitive Science*, 1, 56-61

Riggs, J. K., Ferrand, L., Lancelin, D., Fryziel, L., & Dumur, G. (2006). Subitizing in Tactile Perception. *Psychological Science*, *17*, 271–272.

Tokita, M., & Ishiguchi, A. (2011). Temporal information affects the performance of numerical discrimination: Behavioral evidence for a shared system for numerical and temporal processing. *Psychonomic Bulletin & Review*, *18,* 550–556.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in Human: The psychophysics of Number Representation, *Psychological Science, 10*, 130–137.