Behavioral and Functional Near-Infrared Spectroscopy (fNIRS) Study Testing for Shared Magnitude Processing for Numerosity and Proportion

Nick Yunqi Zhang[1], Patrick G. Kelly[2], Keanan J. Joyner[1], Heather Bortfeld[3], Silvia A. Bunge[1, 4], & Fei Xu[1]
[1]Department of Psychology, University of California, Berkeley
[2]Department of Psychology, University of California, Los Angeles
[3]Department of Psychological Sciences, University of California, Merced
[4]Helen Wills Neuroscience Institute, University of California, Berkeley

Abstract

A large body of research has investigated humans' representation of quantity through the Approximate Number System (ANS). In contrast, it is less clear how proportions are represented, and whether proportion processing is fully analogous to numerosity processing. In the current study, we compared numerosity and proportion discrimination within the same participants, while measuring cerebral hemodynamic responses using functional near-infrared spectroscopy (fNIRS). Behavioral results showed substantial commonalities between the two tasks, including highly comparable response accuracy, and shared ratio-dependent psychophysics. At the same time, proportion judgments were slower than numerosity judgments, suggesting additional processing demands beyond shared magnitude sensitivity. Neural evidence was less diagnostic, perhaps due to insufficient power, although a group-level conjunction analysis revealed overlapping parietal activation centered on the intraparietal sulcus for a general task contrast. Together, these findings speak to the possibility of shared magnitude processing substrate spanning number and proportion, while also cautioning the equivalence of the two forms of quantitative processing.

Keywords: Approximate Number System (ANS); numerosity; proportion; fNIRS; intraparietal sulcus (IPS).

Introduction

Humans extract quantitative information from sensory input in a remarkable and distinctive way. Even when exact quantities are inaccessible practically, people can still make reliable judgments about these quantities using approximate representations. A large body of developmental and comparative research attributes this capacity to an evolutionarily ancient magnitude processing system termed the Approximate Number System (ANS). The ANS has been observed in a wide range of animal taxa, from mammals like monkeys and rats (Jordan & Brannon, 2006; Meck & Church, 1983), to fish (Agrillo et al., 2008), and even invertebrates like honeybees (Howard et al., 2018). For human beings, the ANS emerges early in life: in newborns, infants can discriminate numerosities of 1:3 ratio; by 6 months, 1:2 ratio; by 9 months, 2:3 ratio (Izard et al. 2009; Xu & Spelke, 2000; Xu & Arriaga, 2007). This approximate number sensitivity continues to refine through early childhood (Halberda & Feigenson, 2008).

Across development and across species, the ANS is characterized by a set of key behavioral and neural signatures (for a review, see Odic & Starr, 2018). Behaviorally, numerical discrimination performance exhibits a ratio effect following Weber's law (Pica et al., 2004), such that comparisons of quantities closer to 1:1 ratio are more difficult than larger ratios. Neurally, convergent evidence links numerical processing to a frontoparietal network, where the posterior parietal cortex, specifically the bilateral intraparietal sulcus (IPS), is of particular importance for analog magnitude representation (Piazza et al., 2004; Cantlon et al., 2006; Ansari & Dhital, 2006; Hyde et al., 2010). A ratio effect was also identified, such that harder numerical discrimination (closer to a 1:1 ratio) elicits greater neural activation.

Another line of work reveals an early emerging sensitivity to probabilistic information. As young as 6 months, human infants can anticipate probable outcomes versus improbable ones (Denison et al. 2013; Teglas et al. 2007; Xu & Garcia, 2008). Underpinning such capacity is humans' early sensitivity to proportions and their ability to utilize such information to guide their actions (Denison & Xu, 2014). Much like their numerosity discrimination, 6-months-olds were also capable of differentiating proportions at the same 1:2 ratio (McCrink & Wynn, 2007). Furthermore, discrimination of proportions also exhibits a ratio-dependent behavioral signature reminiscent of the ANS (O'Grady & Xu, 2020; Qu et al., 2024).

Yet, diverging views have been put forth regarding the nature of proportion processing in relation to the ANS. On one view, neuroimaging results suggest that the brain represents proportions in the same magnitude encoding with the ANS, evoking ratio-dependent activations in bilateral IPS (Jacob & Nieder, 2009). Conversely, Matthews, Lewis, and Hubbard (2015) caution against equating the two systems and proposed a separate ratio-processing system (RPS), as performance in ratio related tasks predicted mathematical performance independently of ANS acuity, while Qu et al. (2024), on the other hand, countered that an ANS-based explanation may be more parsimonious.

A key limitation of the existing literature is that numerosity and proportion processing have rarely been examined side by side within the same participants, thus rendering it difficult to determine how the two processes correspond. To address this gap, the present study compared numerosity and proportion discrimination within the same adults and quantified their similarity across multiple levels of analysis. Behaviorally, we measured accuracy and response time (RT) and fit

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

psychometrics to test the ratio-dependent ANS signature. Neurally, we used functional Near-Infrared Spectroscopy (fNIRS) to probe participants' cerebral hemodynamics and tested if the two tasks evoked similar neural profiles.

Methods

Participants

A total of 62 adults (M_{Age} = 21.66 years, *SD* = 3.75; 39 females) participated in the study. Three participants were excluded from the study for poor behavioral performance, defined as overall accuracy falling more than 2 SD below the sample mean. Two additional participants were excluded for incomplete sessions or missing data, yielding a final analysis sample of 57 participants (M_{Age} = 21.26 years, *SD* = 3.03; 36 females). Participants were students enrolled in a psychology course and received course credit for their participation. All participants gave written informed consent prior to the experiment, and all experimental procedures were approved by the Institutional Review Board at the University of California, Berkeley. All participants had normal or corrected to normal vision, no pre-existing neurological or sensorimotor deficits, and no history of self-reported major psychiatric illnesses.

Stimuli

Numerosity Discrimination Task Stimuli for the numerosity discrimination task were images (922 × 461 pixels) pre-generated using a customized Python script (v. 3.9.6; Van Rossum & Drake, 2009). Each image displayed two randomized dot arrays displayed side by side, with one array containing more dots than the other. Stimuli were grouped into three conditions based on the numerosity ratio of the arrays, which were 1.2, 1.5, and 2.0 respectively (Fig. 1). To increase variability in absolute numerosity while preserving ratio structure, the numerosities were randomly generated under two bounded ranges: for half of the stimuli, the total number of dots were constrained to 6–18, and for the other half, the total number of dots were constrained to 12–36. Critically, to prevent reliance on low-level visual heuristics, stimuli were generated under two counterbalanced regimes: each pair of arrays was either equal in mean or total area of the dots, preventing participants from using area as an accurate heuristic. Finally, individual dot sizes varied within a bounded range (±50% of mean area) to further prevent participants from using a salient visual cue in the task.

Proportion Discrimination Task Stimuli for the proportion discrimination task were structurally similar to those used in the numerosity task but included two colors of dots within each array: target dots (red) and non-target dots (gray). The same three ratio conditions (1.2, 1.5, 2.0) were maintained, defined by the proportion of target dots rather than absolute numerosity (see Fig.1). To discourage a numerosity-based strategy, the total number of dots in the two arrays was randomized within the same bounded ranges as the numerosity discrimination task. This ensured that the arrays with higher target proportion are not always the arrays with higher total number of target dots. Surface area controls identical to those in the numerosity task were also applied.

Procedure

Participants were seated approximately 50 cm away from a monitor (1024 × 768 pixels resolution). Following informed consent and completion of a demographic survey, experimenters prepared and fit the fNIRS cap while providing task instructions. Cap placement was guided by anatomical fiducial landmarks, including the inion, nasion, and preauricular points, and alignment was verified using the international 10–20 Cz landmark (Klem et al., 1999). Hair was adjusted as needed to ensure clear optode contact with the scalp. Prior to data collection, signal quality was evaluated using the Aurora fNIRS software (NIRx Medical Technologies, LLC). Channels failing quality criteria were iteratively improved through cap adjustment and hair

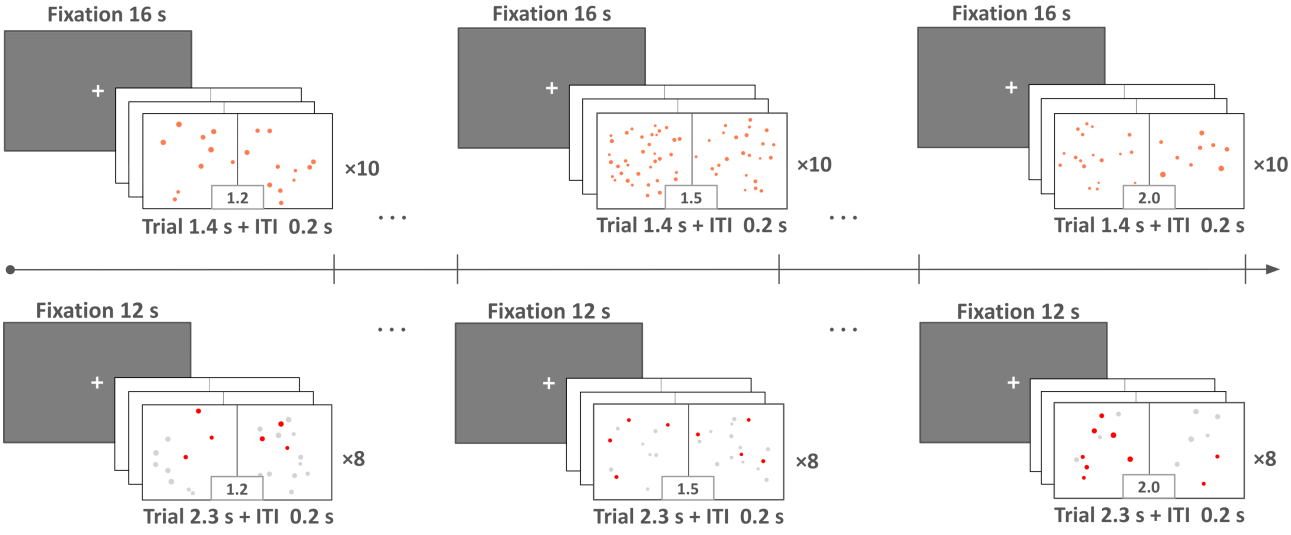

Figure 1: Schematic illustration of the study procedure with stimuli for each condition.

repositioning until no further improvement could be achieved. A black shower cap was then placed over the fNIRS cap to reduce the influence of ambient light.

Participants completed both the numerosity and the proportion discrimination tasks, in a counterbalanced order across subjects. Each task lasted for 16 minutes, with a 5–10 minutes break in between. Before each task, participants completed a 10-trial practice block with experimenter guidance to ensure that they understood the instructions, and responses were made via keyboard button presses (F, J). Together with fNIRS preparation, the total duration of the study was about 60 minutes.

Both tasks used a block design to optimize the fNIRS measurement. Each task consisted of 30 blocks (10 per ratio condition) presented in random order (see Fig.1). For the numerosity discrimination task, each block contained 10 trials. For each trial, the stimulus arrays were displayed for 1400 ms, during which participants were required to respond, followed by a 200 ms inter-trial interval. Before and after each task block, a fixation cross was displayed for 16 s. Participants were instructed to indicate which side had more dots. In the proportion task, we increased stimulus duration while reducing trials per block to maintain equal task duration overall. Each block contained 8 trials, in which the stimulus arrays were displayed for 2300 ms, allowing for response, followed by the same 200 ms inter-trial interval. Before and after each task block, a fixation cross was displayed for 12 s. Participants were instructed to indicate which side had a higher proportion of red dots.

Apparatus

Hemodynamic responses were recorded using a continuous-wave NIRSport2 system (NIRx Medical Technologies, LLC) with dual wavelengths (760 and 850 nm) and sampled at 5.1 Hz. Hemodynamic data was collected via Aurora fNIRS software (NIRx). The montage consisted of 16 sources and 15 detectors, with one detector used for eight short-separation channels, yielding 38 total channels covering lateral prefrontal cortex and posterior parietal cortex (see Fig. 2 A).

Optode layout was designed using fOLD (fNIRS Optodes' Location Decider; Morais et al., 2018). Because fOLD does not directly label some anatomical targets (e.g., DLPFC, IPS), MNI coordinates from prior fMRI studies were converted to their closest 10–10/10–20 scalp labels (Koessler et al., 2009) to support coverage of the targeted ROIs.

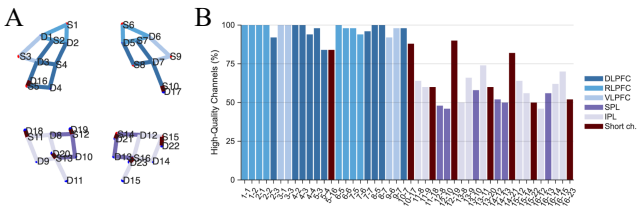

Figure 2: A. fNIRS montage covering frontoparietal ROIs; B. signal quality assessment indicating percentage of high-quality channels of the full sample.

Analysis

Behavioral Analysis Behavioral analyses were conducted in Python (v. 3.9.6) using pandas and SciPy. Performance was assessed using accuracy and response time (RT). Trials with missing responses were coded as incorrect for accuracy analyses and excluded from RT analyses. Participants with accuracy more than 2 SD below the sample mean were excluded from subsequent analysis.

A psychophysical function was used to further characterize participants' behavioral signature as in the ANS literature (Halberda & Feigenson, 2008; Pica et al., 2004). Specifically, percentage correct was modeled as 1 − error rate, where the error function assumed Gaussian approximate number representations of the two compared sets. The model was included a single free parameter, the Weber fraction (w), such that a smaller w indicates higher ANS acuity:

$$P(correct) = 1 - \frac{1}{2}\, erfc\left(\frac{|n_2 - n_1|}{\sqrt{2}\, w \sqrt{n_1^2 + n_2^2}}\right)$$

Neuroimaging Analysis fNIRS data were processed and analyzed in MATLAB (Version 25.2, 2025b, MathWorks Inc., Massachusetts) using the NIRS Brain AnalyzIR Toolbox (Santosa et al., 2018) and custom scripts (available upon request). Raw light-intensity data was first converted into optical density and corrected for motion artifacts using temporal derivative distribution repair (TDDR). Data quality was assessed on a channel-by-channel basis using the QT-NIRS module (Montero-Hernandez & Pollonini, 2019) with thresholds set to qThreshold = 0.6, SCI = 0.6, and PSP = 0.1. Channels failing to meet these criteria were flagged as low quality and excluded from subsequent analyses (numerosity: $M_{excluded}$ = 11.47, SD = 8.84; proportion: $M_{excluded}$ = 10.80, SD = 8.39; Fig. 2 B). Optical density data from retained channels were converted to oxy- (HbO) and deoxy-hemoglobin (HbR) concentration changes using the modified Beer–Lambert law (MBLL) with a partial pathlength factor of 0.1. A wavelet-based filter was then applied to reduce slow drifts and systemic noise. All inferential analyses focused on HbO due to its generally higher signal-to-noise ratio relative to HbR.

Subject-level statistics for HbO concentration were estimated using a general linear model (GLM) in which stimulus onsets were convolved with a canonical hemodynamic response function (HRF) and its temporal and dispersion derivatives. Short-separation channel time-series were included as nuisance regressors to reduce superficial physiological contamination. The model was fit using an autoregressive iteratively reweighted least squares approach (AR-IRLS), which applies pre-whitening to address serially correlated noises and robust regression to down-weight motion-related outliers (Santosa et al., 2018). The resulting beta (β) coefficients from this model were used as subject-level activation estimates. Group-level task-evoked activation was summarized using linear mixed-effects models with task conditions as the fixed effect and participant as the random intercept ($\beta \sim -1 + \text{cond} + (1|\text{subject})$). A weighted regression was performed on the group-level model using subject-level covariance matrices to down-weight the

noisy subjects. To examine task-evoked activation by ratio conditions, the same mixed-effects model was fit with the categorical ratio conditions of 1.2, 1.5, and 2.0. To test for the group-level ratio effect, a mixed-effects model was used with the numerical ratio as the fixed effect and participant as the random intercept ($\beta \sim$ ratio + (1|subject)).

Given the prevalence of excluded channels, HbO time-series were first aggregated within predefined regions of interest (ROIs) to produce ROI-level signals for connectivity analyses. The ROI time-series were then cleaned by regressing out physiological noise measured in short-separation channels, yielding residual signals used for connectivity estimation. Functional connectivity was computed using robust correlation with autoregressive innovations to reduce sensitivity to temporal autocorrelation and outliers. Group-level effects on connectivity were assessed using mixed-effects models fit to correlation coefficients ($r \sim -1$ + cond + (1|subject)).

Statistical significance was evaluated at $p < .05$. Post-hoc tests were corrected for multiple comparisons using the Benjamini-Hochberg (1995) procedure.

Results

Behavioral Results

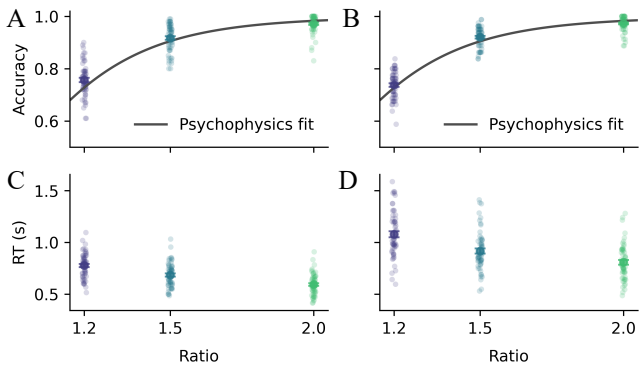

Figure 3: Task performance by condition. A–B. Numerosity and proportion task accuracy with psychophysics fit overlaid. C–D. Numerosity and proportion task RT.

Accuracy was high across tasks, with no evidence for a task difference, (numerosity: $M = .88$, $SD = .04$; proportion: $M = .88$, $SD = .03$; $t(56) = 0.72$, $p = .48$). In contrast, average RTs differed by task: participants responded faster in the numerosity task ($M = 701$ ms, $SD = 119$) than in the proportion task ($M = 949$ ms, $SD = 205$; $t(56) = -12.70$, $p < .001$). Importantly, mean RTs in both tasks were under 1 s, making it unlikely that participants relied on counting. Linear mixed-effects models with participants as random intercept revealed a main effect of ratio for both accuracy ($b = 0.25$, $SE = 0.014$, $p < .001$) and RT ($b = 0.22$, $SE = 0.022$, $p < .001$), and a task × ratio interaction for RT ($b = -0.11$, $SE = 0.031$, $p < .001$), indicating a more negative, stronger ratio effect in the proportion than the numerosity task. The same task × ratio interaction was not observed for accuracy ($b = 0.023$, $SE = 0.019$, $p = .224$). Importantly, we observed significant correlations for participants' mean accuracy ($r(55) = .34$, $p = .009$) and RT ($r(55) = .66$, $p < .001$) across tasks.

A paired t-test showed no significant difference ($t(56) = -0.12$, $p = .908$) between participants' Weber fractions (w) between tasks (numerosity: $M = .21$, $SD = .06$; proportion: $M = .21$, $SD = .04$), and psychometric fits (Fig. 3 A–B) were excellent in both tasks (numerosity: $R^2 = .96$; proportion: $R^2 = .99$). Moreover, there was a moderate but significant correlation across participants between w from the two tasks ($r(55) = .32$, $p = .015$).

Neuroimaging Results

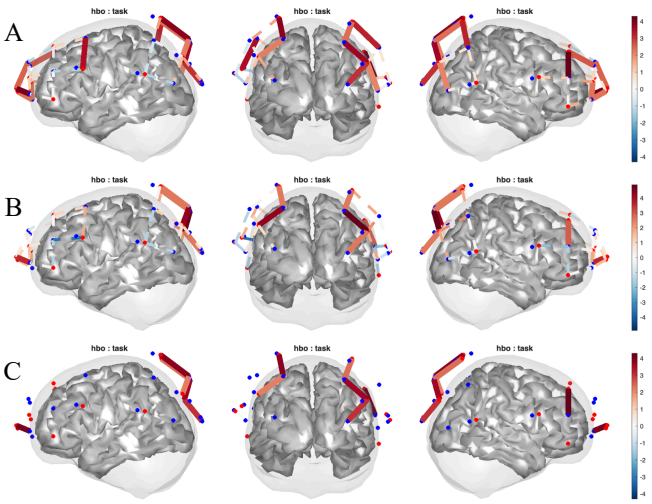

Figure 4: Channel level parametric maps. Left lateral, posterior, and right lateral views show task vs. rest activation for the numerosity task (A), proportion task (B), and their conjunction (C). Solid lines indicate significant channels with FDR corrected $p < .05$. Color bars indicate t-statistics ranging from −4 to 4.

Task Activation Both the numerosity and the proportion discrimination tasks elicited robust task-evoked activation across prefrontal and parietal cortices relative to baseline (see Fig. 4 A–B), with the strongest and most consistent effects localized to bilateral IPS. The conjunction analysis revealed substantial overlap between the tasks in bilateral IPS (Fig. 4 C), indicating a shared core network supporting both forms of magnitude processing. Outside of the parietal cortex, overlapping activation was comparatively limited, confined to two prefrontal channels, one in right DLPFC and one in left RLPFC, indicating comparatively focal prefrontal overlap. Overall, these results suggest that the two tasks recruit parietal cortex centered on IPS, with sparser and more task-variable recruitment of prefrontal regions. Further research is needed to test whether this level of overlap is greater than for numerical vs. non-numerical tasks.

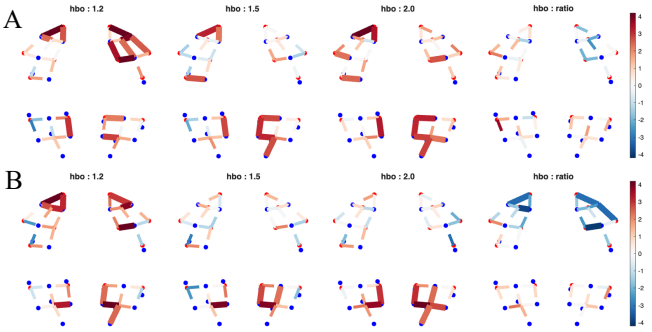

Figure 5: Channel level parametric maps show numerosity (A) and proportion (B) activations by ratio conditions of 1.2, 1.5, and 2.0, along with the corresponding ratio effect.

Task Activation by Condition and Ratio Effect When examined separately by condition (Fig. 5), both tasks showed broadly similar activation topographies across prefrontal and parietal cortex. However, ratio-dependent modulation of task-evoked activation was relatively limited (Fig. 5, the 4th column). Notably, despite robust overall task activation, bilateral IPS channels did not exhibit reliable ratio effects. In contrast, ratio-related modulation—reduced activation for higher (i.e., easier) ratios—was evident in a subset of bilateral prefrontal channels during the proportion task, whereas ratio effects in the numerosity discrimination task were weaker and less consistent. This neural pattern parallels the behavioral findings in response time, where ratio effects were also more pronounced in the proportion task. Thus, under the present design, ratio sensitivity was more apparent in the prefrontal cortex, particularly during proportion judgments, whereas IPS activation appears robust but comparatively insensitive to ratio manipulations.

Function Connectivity ROI-to-ROI connectivity analysis revealed strong within-ROI and contralateral homotopic connectivity for both tasks (Fig. 6). Additionally, moderate frontoparietal connectivity was observed, particularly between left parietal and right prefrontal regions, suggesting coordinated activity between core magnitude-processing and higher-order control systems.

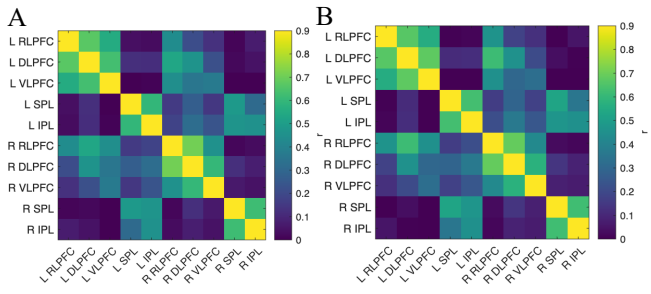

Figure 6: ROI-to-ROI functional connectivity matrices of the numerosity task (A) and the proportion task (B).

Neural Similarity To quantify correspondence between numerosity and proportion processing beyond spatial overlap, we computed cross-task similarity for both task-evoked activation and functional connectivity (Fig. 6). For activation, similarity was assessed by correlating channel-wise task vs. rest *t*-statistics across tasks. At the group level, activation patterns were strongly correlated, $r(36) = .75$, $p < .001$, indicating similar spatial profiles of task-evoked activation. At the participant level, repeated-measures correlation revealed a moderate but reliable correspondence in activation patterns, $r_{rm}(1286) = .34$, $p < .001$.

For functional connectivity, similarity was assessed by correlating ROI-to-ROI connectivity matrices across tasks. Group-level connectivity patterns were highly correlated, $r(43) = .99$, $p < .001$, and participant-level connectivity estimates showed strong cross-task consistency, $r_{rm}(1055) = .83$, $p < .001$.

Discussion

The present study directly compared numerosity and proportion discrimination within the same participants using matched behavioral and neural measures. Behaviorally, the two tasks showed substantial commonality: accuracy was nearly identical, both tasks showed robust ratio-dependent modulation that was well captured by the same Weber-based psychometrics, individual performance and Weber fractions were also significantly correlated across tasks. Neurally, both tasks elicited robust task-evoked activation across frontoparietal cortex, with the most consistent overlap localized to bilateral IPS; activation topographies and functional connectivity patterns were also similar across tasks. However, as ratio-dependent modulation was limited in the neural data, these cross-task similarities are best interpreted as evidence for shared frontoparietal engagement rather than shared magnitude encoding per se. Taken together, these findings suggest substantial overlap between numerosity and proportion discrimination, with ratio-dependent psychophysics providing the most direct evidence for a shared magnitude-processing architecture and neural similarity providing convergent but less diagnostic support.

These findings are consistent with prior neuroimaging work linking magnitude processing to a frontoparietal network and suggesting that computing proportions engages overlapping circuitry. In particular, Jacob and Nieder (2009), using an fMRI adaptation paradigm, reported evidence that bilateral IPS encoded nonsymbolic proportions in a similar analog magnitude representation to that of numerosities. The present study extends this literature in two ways. First, it compared numerosity and proportion discrimination within the same participants, allowing cross-task correspondence to be tested directly in matched decision tasks rather than a passive adaptation setting. Second, it quantified similarity across multiple levels of analysis from task performance, Weber-based psychophysics, activation topographies, to functional connectivity, providing a more detailed characterization of where numerosity and proportion processing converge.

It is this approach that supports a more nuanced view. While commonalities in behavioral and neural data were substantial, several features of the data point to meaningful divergence superimposed on that shared core. Behaviorally, proportion judgments were significantly slower than numerosity judgments despite comparable accuracy. RTs also showed a more pronounced ratio dependence in the proportion task, suggesting that increasing difficulty disproportionately slows proportion discriminations. This dissociation suggests that the proportion task imposes additional processing demands, potentially reflecting the need to compute, maintain, or compare complicated part-to-whole relations rather than simple numerosities. Neurally, ratio-dependent modulation was minimal in IPS despite robust parietal task activation, whereas ratio effects emerged more clearly in a subset of bilateral prefrontal channels during the proportion task (and were weaker for the numerosity task). One interpretation is that proportion discrimination more strongly recruits prefrontal mechanisms associated with executive control and decision processes such as integrating multiple sources of information, maintaining task-relevant representations, or resolving competition between partially diagnostic cues. Under this account, IPS provides a shared substrate, potentially supporting magnitude-related processing, while prefrontal engagement reflects task-dependent control demands that are amplified when judgments require proportion-based inference.

Our individual-level data further underscore partial, rather than complete, correspondence in the two tasks. Specifically, participant-level correlations in ANS acuity and activation topography were significant but only moderate in magnitude. However, this should not be taken as evidence against shared mechanisms. Instead, it likely reflects a combination of imperfect psychometric reliability in behavioral and neural measures, measurement noise inherent to fNIRS (especially in posterior and parietal regions), and potential time-on-task effects (fatigue) that add variance unrelated to the constructs of interest. In contrast, the much stronger participant-level correspondence observed for functional connectivity suggests that network organization may be a more stable within-person feature than fine-grained activation amplitude patterns or estimated psychophysical parameters in the current design (Kelly et al., 2026), though potentially also less diagnostic of the neural signature of interest.

This same methodological perspective is important for interpreting the limited evidence of parietal ratio modulation. In fMRI, IPS ratio effects are often treated as a hallmark signature of ANS-like computations. However, subtle parametric modulations can be difficult to detect with fNIRS given lower spatial specificity, vulnerability to systemic noise, and strong dependence on optode–scalp coupling—limitations that are especially pronounced in posterior cortex (Beauchamp et al., 2011; Skau et al., 2022). In the present study, the prevalence of low-quality parietal channels likely reduced sensitivity to ratio effects even when robust task activation was clearly present. This interpretation aligns with prior fNIRS findings (e.g., Horaguchi et al., 2008), where reliable parietal activation was observed without detectable ratio modulation. Accordingly, the absence of IPS ratio modulation here is most parsimoniously interpreted as a limitation in sensitivity rather than strong evidence that ratio-dependent magnitude coding is absent in parietal cortex. Nevertheless, this limitation constrains the strength of inferences that can be drawn from the activation results. Specifically, while task-evoked overlap and group-level topographic similarity indicate shared frontoparietal engagement, the lack of reliable IPS ratio modulation in the present paradigm constrains our interpretation to a shared network recruitment, perhaps only suggestive of, rather than clearly evidencing for a shared magnitude coding mechanism per se.

Several limitations should be considered when interpreting these findings. First, posterior channel signal quality constrained inference about IPS and parietal subregions, reducing power to detect more subtle effects such as parametric ratio modulation. Beyond being a technical issue, this limitation has equity implications: fNIRS signal quality can vary systematically with hair characteristics and skin pigmentation, potentially producing unintended under-representation of some groups and reducing generalizability. Second, although each task lasted only 16 minutes, the full session was approximately 60 minutes including fNIRS setup and both tasks, which may have introduced fatigue or reduced attention and thereby attenuated within-task reliability and cross-task associations. Follow-up studies could distribute tasks across separate visits, increase the number of trials and blocks per condition, optimizing session duration by balancing measurement stability, statistical power, and participant burden. These design changes would, in return, enable more direct assessments of reliability (e.g., split-half estimates in a larger dataset or test-retest across sessions), helping to determine whether moderate cross-task correlations and the weak neural ratio modulation reflect true features of the data or are constrained by measurement noise.

In summary, numerosity and proportion discrimination exhibited similar behavioral signatures that were well described by a common ANS framework. The fNIRS results provided complementary evidence for overlapping frontoparietal engagement, particularly within bilateral IPS, alongside highly similar activation topographies and functional network organization across tasks. While the neural results remain inconclusive, particularly with the absence of IPS ratio modulation, these results speak to the possibility of a shared magnitude-related system that spans numerosity and proportion, while also indicating task-specific demands superimposed on that shared architecture, most notably slower response and more pronounced prefrontal ratio modulation during proportion discrimination.

References

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. Animal Cognition, 11(3), 495–503.

Ansari, D., & Dhital, B. (2006). Age-related Changes in the Activation of the Intraparietal Sulcus during Nonsymbolic Magnitude Processing: An Event-related Functional Magnetic Resonance Imaging Study. Journal of Cognitive Neuroscience, 18(11), 1820–1828.

Beauchamp, M. S., Beurlot, M. R., Fava, E., Nath, A. R., Parikh, N. A., Saad, Z. S., Bortfeld, H., & Oghalai, J. S. (2011). The developmental trajectory of brain-scalp distance from birth through childhood: Implications for functional neuroimaging. *PloS One*, *6*(9), e24981.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300.

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional Imaging of Numerical Processing in Adults and 4-y-Old Children. PLoS Biology, 4(5), e125.

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5-and 6-month-olds. Developmental Psychology, 49(2), 243.

Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. Cognition, 130(3), 335–347.

Gorniak, S. L., & Pollonini, L. (2026). Developing an inclusive approach to fNIRS research in women of color. Neurophotonics, 13(S1), S13003–S13003.

Horaguchi, T., Ogata, Y., Watanabe, N., & Yamamoto, M. (2008). Behavioral and near-infrared spectroscopy study of the effects of distance and choice in a number comparison task. Neuroscience Research, 61(3), 294–301.

Hyde, D. C., Boas, D. A., Blair, C., & Carey, S. (2010). Near-infrared spectroscopy shows right parietal specialization for number in pre-verbal infants. NeuroImage, 53(2), 647–652. https://doi.org/10.1016/j.neuroimage.2010.06.030

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. Proceedings of the National Academy of Sciences, 106(25), 10382–10385.

Jacob, S. N., & Nieder, A. (2009). Notation-independent representation of fractions in the human parietal cortex. Journal of Neuroscience, 29(14), 4652–4657.

Jordan, K. E., & Brannon, E. M. (2006). A common representational system governed by Weber's law: Nonverbal numerical similarity judgments in 6-year-olds and rhesus macaques. Journal of Experimental Child Psychology, 95(3), 215–229.

Kelly, P. G., Bortfeld, H., Joyner, K. J., & Bunge, S. A. (2026). Evaluating the reliability of functional near-infrared spectroscopy data in the context of a reasoning paradigm. bioRxiv, 2026.01.16.699971.

Klem, G. H., Lüders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. Electroencephalography and Clinical Neurophysiology. Supplement, 52, 3–6.

Koessler, L., Maillard, L., Benhadid, A., Vignal, J. P., Felblinger, J., Vespignani, H., & Braun, M. (2009). Automated cortical projection of EEG sensors: anatomical correlation via the international 10–10 system. *Neuroimage*, *46*(1), 64–72.

Lewis, M. R., Matthews, P. G., & Hubbard, E. M. (2016). Neurocognitive architectures and the nonsymbolic foundations of fractions understanding. In Development of mathematical cognition (pp. 141–164). Elsevier.

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. Psychological Science, 18(8), 740–745.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. Journal of Experimental Psychology: Animal Behavior Processes, 9(3), 320.

Odic, D., & Starr, A. (2018). An introduction to the approximate number system. Child Development Perspectives, 12(4), 223–229.

O'Grady, S., & Xu, F. (2020). The development of nonsymbolic probability judgments in children. Child Development, 91(3), 784–798.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. Neuron, 44(3), 547–555.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. Science, 306(5695), 499–503.

Qu, C., Clarke, S., Luzzi, F., & Brannon, E. (2024). Rational number representation by the approximate number system. Cognition, 250, 105839.

Skau, S., Helenius, O., Sundberg, K., Bunketorp-Käll, L., & Kuhn, H.-G. (2022). Proactive cognitive control, mathematical cognition and functional activity in the frontal and parietal cortex in primary school children: An fNIRS study. *Trends in Neuroscience and Education*, *28*, 100180.

Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. Proceedings of the National Academy of Sciences, 104(48), 19156–19159.

Van Rossum, G., & Drake, F. L. (2009). Python 3 reference manual. Scotts Valley, CA: CreateSpace, 926.

Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. British Journal of Developmental Psychology, 25(1), 103–108.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. Proceedings of the National Academy of Sciences, 105(13), 5012–5015.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. Cognition, 74(1), B1–B11.

Zimeo Morais, G. A., Balardin, J. B., & Sato, J. R. (2018). fNIRS Optodes' Location Decider (fOLD): A toolbox for probe arrangement guided by brain regions-of-interest. Scientific Reports, 8(1), 3341.