AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Human Perception and Performance

ISSN: 0096-1523

2023, Vol. 49, No. 1, 108–128
https://doi.org/10.1037/xhp0001062

Investigating the Relationships Between Temporal and Spatial Ratio Estimation and Magnitude Discrimination Using Structural Equation Modeling: Evidence for a Common Ratio Processing System

Rebekka Lagacé-Cusiac, Paul F. Tremblay, Daniel Ansari, and Jessica A. Grahn
Department of Psychology, University of Western Ontario

Humans perceive ratios of spatial and temporal magnitudes, such as length and duration. Previous studies have shown that spatial ratios may be processed by a common ratio processing system. The aim of the current study was to determine whether ratio processing is a domain-general ability and consequently involves common processing of temporal and spatial magnitudes. Two hundred seventy-five participants completed a battery of spatial and temporal ratio estimation and magnitude discrimination tasks online. Structural equation modeling was used to analyze the relationship between ratio processing across domains while controlling for absolute magnitude discrimination ability. The four-factor higher order model, consisting of spatial and temporal magnitude and ratio processing latent variables, showed adequate local and global fit, $\chi^2(44) = 41.41$, $p = .626$, root mean square error of approximation = .000. We found a significant relationship ($r = .63$) between spatial and temporal ratio processing, suggesting that ratio processing may be a domain-general ability. Additionally, absolute magnitude processing explained a large part (60–66%) of the variance in both spatial and temporal ratio processing factors. Overall, findings suggest that representation of spatial and temporal ratios is highly related and points toward a common ratio processing mechanism across different types of magnitudes.

Public Significance Statement

Many of our day-to-day activities require us to process ratios. For example, we often use progress bars to assess progress on a certain task (e.g., downloading a file). These ratios can take the form of both ratios in space and time. We empirically investigated people's ability to perceive and represent ratios in both space and time. Our results show that the ability to perceive and represent spatial and temporal ratios is highly related. This has implications for theories of both absolute and relative magnitude processing for both spatial and temporal magnitudes.

Keywords: nonsymbolic spatial and temporal ratio processing, proportional reasoning, number line estimation, a theory of magnitude (ATOM), structural equation modeling (SEM)

Supplemental materials: https://doi.org/10.1037/xhp0001062.supp

Which queue is the shortest at the movies? How much time will it take to run an errand? Many of our day-to-day decisions are based on numerical information. While some decisions are based on *absolute* magnitudes, such as a single length or duration, other decisions are based on *relative* magnitudes (i.e., the magnitude of the relationship between two absolute magnitudes, otherwise known as proportions or ratios). For example, we can easily tell from the battery icon on our electronic device how much charge is left by comparing the length of the filled bar to the total length of the battery icon, regardless of the overall size. The ratio between two lengths is an example of a relative magnitude in the spatial domain. However, humans also use ratios in the temporal domain. For example, rhythms in Western music are commonly composed of notes with proportional durations, which is why we can recognize tunes despite tempo changes. When a tune is slowed down, all durations are lengthened such that the relative, or proportional, relationships are maintained. Furthermore, some studies have shown that adults can accurately represent, on a bounded line, temporal

This article was published Online First October 20, 2022.

Rebekka Lagacé-Cusiac https://orcid.org/0000-0003-1322-7792

Preregistration for this study can be found at https://osf.io/h6gts. Materials, code, and data can be found at https://osf.io/jbqta/. We have no known conflict of interest to disclose. The results of this study have been disseminated on the University of Western Ontario dissertation database. Funding sources included an NSERC Discovery Grant to Jessica A. Grahn (2016-05834), an Understanding Human Cognition Scholar Award from the McDonnell Foundation to Jessica A. Grahn, an NSERC Discovery Grant to Daniel Ansari (RGPIN: 342192), and a Klaus J. Jacobs Foundation Fellowship to Daniel Ansari.

Correspondence concerning this article should be addressed to Rebekka Lagacé-Cusiac, Department of Psychology, University of Western Ontario, Western Interdisciplinary Research Building (WIRB), Perth Drive, London, Ontario N6A 3K7, Canada. Email: rlagacec@uwo.ca

ratios present in a pair of serial intervals delimited by three tones (Nakajima, 1987; Nakajima et al., 1988).

Given the wide range of stimuli for which we perceive and use ratios, one may ask how these relative magnitudes are processed across temporal and spatial domains. Previous research indicates that ratios may be processed by a general ratio processing system (RPS; Lewis et al., 2015). However, this field of research has mostly focused on ratios that are symbolic (e.g., fractions) and visuospatial and nonsymbolic (e.g., ratios of lengths). Therefore, little is known about ratio processing mechanisms in other domains such as time. This leaves unanswered the question of whether ratios are processed by the same mechanism across magnitudes in space (e.g., numerosity, length, area) and time (e.g., duration).

Magnitude Processing Across Space and Time

It has been previously proposed that numerical, spatial, and temporal magnitudes may be processed by a generalized magnitude system. A theory of magnitude (ATOM) is one such theory, which suggests that numerical, spatial, and temporal magnitudes share a common analog system in the brain (Walsh, 2003). The theory also suggests that these different types of magnitudes are processed by common neural correlates, likely in the fronto-parietal network (Bueti & Walsh, 2009; Walsh, 2003). Since this theory was put forward, numerous groups have tested its predictions using both behavioral and neuroimaging paradigms. While some studies support the theory of a generalized magnitude system, others cast doubt on its validity (see Hamamouche & Cordes, 2019).

From a behavioral perspective, interference effects between numerical, spatial, and temporal magnitudes (i.e., the degree to which a judgment on a magnitude is influenced by the presence of another magnitude) are often inconsistent and asymmetric (Agrillo et al., 2010; Alards-Tomalin et al., 2016; Cai & Connell, 2015, 2016; Casasanto & Boroditsky, 2008; Fabbri et al., 2012; Gijssels et al., 2013; Ishihara et al., 2008; Srinivasan & Carey, 2010). For example, some studies show a greater influence of duration on judgments of spatial magnitude (Cai & Connell, 2015), while other studies show the opposite effect in which spatial magnitude only affects perceived duration (Casasanto & Boroditsky, 2008; Gijssels et al., 2013). Similarly, studies on cross-modal adaptation between space, time, and number either do not find any adaptation effect (Anobile et al., 2018) or find asymmetrical effects in which adaption in one dimension affects judgments in another dimension, but not vice versa (Tsouli et al., 2019). Finally, studies investigating individual differences related to temporal and spatial magnitude acuity have reported similar null or inconsistent results (Anobile et al., 2018; Mendez et al., 2011; Odic et al., 2016). For example, Mendez et al. (2011) compared performance on length and duration categorization tasks in humans and monkeys and found that length categorization was correlated with duration categorization, but only for certain pairs of lengths and durations (Mendez et al., 2011).

From a neuropsychological perspective, some populations, such as patients with spatial neglect, show processing deficits across spatial and temporal magnitudes (Basso et al., 1996), while other populations, such as patients with Parkinson's disease, show evidence for a double dissociation between spatial and temporal magnitude processing (Dormal, Grade, et al., 2012). Evidence from neuroimaging studies suggests that neural correlates associated with spatial and temporal magnitude processing partially overlap in the frontal and parietal cortices (Dormal, Dormal, et al., 2012; Hayashi et al., 2013; Skagerlund et al., 2016). However, these key brain areas have also been linked to other domain general processes, and it is therefore difficult to conclude that these areas constitute a generalized system specific to magnitude processing (Van Opstal & Verguts, 2013).

Given the inconsistent behavioral and neuropsychological evidence, some have offered a more nuanced explanation for a generalized magnitude system in which different types of magnitude are initially processed by separate mechanisms but share a common magnitude comparison system in later processing steps (Cona et al., 2021; Van Opstal & Verguts, 2013). Others reject the idea of a common magnitude system altogether and suggest that associations between spatial and temporal magnitude processing might be better explained by domain-general processes such as working memory or decision-making (Marcos & Genovesio, 2017; Van Opstal & Verguts, 2013).

A Generalized Ratio Processing System

Although it remains unclear whether spatial and temporal magnitudes are processed or represented in a general magnitude system, most of the research on these systems has studied perception and representations of absolute magnitudes, leaving open the question of how relative magnitudes (i.e., ratios) are processed across spatial and temporal domains. From a physical perspective, absolute magnitudes are concrete and orthogonal across different types of magnitudes. For example, it is impossible to convert a length to a duration without assigning an arbitrary conversion rule because these magnitudes exist in different dimensions (e.g., 1 cm is not equal to 1 s, nor 2, nor 3, etc.). In contrast, ratios are abstract, second-order magnitudes, and therefore their value stays constant regardless of the type of magnitude from which they have been derived (Bonn & Cantlon, 2017; Meng et al., 2019; Park et al., 2021). For example, a ratio of a half (1/2) does not change whether it is depicted by a bisected line, a set of dots in which half are one color and the other half is a different color, or an auditory interval that is half the duration of another interval.

Given ratios' abstract nature, one might predict that there exists a system that extracts and represents ratio magnitudes, regardless of their format (i.e., the type of magnitude used to depict them). Multiple psychophysical studies by Stevens have shown that humans can estimate relative magnitudes in several domains (e.g., Stevens & Galanter, 1957; Stevens, 1960). More recently, Lewis et al. (2015) have proposed an RPS they defined as "a set of neurocognitive architectures that support the representation and processing of nonsymbolic ratios" (Lewis et al., 2015, p. 144). This theory is derived from recent developments in the study of ratio processing and proportional reasoning. For example, Vallentin and Nieder (2010) showed that rhesus monkeys can discriminate nonsymbolic proportions in a spatial proportion-discrimination task. In this study, monkeys were shown a pair of lines representing a specific ratio followed by a second pair of lines representing either the same or a different ratio. The task was to indicate whether the ratio of the second stimulus matched the ratio of the first stimulus. The monkeys performed well above chance, and their performance resembled that of human subjects on all trained ratios as well as novel, untrained ratios, indicating that they had generalized the concept of proportionality (Vallentin & Nieder, 2010). Additionally, single-cell recordings collected during the task suggest the presence of ratio selective neurons in the prefrontal cortex (Vallentin & Nieder, 2010). The authors later replicated these findings and found similar ratio-tuned neurons in the posterior parietal cortex, a brain region often associated with magnitude processing (Vallentin & Nieder, 2010). Results from these studies provide strong support for an innate RPS shared with nonhuman primates. However, such a mechanism cannot necessarily be generalized to other types of magnitudes given that ratios were only depicted using line length.

Subsequent neuroimaging studies in humans have provided support for a common ratio processing system and extended them to the study of other magnitudes such as numerosity. For example, one study using a functional magnetic resonance imaging (fMRI) adaptation design found that humans encode relative magnitudes in the same areas known to encode absolute magnitudes (Jacob & Nieder, 2009b). In this study, the same ratio with varying overall sizes was repeatedly presented to participants, causing the signal in brain areas involved in ratio processing to decrease (a phenomenon often referred to as neural adaptation). Then, after multiple presentations of the same ratio, a comparison ratio was presented causing the signal in these areas to recover (i.e., increase). Participants showed this adaptation response to nonsymbolic ratios depicted using both length and numerosity (i.e., sets of dots and triangles). More importantly, the same adaptation pattern was found for both magnitude types (length and numerosity) in the same brain areas (Jacob & Nieder, 2009b). Another study using the same fMRI adaptation paradigm with number and word fractions (e.g., 3/6 and "a half") uncovered a similar pattern of activity, even when number and word fractions were mixed across trials (Jacob & Nieder, 2009a). These results converge with evidence from previous studies indicating that relative magnitudes are perhaps processed by a higher order mechanism that is invariant to format or type of magnitude. In other words, once absolute magnitudes are encoded, quantifying the relationship between magnitudes might be done by a single higher order mechanism whether they are symbolic (e.g., number and word fractions) or nonsymbolic (e.g., numerosity or length). However, the current body of literature has mainly focused on ratio processing in the spatial domain, and it is therefore unknown whether ratios in time might be processed by the same ratio processing mechanism.

Current Study

Previous research on magnitude processing across spatial and temporal domains has mostly been conducted on absolute magnitudes, and little is known about how relative magnitudes (i.e., ratios) are processed across domains. Furthermore, research on ratio processing has been limited to spatial magnitudes such as length and numerosity. Therefore, the aim of the current study was to bridge this gap by investigating how spatial and temporal ratio processing are related. More specifically, we investigated whether ratio processing is a domain-specific mechanism (i.e., ratios processed separately for each type of magnitude) or domain-general mechanism (i.e., ratios processed by a unique mechanism independent of magnitude type).

To investigate this question, adult human participants completed 12 tasks measuring ratio estimation (RE) and absolute magnitude discrimination, hereafter magnitude discrimination (MD), both in the spatial and temporal domain. These tasks included three spatial RE tasks (i.e., estimating the ratio between two lengths, areas, and numerosities), three temporal RE tasks (i.e., estimating the ratio between two durations), three spatial MD tasks (i.e., discriminating the longest/largest of two lines, areas, and numerosities), and three temporal MD tasks (i.e., discriminating the longest of two durations). If spatial and temporal ratios are processed by a unique, ratio-specific mechanism, then individuals' temporal and spatial RE ability were expected to correlate even after controlling for absolute magnitude processing. In other words, an individual who is more accurate at estimating spatial ratios, such as the relative length between two lines, would also be more accurate at estimating temporal ratios, such as the relative duration between two intervals, when controlling for their ability to perceive and process absolute lengths and durations. By controlling for absolute magnitude processing, we eliminated the possibility that the relationship between ratio processing in space and time is explained by the precision with which people perceive absolute spatial and temporal magnitudes.

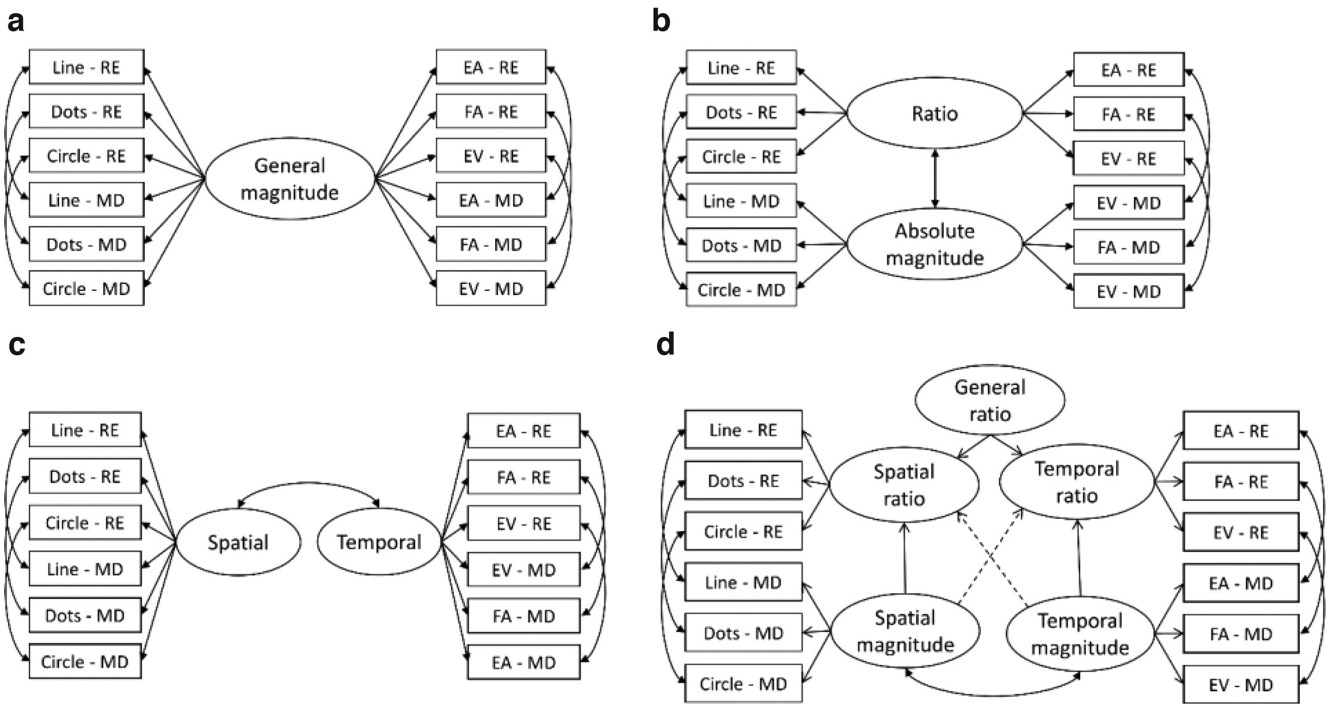
To test the hypothesis above, we developed and tested four competing structural equation models (SEMs): a single-factor confirmatory factor analysis (CFA) model, two two-factor CFA models, and a four-factor SEM model. These are shown in Figure 1 and described in more detail below. In deriving a set of competing models, we originally proposed to include a bifactor model separating a general factor (domain-general ratio processing) from the two specific factors (spatial ratio processing and temporal ratio processing). However, we became aware during the analysis stage, and upon further investigation of the literature on model identification, that a bifactor model with two specific factors and three indicators each is unidentified, meaning that it cannot provide a unique solution. Therefore, this bifactor model has therefore been omitted from the current report despite its inclusion in the preregistered report, which can be found at https://osf.io/h6gts.

Analysis of SEM models yields two types of information: model fit (i.e., how well does the model fit the data) and parameter estimates (i.e., the magnitude of the relationships between variables) along with their standard errors, which can be used for statistical tests. Although model fit was examined as a first necessary step to assure the adequacy of each measurement model proposed, the relevant hypotheses were confirmed based on the magnitude and statistical significance of the parameter estimates of the retained model.

Single-Factor CFA Model

Since all tasks involve making judgments about quantity, a single-factor model tested whether performance on the 12 tasks could be explained by a single general magnitude processing factor (Figure 1a). This model was not expected to fit the data as well as the subsequent models because it assumes that all tasks can be explained by a single latent variable. Thus, it does not account for differentiation between spatial and temporal processing or absolute magnitude versus ratio processing. In this model, as well as all subsequent models, a correlation parameter was included among the residuals of analogous estimation and discrimination

Figure 1
Hypothesized Models



Note. EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination; RE = ratio estimation. (a) The single-factor model assumes that performance on all tasks can be explained by a general magnitude factor. (b) The two-factor model assumes that performance on the tasks can be explained by two correlated latent factors (absolute and ratio magnitude). (c) The two-factor model assumes that performance on the tasks can be explained by two correlated latent factors (spatial and temporal magnitudes; not preregistered). (d) The four-factor higher order model assumes that performance on the tasks can be explained by four factors. Single-headed arrows control for absolute magnitude processing ability in both ratio processing factors.

tasks to account for common variance due to similar stimuli. For example, the residuals of the line length RE task were allowed to correlate with the residuals of the line length MD task because the same shape was used.

Two-Factor CFA Models

Next, a two-factor CFA model tested whether performance on the tasks could be explained by two correlated factors: a general ratio processing factor and an absolute magnitude processing factor (Figure 1b). This model tested the possibility that ratio processing and absolute magnitude processing are related but separable constructs. Put differently, this model tested whether the data could be explained solely by the current theories of a generalized magnitude system (e.g., ATOM; Walsh, 2003) and RPS (e.g., Lewis et al., 2015). The two-factor model was expected to fit the data significantly better than the single-factor model. However, it was not expected to fit the data as well as the next model because it does not account for the hypothesized differentiation between performance on spatial versus temporal tasks. In addition, a second two-factor model (not preregistered) with spatial magnitude processing and temporal magnitude processing as latent variables was fit to test whether performance on the tasks could be explained only by stimulus type (spatial or temporal). This model assumes that magnitudes, whether they are absolute or relative, are processed similarly within domains and differently across the spatial and temporal domains (Figure 1c).

Four-Factor Higher Order SEM Model

The fourth model was an SEM rather than a CFA and was therefore modeled in two steps. The first step consisted of assessing the measurement model by fitting a four-factor CFA to the data. The four latent variables included in the model were spatial ratio processing, temporal ratio processing, spatial magnitude processing, and temporal magnitude processing.

The second step consisted of fitting the structural model using the same four latent variables as the four-factor CFA model but now specifying specific relationships between these variables. First, the two ratio latent variables were regressed onto the latent absolute magnitude variables (represented with the single-headed arrows) because we assumed that participants would process the absolute magnitudes of the stimuli before extracting the ratio, and therefore absolute magnitude ability would explain ratio processing ability. Adding these paths allowed us to test the strength of the relationship between spatial and temporal ratio processing when controlling for spatial and temporal magnitude processing. In addition to controlling for *within-domain* absolute magnitude processing (e.g., spatial magnitude on spatial ratio), we also added dashed paths controlling for *between-domain* absolute magnitude processing (e.g., spatial magnitude to temporal ratio). Given that the literature is divided on how different types of magnitudes are processed, we included these paths as they might control for additional variance related to absolute magnitude processing ability and general cognitive ability. Since we did not expect the coefficients for the dashed paths to be significant, we estimated and compared two models (one with the dashed paths and one without the dashed paths) and retained the model with the best fit.

Last, a second-order factor (i.e., general ratio processing) influencing spatial and temporal ratio processing variables was included to explain the common variance between spatial and temporal ratio processing (Figure 1d). If ratio processing is a domain-general mechanism, large coefficients (i.e., loadings) are expected between the general ratio processing factor and the two ratio factors (spatial and temporal ratio processing). Equality constraints were placed on the two higher-order loadings for the model to be locally identified.

Method

Participants

Three hundred twenty-seven participants were recruited from the online survey panel Prolific. Thirty-nine participants withdrew before the start of the study due to technical difficulties, and 13 participants withdrew partway through the study either due to technical difficulties or by choice. The final sample consisted of 275 participants between 18 and 50 years old (M = 27.68, SD = 8.33; 106 women, 166 men, three nonbinary). Participants were residents from the United Kingdom (35.7%), Portugal (32.5%), United States (14.8%), Spain (5.8%), and South Africa (4.0%), as well as Ireland, Belgium, Canada, France, Germany, and Sweden (remaining 7.2%). To be eligible, participants had to be a minimum of 18 years old and self-report normal hearing and normal or corrected to normal vision. Participants also required access to a laptop or desktop computer with a keyboard and sound. Sampling on Prolific was also restricted to adults who were fluent in English to limit cases in which the participants did not understand the instructions well enough to execute the tasks. Sampling was also restricted to adults between the ages of 18 and 50 to limit the potential developmental confounds associated with an older population. Data were collected from April 24 to May 13, 2021. Participants were paid £7.50 for their participation. The study was approved by the nonmedical research ethics board at the University of Western Ontario.

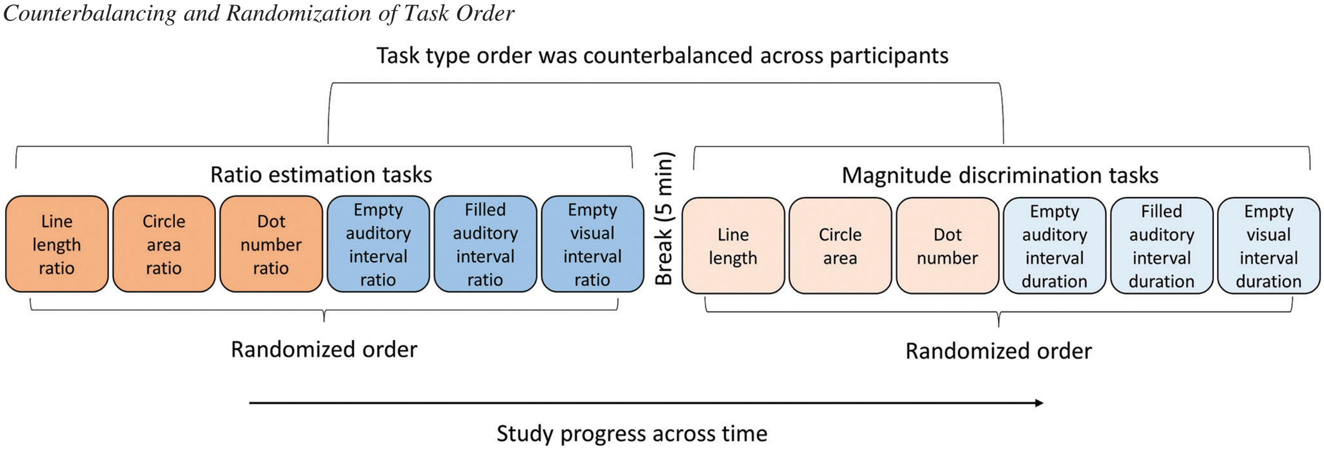
Study Design and Materials

Participants completed six RE tasks and six MD tasks. The study design is depicted in Figure 2. Tasks were grouped by task type (e.g., they completed all RE tasks and then all MD tasks), and the task type order was counterbalanced across participants. The order of tasks within each task type was randomized for each participant. Participants were permitted to take a 5-min break between the two sections. Once participants had completed all 12 tasks, they completed a short demographics questionnaire. The entire study took approximately 1 hour to complete. The study was programmed using the free software PsychoPy Version 2020.2.10 (Peirce et al., 2019) and hosted on the platform Pavlovia. The auditory stimuli for the various auditory tasks were generated using MATLAB (Version 2019a).

RE Tasks

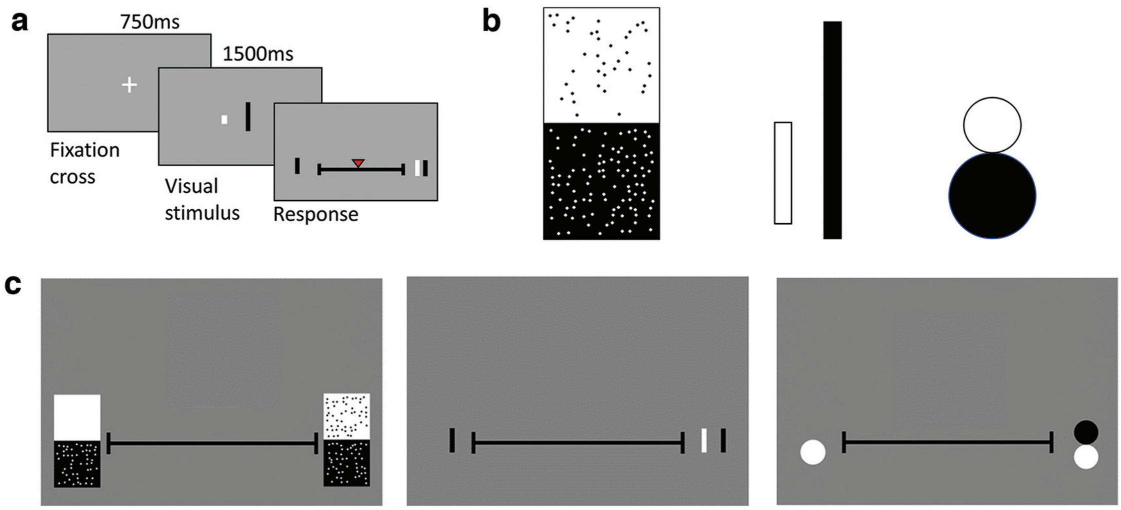
The RE tasks were a variation on the number line task commonly used in numerical cognition research (Siegler & Opfer, 2003). For all RE tasks, participants were presented a spatial or temporal ratio and then asked to represent that ratio on a bounded line (Figure 3a). In each trial, participants could click anywhere on the line and subsequently adjust their estimation if needed. Participants then pressed on the space bar to continue to the next trial. At the start of each RE task, participants were instructed to try to use the entire response line throughout the trials. There were three spatial RE tasks (i.e., pairs of dot arrays, line lengths, and circle areas) and three temporal RE tasks (i.e., auditory and visual durations with "empty" time intervals and auditory duration with "filled" intervals). Thus, all spatial RE tasks were visual tasks, and two temporal RE tasks were auditory and one was visual.

Figure 2
Counterbalancing and Randomization of Task Order



Note. See the online article for the color version of this figure.

Figure 3
Illustration of a Spatial RE Trial and Stimuli



Note. (a) Spatial ratio estimation (RE) trial. (b) Example stimuli for the dot array, line length, and circle area tasks, respectively. (c) Response screen for dot array, line length, and circle area. Stimuli on the left and right of the response line corresponded to ratios of 0:1 and 1:1, respectively. See the online article for the color version of this figure.

For spatial RE tasks, participants were presented with a pair of stimuli: One of the stimuli represented the part, while the other represented the whole (Figure 3b). The participants' task was to represent the part:whole ratio on a bounded line (adapted from Meert et al., 2012; Möhring et al., 2016). For example, if the stimulus corresponding to the part was half the size of the stimulus corresponding to the whole, then the participant would respond by marking the middle of the line. For each trial, the visual stimuli were presented for 1,500 ms and then replaced by the response line. At each end of the line was a figure showing a ratio of 0:1 on the left and 1:1 on the right (Figure 3c).

For temporal RE tasks, participants were presented a divided interval. Divided intervals were denoted either by three empty or filled tones or three brief flashes (see Figure 4). Participants' task was to represent the ratio of the divided interval using a bounded line (adapted from Nakajima, 1987). Participants were instructed to estimate when the second tone/flash occurred in relation to the first and third tones. For example, if the second tone/flash was presented halfway between the first and third tones/flashes, then the participant would respond by marking the middle of the line.

Each part-whole pair was created from 11 possible ratios (1/12 to 11/12). Each whole stimulus in the part-whole pair had three overall magnitudes that were randomized throughout the task. This resulted in a total of 33 trials (3 total magnitudes × 11 ratios) per task. Note that for temporal tasks, the whole corresponded to the total duration of the divided interval and the part corresponded to the interval between the first and second tones/flashes. Spatial stimuli were adapted from (Matthews et al., 2016; Park & Matthews, 2020; Park et al., 2021). Magnitudes used for the various spatial and temporal stimuli can be found in Table 1.

For tasks with empty auditory and visual intervals, durations were measured from the offset of the tone/flash to the onset of the subsequent tone/flash. For visual stimuli, flash duration was two frames with a refresh rate of 60 Hz (~32 ms). For auditory stimuli, the tone duration for empty intervals was matched to the flash duration (~32 ms). For tasks with filled intervals, the duration of each tone was equal to the length of the specified duration followed by a silence of 16 ms (to demarcate the onset of the next tone). The third tone in the filled interval stimuli was 200 ms across all ratios and total durations. Tones of 500 Hz were used in both the empty and filled tasks and had 10-ms linear onset/offset ramps.

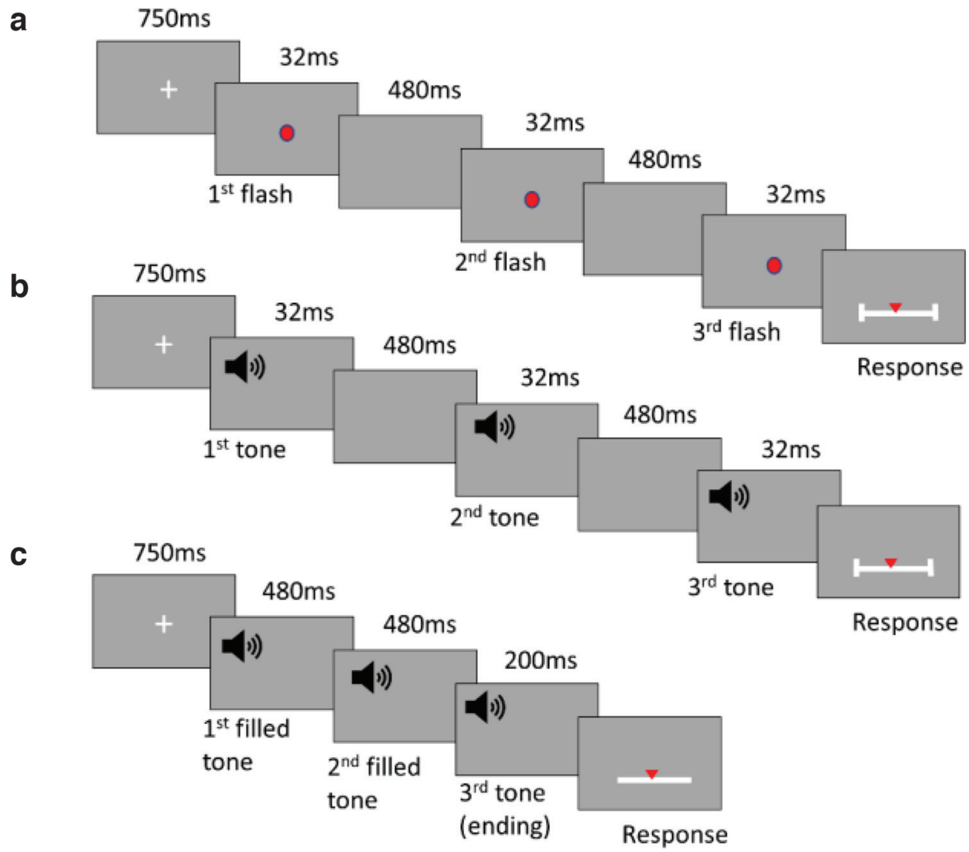
MD Tasks

To account for spatial and temporal absolute magnitude processing ability, participants completed six MD tasks, each created to be analogous to the six RE tasks. For all MD tasks, participants had to indicate which one of two stimuli was the largest/longest. They were instructed to press the "F" key if the first/stimulus on the left was larger/longer or the "J" key if the second/stimulus on the right was larger/longer. Their response immediately triggered the start of the next trial. Participants were instructed to respond as quickly as possible.

For spatial MD tasks, participants were presented a pair of visual stimuli and asked to indicate which of the two was the largest (i.e., circle area), was the longest (i.e., line length), or had the greatest quantity (i.e., number of dots; see Figure 5). The pair of stimuli were presented simultaneously for 1,000 ms. For temporal MD tasks, two intervals were presented serially, separated by ~2,400 ± 150 ms, and participants indicated which of the two intervals was longer (see Figure 6).

Stimuli for the discrimination task were created using eight standard magnitudes and five comparison ratios. As a result, each task was composed of 40 trials (5 comparison ratios × 8 standards). Each trial consisted of comparing a standard to a comparison magnitude, which was obtained by multiplying the standard magnitude with one of the comparison ratios. For example, given

Figure 4
Illustration of Temporal RE Trial for Three Types of Temporal Intervals



Note. RE = ratio estimation. Temporal RE trials for (a) empty visual intervals, (b) empty auditory intervals, and (c) filled auditory intervals. Examples are for a ratio of .5 and a total duration of 960 ms. A blank screen lasting 750 ms immediately preceded and followed the first and last flash/tone, respectively (not depicted in figure). See the online article for the color version of this figure.

a standard of 1 s and the five comparison ratios 1:1.20, 1:1.25, 1:1.30, 1:1.40, and 1:1.60, participants were presented the standard-comparison duration pairs of 1 s and 1.20 s, 1 s and 1.25 s, etc.

Comparison ratios, which were determined based on previous piloting, varied across tasks (e.g., they were different for the line length and the circle area discrimination tasks) but remained constant across all participants. Standard and comparison magnitudes spanned the range of the magnitudes presented in the RE tasks. The range of standard magnitudes for each task is listed in Table 1. The side on which the correct response was presented (or order in the case of temporal stimuli) was counterbalanced so that an equal number of larger/longer trials was presented on both sides (or in both orders in the case of temporal tasks). The side/order of presentation of the stimulus pairs was also counterbalanced across participants.

Practice Trials

Participants completed three practice trials for each RE task. In these practice trials, participants were shown a stimulus pair and asked to estimate the ratio for that pair using the response line. After they responded, a green line appeared on the response line indicating the correct answer. The same three ratios were given for every practice trial set (i.e., .25, .5, .75). Practice trials were done on the same total magnitude levels across all participants.

Participants also completed three practice trials at the beginning of each MD task. In these practice trials, participants were shown a pair of stimuli (i.e., the standard and a comparison) and indicated which was the largest/longest. After they responded, feedback was given indicating correctness (i.e., "correct" or "incorrect").

Attention Checks

Given that the study was conducted online, each task included one attention trial to verify that participants were not simply clicking through instead of paying attention to the task. For all attention trials, participants saw a screen after the stimulus presentation displaying "Attention check!" which lasted 1 s. For RE tasks, participants were then instructed to place their cursor either to the extreme left or right of the response line. The attention trial stimulus ratio for ratio tasks was always .5 so that the attention check response would not be confounded by actual estimations. The side of the correct response (i.e., left or right) was decided randomly for each trial.

Table 1
Task Parameters for Ratio Estimation and Magnitude Discrimination Tasks

	Magnitude discrimination		Ratio estimation	
Tasks	Comparison ratios	Range of standards	Stimulus ratios	Total magnitudes (denominator)
Dot number	1.09 (12:11), 1.10 (11:10), 1.12 (9:8), 1.14 (8:7), 1.25 (5:4)	48–133 dots	1/12, 2/12, 3/12, 4/12, 5/12, 6/12, 7/12, 8/12, 9/12, 10/12, 11/12	75, 100, 125 dots
Line length	1.01, 1.02, 1.03, 1.06, 1.12	75, 100, 125 pixels		75, 100, 125 pixels
Circle area	1.02, 1.04, 1.06, 1.08, 1.18	50–100 pixels		50, 75, 100 pixels
Empty auditory intervals	1.20, 1.25, 1.3 1.4, 1.6	200, 300, 400, 500, 600, 700, 800, 900 ms		480, 960, 1,440 ms
Filled auditory intervals	1.20, 1.25, 1.3 1.4, 1.6	200, 300, 400, 500, 600, 700, 800, 900 ms		480, 960, 1,440 ms
Empty visual intervals	1.20, 1.25, 1.3, 1.4, 1.6	400, 450, 500, 550, 600, 700, 800, 900 ms		960, 1,200, 1,440 ms

Note. Circle area values in table are given using radius length. However, all circle stimuli pairs were calculated based on circle area.

For all MD task attention checks, participants were instructed to press either the "F" or "J" key, regardless of the stimulus presented for that trial. The attention trial stimulus for the MD tasks was drawn from the easiest ratio bin, for which the difference between the stimulus pair was the largest and easiest to identify. The specific key participants were instructed to respond with (i.e., "F" or "J") corresponded to the incorrect answer for the stimulus pair presented.

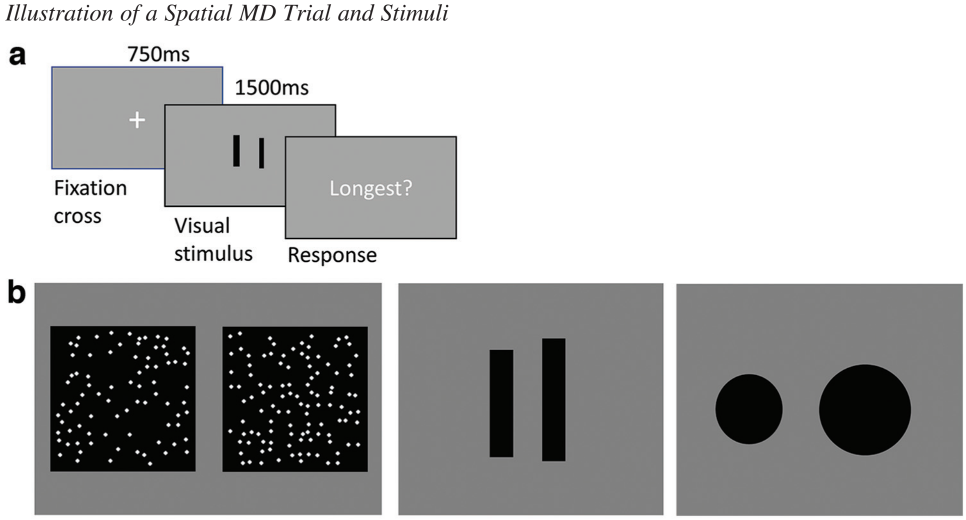
Demographics

Once participants completed all RE and discrimination tasks, they completed a demographics questionnaire. Information such as gender, age, years of education, hearing, and music experience (e.g., years of formal music training and years of music practice) were collected. Participants were also asked whether they understood how to perform the tasks, how difficult they perceived the tasks to be using a 3-point Likert scale (*easy*, *neutral*, or *difficult*), and whether they experienced any technical difficulties with either the auditory or visual stimuli during the experiment. This information was collected to support decisions regarding data exclusion during data preprocessing.

A Priori Power Analyses

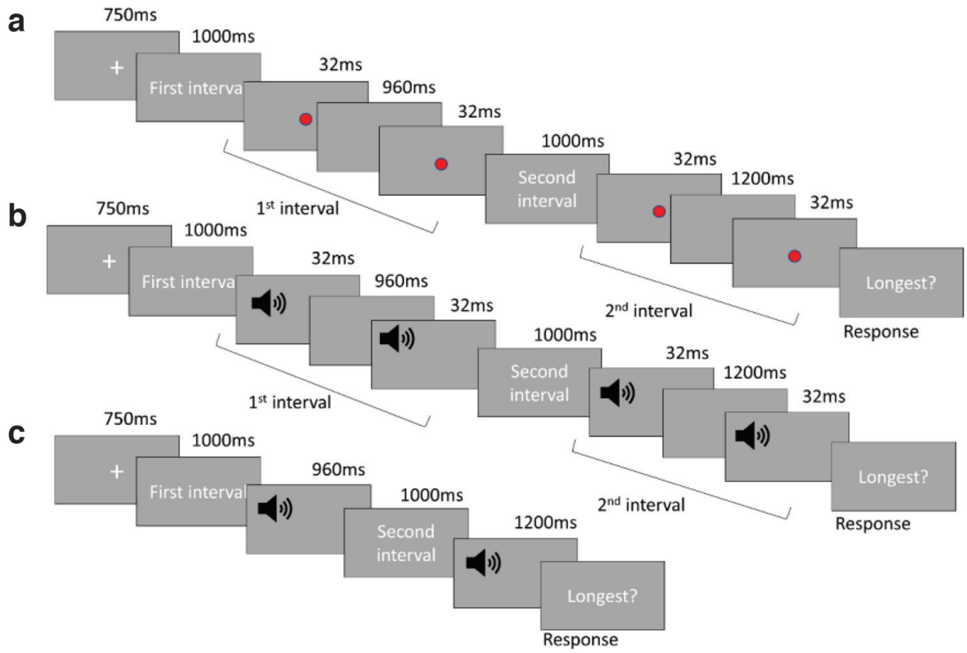
Using Mplus, Monte Carlo simulations were conducted for all proposed models. Sample size was decided based on the results of the power analysis for a four-factor model with correlated residuals as that was the main model of interest. This model is equivalent to the higher order model presented in Figure 1d but has correlated residuals between the two ratio factors instead of a higher order factor, which facilitates specification and interpretation of the coefficients and effect sizes. As there was no previous literature on the relationship between spatial and temporal ratio processing, we set the value to the smallest effect size of interest. Results from the simulations showed that a sample size of 275 was appropriate to detect a minimum correlation of .25 between the spatial and temporal ratio latent factors, controlling for the magnitude

Figure 5
Illustration of a Spatial MD Trial and Stimuli



Note. MD = magnitude discrimination. (a) Example of an MD trial. (b) Example stimuli for the dot array, line length, and circle area MD tasks. See the online article for the color version of this figure.

Figure 6
Illustration of Temporal MD Trial for Three Types of Temporal Intervals



Note. MD = magnitude discrimination. Temporal MD trials for (a) empty visual, (b) empty auditory, and (c) filled auditory intervals. Examples are for a comparison ratio of 1.25 and a total duration of 960 ms. A blank screen lasting 700 to 750 ms immediately preceded and followed each interval (not depicted in figure). Finally, a jitter lasting up to 150 ms was added in between the presentation of the first and second intervals. See the online article for the color version of this figure.

latent variables, with a power of .8 at the standard .05 alpha error probability. All 1,000 samples generated in the simulation reached convergence. Other relationships, such as spatial ratio-magnitude processing and spatial-temporal magnitude processing, were estimated based on previous literature (see online supplemental materials). Power analyses for this model can be found at https://osf.io/374zu/.

Preprocessing

Trials with long reaction times (greater than 30 s for RE tasks and 10 s for MD tasks) were excluded, as well as trials with significantly large errors in the RE tasks (for details, see online supplemental materials). This resulted in the exclusion of one trial in 20% of responses (i.e., tasks) across all tasks and participants, two trials in 2.33% of responses, and three to five trials in .88% of responses. Additional preprocessing steps were implemented to identify and correct RE responses with incorrect estimation patterns (i.e., scale inversion or half scale patterns). Out of all RE tasks across all participants, .8% of responses were corrected for scale inversion, and 3.15% of responses were corrected for half scale patterns (i.e., participants only used first half or last half of scale).

Aggregate scores were then calculated. For RE tasks, the absolute error (stimulus ratio − estimated ratio) was calculated for each RE trial. Absolute error was then averaged across trials for each task separately to obtain the average absolute error (average |error|). For discrimination tasks, the proportion of correct trials was calculated separately for each task. This score was then reversed prior to conducting the main analyses by subtracting them from 1 so that both aggregate scores would have the same direction (i.e., lower scores indicate better performance).

Once the aggregate scores for both RE and MD tasks were calculated, we identified responses with low overall accuracy (i.e., response patterns with slopes less than .3 for RE tasks and proportion correct less than .55 for MD tasks) and excluded these if participants indicated they did not understand the task, had technical difficulties (e.g., did not hear all tones), or failed the attention check. This resulted in the exclusion of 20 responses (i.e., tasks) across 13 participants. Additionally, four participants were entirely excluded from the analysis because they did not complete most of the tasks properly, did not understand how to complete most of the tasks, and/or showed signs of noncompliance (i.e., failed attention checks and long RTs on most tasks). This resulted in a final sample of 271 participants. Of these 271 participants, 258 participants had complete data sets (i.e., an aggregate score for each task).

Additional details about the preprocessing steps described below as well as a summary of trial and response exclusions are available at the following link: https://osf.io/uwd2t/. Visualizations of responses of the retained sample are available at the following link: https://osf.io/7t8sr/.

Main Analyses

Model Estimation

All analyses were conducted using the software R (Version 4.1.1) and the lavaan R package Version .6.8 (Rosseel, 2012). Models were estimated using a robust maximum likelihood estimator. This method provides robust standard errors (Huber-White) and scaled fit statistics for data and is appropriate for data with slight deviations from multivariate normality (Savalei, 2014; Savalei & Falk, 2014; Yuan et al., 2015). Once models were estimated, we verified that solutions were admissible and empirically identified (e.g., all standardized correlations were below 1, no negative variances). Missing data were managed by using full information maximum likelihood (FIML), which assumes that data are missing completely at random (MCAR) or at random when controlling for auxiliary variables (MAR). Missing data were due to a few cases of technical issues, noncompliance, or misunderstanding of the tasks. We have no reason to believe these missing data do not satisfy the MCAR/MAR assumption.

Model Evaluation

Each model was fit and assessed individually using global and local fit indices. Four global fit indices were considered: chi-squared test, comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean residual (SRMR). A significant chi-squared test ($p < .05$) indicates that the model significantly differs from the data and therefore fits the data poorly. Given that this statistical test is sensitive to sample size, large samples can result in rejecting the model for small discrepancies between the model and data. CFI values greater than .95, RMSEA values lower than .06 (with .05 indicating close fit), and SRMR values lower than .08 were used as thresholds indicating a model with reasonable fit (Hu & Bentler, 1999; Kline, 2015). Local fit was analyzed by looking at the residual correlation matrix (i.e., the difference between model correlation matrix and the data correlation matrix). As a rule of thumb, absolute residual correlations greater than .10 may indicate poor local fit (Kline, 2015). Differences in fit between the competing nested models were tested using the Satorra-Bentler scaled chi-square difference test.

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All study materials, data, and analysis code have been made publicly available on the Open Science Framework and can be accessed at https://osf.io/jbqta/. Data were analyzed using R Version 4.1.1 (R Core Team, 2021), the lavaan package Version 0.6.8 (Rosseel, 2012), and the influence.-SEM package Version 2.2 (Pastore & Altoe, 2018).

This study's design, hypotheses, and analysis plan were preregistered (see https://osf.io/h6gts). There are three major ways in which the current analysis diverges from preregistered report. First, the original report included a bifactor model that was then omitted because it was found to be unidentified. Second, univariate outliers were initially to be treated as missing data and multivariate outliers excluded altogether, in part to handle problems related to multivariate nonnormality. Aguinis et al. (2013) provided alternative ways of handling outliers instead of completely excluding them from the analysis (which could bias results). Therefore, the analyses and results reported here follow the best practice recommendations listed in (Aguinis et al., 2013). Third, a second two-factor CFA model was added to the list of hypothesized models to test whether the data could be represented by a spatial and temporal factor. In the aim of transparency, results from the original analyses plan are also available at this link: https://osf.io/r5u3n/.

Results

Descriptive Statistics

Descriptive statistics for each task are listed in Table 2. Before fitting the various models, we inspected the data for evidence of multivariate nonnormality. Because the assumption of multivariate normality cannot be directly tested, univariate and bivariate nonnormality were taken as indirect indicators of multivariate nonnormality (Kline, 2015). Tasks deviating substantially from univariate normality were identified as tasks with a skew greater than $\pm$ 2 and kurtosis greater than 4. Additionally, we visually inspected the bivariate scatterplots and quantile-quantile plots for all task pairs for evidence of bivariate normality, linearity, and homoscedasticity of the residuals (Kline, 2015). From these inspections, we found evidence of deviation from univariate and bivariate normality, indicating that the assumption of multivariate normality was likely violated. To address the violation of this assumption, a robust maximum likelihood estimator was used to fit the hypothesized models. Robust maximum likelihood corrects standard errors and model fit statistics in the case of deviation from multivariate normality even in the presence of missing data (Savalei, 2014). In addition to verifying the assumption of multivariate normality, the data were screened for extreme bivariate and multivariate collinearity. Table 3 displays the bivariate correlation matrix for all 12 tasks. All tasks had low to moderate correlation coefficients (range = .14–.63), and there was no evidence of extreme bivariate collinearity (all correlations were below .85; Brown, 2006).

Main Analyses

Table 4 summarizes goodness-of-fit statistics for each model estimated. All models were shown to be empirically identified. Fully standardized parameter estimates are reported in the path diagrams and can therefore be interpreted as correlations in the case of double-headed arrows and standardized regression coefficients in the case of single-headed arrows. Complete standardized and unstandardized solutions for all models can be found in the online supplemental materials.

Single-Factor CFA Model

We first tested the theory that all tasks are explained by a single general magnitude processing factor (see Figure 7). This model yielded a poor fit according to the chi-squared statistic, CFI, and RMSEA (Hu & Bentler, 1999; Kline, 2015). The SRMR was at the limit of what is considered reasonable fit (Hancock & Mueller, 2008). Finally, local fit testing showed four instances of poor local fit in which the residual correlation was greater than $\pm$ .10. Thus, the one-factor model could not adequately explain the participants' performance on the various tasks.

Two-Factor CFA Models

Next, we tested two two-factor models. The first two-factor model tested whether the data could be explained by the following two underlying factors: a general ratio processing factor and a general (absolute) magnitude processing factor (Figure 8a). Similar to the previous model, this two-factor model showed poor fit according to the chi-squared statistic. RMSE, CFI, and SRMR all

Table 2

Descriptive Statistics for Average Absolute Error (RE Tasks) and Proportion Incorrect (MD Tasks)

Tasks	N	M (SD)	Median	Range (min–max)	Skewness	Kurtosis	Cronbach's α
Spatial RE							
Circle-RE	270	.116 (.035)	.112	.052–.249	0.95	4.14	.83
Dot-RE	270	.125 (.037)	.118	.055–.265	0.93	4.05	.78
Line-RE	267	.098 (.040)	.089	.034–.314	1.41	6.56	.90
Temporal RE							
EA-RE	266	.090 (.037)	.080	.042–.281	2.00	8.31	.90
FA-RE	269	.115 (.052)	.103	.038–.294	1.20	3.93	.91
EV-RE	269	.094 (.047)	.081	.036–.344	2.32	9.75	.94
Spatial MD							
Circle-MD	271	.164 (.071)	.150	.025–.425	0.62	3.87	.41
Dot-MD	271	.158 (.071)	.150	0–.4	0.68	3.44	.48
Line-MD	271	.162 (.071)	.150	.025–.375	0.50	2.88	.43
Temporal MD							
EA-MD	269	.177 (.110)	.150	0–.575	1.06	4.29	.74
FA-MD	269	.138 (.081)	.125	0–.4	0.74	3.25	.64
EV-MD	270	.180 (.107)	.150	0–.575	1.02	3.81	.69

Note. EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination; RE = ratio estimation; min = minimum; max = maximum. The mean score for MD tasks refers to the proportion of incorrect responses. These means were transformed from the proportion of correct responses to make the direction of scores constant across the RE tasks and the magnitude discrimination tasks (lower scores indicate a better performance). The mean score for RE tasks refers to the averaged absolute error.

indicated adequate fit. In terms of local fit, there were three instances of poor local fit. Finally, when compared to the previous one-factor model, the nested chi-squared difference test indicated that the two-factor model fit the data significantly better than the one-factor model, $\chi^2(1) = 57.76$, $p < .001$. The second two-factor model tested whether the data could be modeled using two different factors: general spatial and general temporal processing (Figure 8b). In this case, the model assumes that performance on the set of tasks can be explained by domain-related factors. This model yielded slightly worse model fit than the previous two-factor model given that the RMSEA was now above the cutoff. However, it still fit the data significantly better than the one-factor model, $\chi^2(1) = 25.91$, $p < .001$.

Four-Factor CFA Model

The four-factor CFA model showed the best model fit according to all fit indices. The four-factor model also had a significantly better fit than both two-factor models, ratio and magnitude factors: $\chi^2(5) = 44.32$, $p < .001$; spatial and temporal factors: $\chi^2(5) = 64.73$, $p < .001$. All residual correlations were below ± .10, indicating good local fit. Figure 9 shows the parameter estimates for this model.

All factor loadings were between .39 and .85. Higher loadings within a given latent variable, ideally above .40 or even .50, indicate higher commonality among the indicator variables. Only one of the subtests, dot MD, had a relatively lower loading on the spatial magnitude latent variable (.39). Composite reliability scores, as measured by McDonald's omega, were greater than .7 for three

Table 3

Bivariate Correlations for All Tasks From the FIML Observed Covariance Matrix

Tasks	1	2	3	4	5	6	7	8	9	10	11
1. Circle-RE	1										
2. Line-RE	.41	1									
3. Dot-RE	.34	.57	1								
4. EA-RE	.32	.58	.39	1							
5. FA-RE	.25	.47	.36	.61	1						
6. EV-RE	.27	.49	.44	.63	.59	1					
7. Circle-MD	.25	.38	.26	.27	.23	.27	1				
8. Line-MD	.14	.29	.22	.27	.27	.27	.27	1			
9. Dot-MD	.24	.21	.24	.28	.16	.23	.23	.15	1		
10. EA-MD	.31	.44	.34	.58	.51	.45	.33	.35	.28	1	
11. FA-MD	.25	.39	.35	.43	.43	.48	.30	.27	.23	.59	1
12. EV-MD	.28	.39	.41	.49	.49	.56	.31	.30	.27	.55	.50

Note. FIML = full information maximum likelihood; EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination; RE = ratio estimation. All correlations were statistically significant at $p < .001$, except for the correlations between FA-RE and dot-MD ($p < .01$), line-MD and circle-RE ($p < .01$), and line-MD and dot-MD ($p < .01$). These exceptions are underlined in the table. Colors do not have a gradient scale but were included to help visualize correlation clusters.

Table 4
Goodness-of-Fit Statistics for All Models

Models	χ^2 (*df*)	rRMSEA [90% CI]	rCFI	rSRMR	AIC	BIC
1-factor CFA	125.77 (48), $p < .001$	0.083 [0.065, 0.101]	0.916	**0.046**	−10,097.60	−9,946.32
2-factor CFA (magnitude ratio)	79.93 (47), $p = .002$	**0.055 [0.033, 0.075]**	**0.964**	**0.038**	−10,147.67	−9,992.78
2-factor CFA (spatial temporal)	92.63 (47), $p < .001$	0.064 [0.044, 0.083]	**0.951**	**0.040**	−10,134.23	−9,979.34
4-factor CFA and SEM	**39.67 (42), $p = .574$**	**0.000 [0, 0.041]**	**1.000**	**0.027**	−10,183.42	−10,010.52
4-factor SEM (trimmed)	**40.41 (44), $p = .626$**	**0.000 [0, 0.038]**	**1.000**	**0.027**	−10,187.16	−10,021.47

Note. r = robust; RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root mean residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; CFA = confirmatory factor analysis; SEM = structural equation model. Bolded values are fit statistic values indicating good fit according to the criteria described in the Method section. The chi-squared statistic was corrected using the Mplus variant of the Yuan-Bentler correction factor. Chi-squared scaling factors were between 1.138 and 1.164.

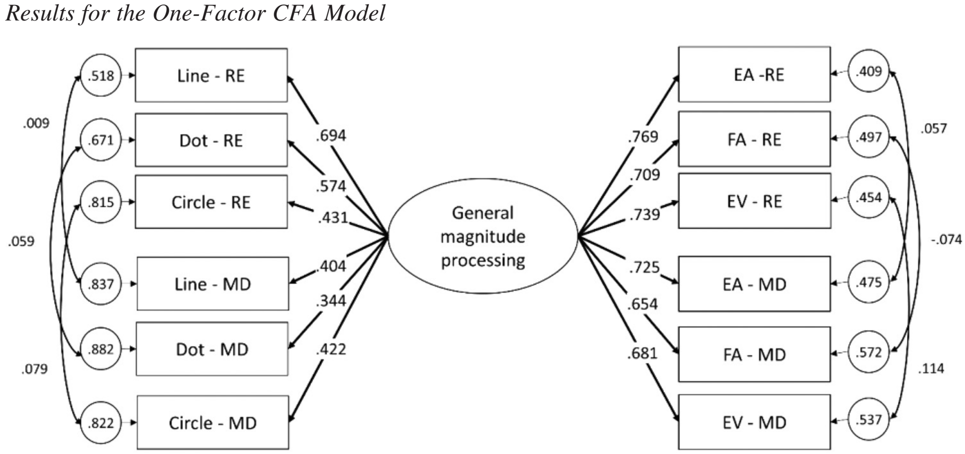
out of the four factors (.81, .79, and .73 for the temporal ratio, temporal magnitude, and spatial ratio factors, respectively), indicating that these indicators were reliable measures of the latent construct. In contrast, the spatial magnitude factor showed poor composite reliability (.46). Inspection of the loadings for that variable revealed smaller loadings (.39, .48, and .54) than in other latent variables, thus explaining the lower reliability. The model also revealed that two of the four correlations between the latent factors were greater than .80, which may indicate that the constructs are not empirically distinguishable. However, following the procedure proposed by Rönkkö and Cho (2022), we found that only the correlation between spatial and temporal magnitude suggested weak discriminant validity. Given the nature of the tasks (i.e., their clear conceptual/operational distinction), we chose to keep these two latent variables separate. In the end, the four-factor measurement model was retained for the structural analyses.

SEM Models With a Higher Order Latent Variable

The first model is a full SEM model consisting of the four-factor CFA model with regression paths going from the magnitude to the ratio latent variables instead of correlations. We first tested the significance of the paths going across domains and magnitude type (i.e., from temporal magnitude to spatial ratio processing and spatial magnitude to temporal ratio processing). To do this, we fit a trimmed model in which these paths were constrained to zero. According to fit indices, the resulting trimmed model (Figure 10a) had comparable fit to the previous SEM model containing all paths (not depicted). In other words, the fit of the trimmed model was not significantly worse than the model with all paths included, $\chi^2(2) = .306$, $p = .86$. In addition, all residual correlations for the trimmed model were below .10, indicating good local fit. Because the fit indices indicated that these two structural models were comparable, we retained the most parsimonious model (i.e., the trimmed model).

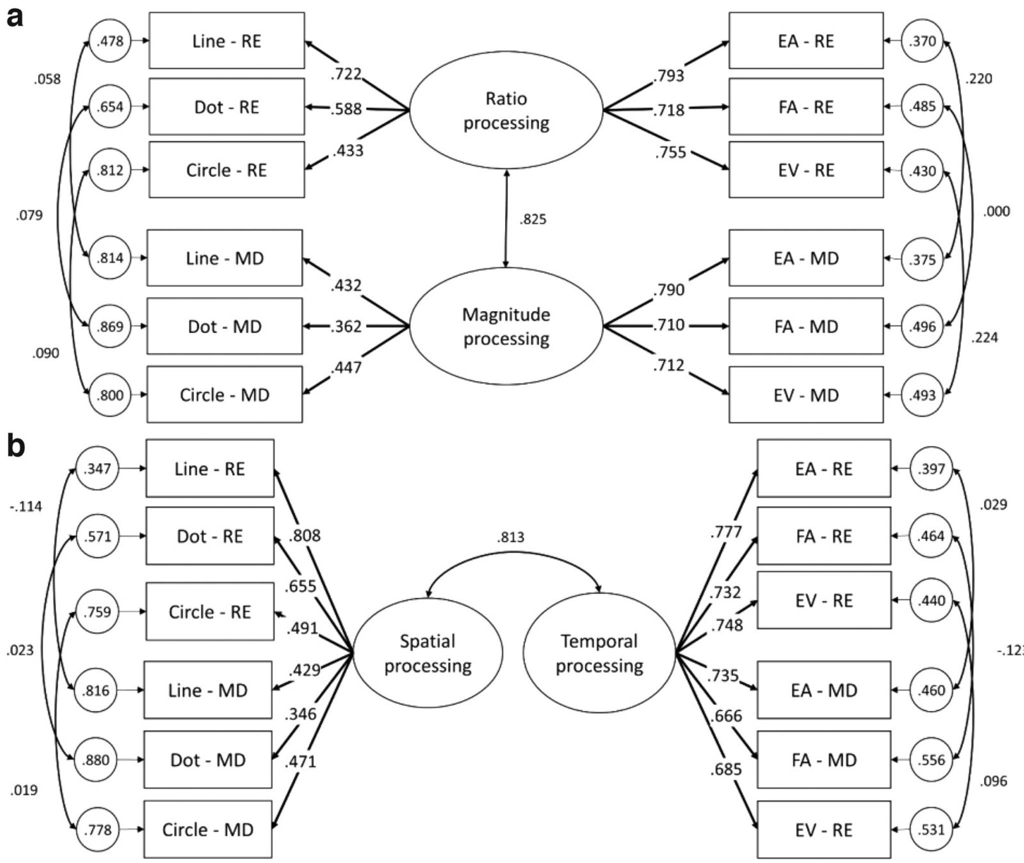
Figure 10a depicts the parameter estimates for the retained structure model. The residual correlation between spatial ratio and temporal ratio was high (.632, *SE* = .119, $p < .001$), indicating that there is a significant relationship between spatial and temporal ratio processing once we control for absolute magnitude processing. This is a slight decrease from the correlation between these two variables in the four-factor CFA model in which the relationship is not controlled for absolute magnitude processing (.774, *SE* = .046, $p < .001$). Spatial magnitude processing significantly predicted spatial ratio processing ($\beta = .777$, *SE* = .053, $p < .001$, 95% confidence interval [CI; .673, .880], 99% CI [.641, .913]). This means a 1-unit standard deviation change in spatial magnitude processing ability was related to a .777 standard deviation unit change in spatial ratio processing ability. Temporal magnitude significantly predicted temporal ratio processing ($\beta = .811$, *SE* = .044, $p < .001$, 95% CI [.725, .897], 99% CI [.698, .924]). This means a 1-unit standard deviation change in temporal magnitude

Figure 7
Results for the One-Factor CFA Model



Note. CFA = confirmatory factor analysis; EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination; RE = ratio estimation. Parameter estimates are fully standardized.

Figure 8
Results for Two Two-Factor CFA Models



Note. Two-factor CFA models with (a) ratio and absolute magnitude factors and (b) spatial and temporal factors. CFA = confirmatory factor analysis; EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination; RE = ratio estimation. Parameter estimates are fully standardized.

processing ability was related to a .811 standard deviation unit change in temporal ratio processing ability. Finally, spatial and temporal magnitude processing ability were significantly correlated (.855, $SE = .053$, $p < .001$).

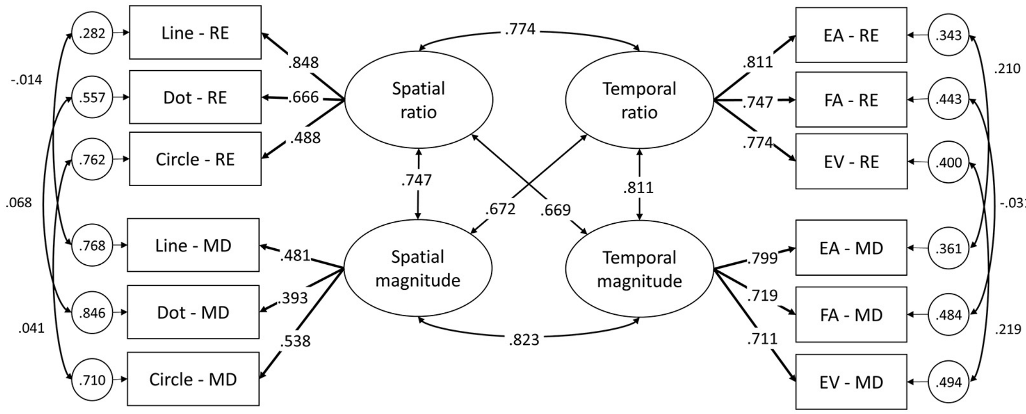
We also considered an equivalent model in which the correlation between the spatial and temporal ratio latent variables was specified as a higher order factor with equality constraints imposed on the loadings (see Figure 10b). Though this model is equivalent (i.e., it is another way of specifying the correlation between two latent variables and results in the same goodness fit), it assumes that the spatial and temporal ratio factors are affected by a common ratio processing factor. Spatial and temporal ratio processing were significantly related by the general ratio processing factor ($\beta = .501$, $SE = .059$, $p < .001$, 95% CI [.385, .617], 99% CI [.348, .653], and $\beta = .465$, $SE = .070$, $p < .001$, 95% CI [.329, .602], 99% CI [.286, .645], respectively). Overall, in this higher order SEM model, the respective magnitude and general ratio factor explained in sum 85% of the variance for the spatial ratio processing and 87% of the variance for temporal ratio processing. Finally, sensitivity analyses were conducted to determine whether the results could be due to influential outliers or the violation of the multivariate normality assumption. Though we did detect some influential outliers, exclusion of these outliers did not qualitatively change the model parameter estimates significantly (see online supplemental materials). This indicates that results were not due to influential outliers or the violation of the multivariate normality assumption.

Exploratory Analyses

Number Line Estimation Response Bias

One possible explanation for the correlation between the spatial and temporal ratio factors is that participants had a similar response bias unrelated to their ratio processing ability. For example, a person might have a bias away from the response line extremities (i.e., their estimations are biased toward the middle of the response line) thereby causing them to have a lower slope while still being highly precise (Figure 11b). This leads to a worse average absolute error score even though they may be just as precise as an individual with a slope near 1 (Figure 11a). If people have consistent biases, the use of the number line for all ratio tasks may inflate correlations between the two ratio processing factors. To test whether the results found in the previous section may be influenced by response bias, we refit the final model using R^2 as an

Figure 9
Results for Four-Factor CFA Model



Note. CFA = confirmatory factor analysis; EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination; RE = ratio estimation. Parameter estimates are fully standardized.

accuracy measure instead of average absolute error. R^2 was extracted from a linear model fit to each participant's response in every task following data cleaning. The advantage of using R^2 as a measure of RE accuracy is that it measures precision relative to the slope for each task and participant. It is also robust to problems created to incorrect response patterns (i.e., scale inversion and half scale patterns). The disadvantage is that this measure might not capture the accuracy of the individual's RE relative to the true stimulus ratio. Also, R^2, like average absolute error, does not differentiate response biases when the participant also has low precision (Figure 11c and d).

When using R^2 as a measure of RE accuracy, model fit remained adequate, $\chi^2(44) = 52.598$, $p = .175$, CFI = .991, RMSEA = .027, 90% CI [.00, .057], SRMR = .031, Akaike information criterion (AIC) = −5,521.09, Bayesian information criterion (BIC) = −5,355.39. Parameter estimates were similar to the ones previously reported, with two notable exceptions (for complete unstandardized and standardized solutions, see the online supplemental materials). First, the loading for the circle RE task went from $\beta = .488$ ($SE = .062$) to $\beta = .726$ ($SE = .057$). The spatial ratio factor now showed adequate composite reliability (.81). Second, the residual correlation between spatial and temporal ratio processing went from .632 ($SE = .119$) to .728 ($SE = .117$).

Taken together, these results indicate that initial findings are robust to the accuracy measure used and somewhat robust to linear response biases. It does not appear that response bias inflated the relationships between tasks. Furthermore, they suggest that R^2 might also be a more reliable measure of ratio processing than average absolute error as it is less vulnerable to certain types of response bias.

Reliability of Comparison Ratios

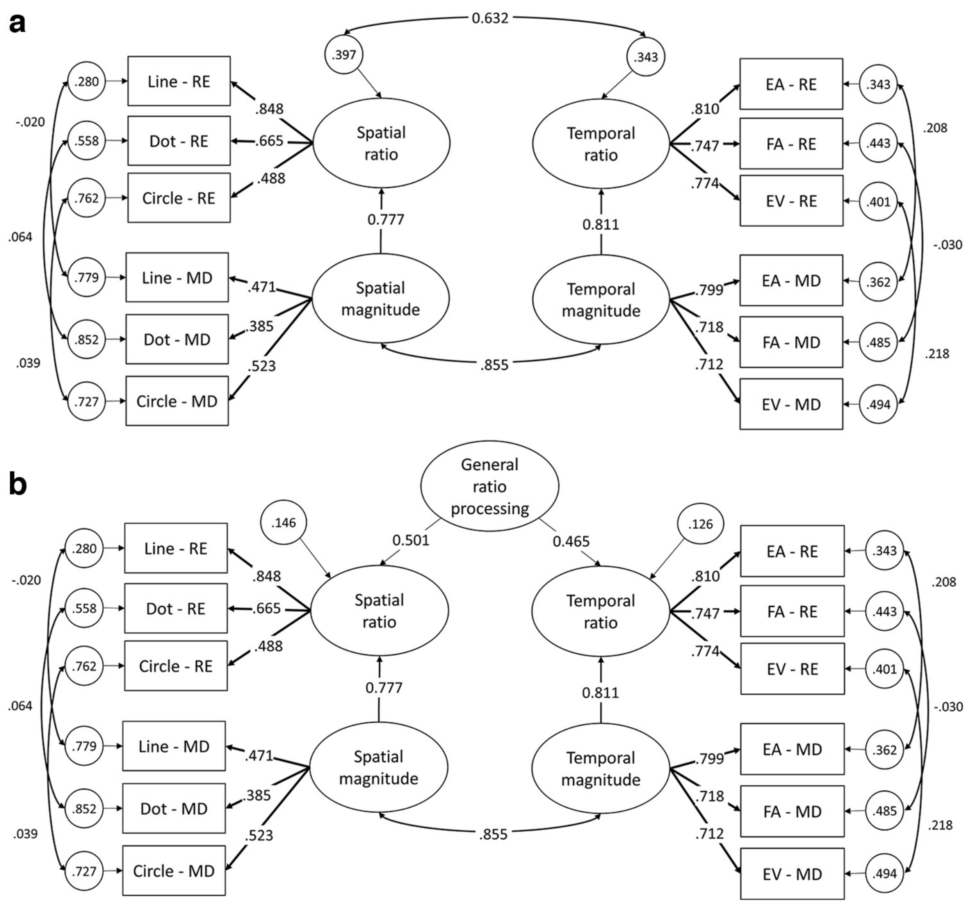
Indicators of the spatial magnitude factor were shown to have poor reliability. To investigate the source of this low reliability, we tested whether poor reliability in the spatial MD tasks could be explained by the comparison ratios used in this study. When designing the study, we chose sets of comparison ratios on which participants performed above chance but that were difficult enough to avoid ceiling effects. However, we had not verified that these comparison ratios had adequate reliability. Therefore, we used a CFA to examine the reliability of the comparison ratios. Descriptive statistics for the comparison ratios are shown in Table S9. On average, participants were above chance on all comparison ratios in each task ($p < .001$).

To investigate the reliability of the comparison ratios, we estimated a two-factor CFA model with spatial magnitude and temporal magnitude as latent variables. Each factor was composed of three latent subfactors (i.e., the tasks): line length, circle area, and dot number discrimination for the spatial magnitude factor and empty auditory, full auditory, and empty visual interval discrimination for the temporal magnitude factor. Each subfactor was composed of five indicators corresponding to the comparison ratios. Indicators corresponded to the proportion of correct trials across all standard magnitudes for a given comparison ratio. The model was estimated using FIML to handle missing data and robust maximum likelihood to handle the nonnormality of the factor indicators. The model yielded adequate fit, $\chi^2(398) = 399.828$, $p = .47$, CFI = .998, RMSEA = .004, 90% CI [.00, .022], SRMR = .047, AIC = −11,213.35, BIC = −10,863.95. The model along with fully standardized parameter estimates are shown in Figure 12 (see Table S10 for complete solution). From this figure, we noticed two potentially problematic comparison ratios. These two loadings, which were the hardest comparison ratios for both the line ($\beta = .110$, $SE = .072$, $p = .133$) and circle discrimination tasks ($\beta = .112$, $SE = .064$, $p = .148$), were not significantly different from zero. This indicates that these comparison ratios might not be appropriate or reliable measures of MD even though participants were, on average, above chance. In practice, these comparison ratios might have been too small for participants to discriminate the difference without guessing, thereby introducing noise into the measures. When recalculating the task reliability after these exclusions, Cronbach's alpha for both the line and circle tasks remained low (.44 and .47, respectively).

We then wanted to investigate whether these poor comparison ratios impacted the results found in the previous SEM analysis. We recalculated the aggregate scores excluding the comparison ratios with nonsignificant loadings (line 1.01 and circle 1.02) and

Figure 10

Retained Four-Factor SEM Model With a Higher Order Variable and Equivalent Model With Correlated Residuals



Note. (a) SEM model with a higher order latent variable (general ratio processing). (b) Equivalent four-factor SEM model with a correlation between residuals of spatial and temporal ratio variables. SEM = structural equation modeling; EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination; RE = ratio estimation. All parameters shown in the figure are significant ($p < .001$), except for the loading between circle ratio and spatial ratio processing (p = .001). Values in circles are residual variances.

reestimated the final model with correlated residuals. Model fit remained adequate, $\chi^2(44)$ = 38.779, p = .694, CFI = 1.000, RMSEA ≤ .001, 90% CI [.00, .030], SRMR = .026, AIC = −10,147.82, BIC = −9,982.12, and most parameter estimates showed little change, qualitatively speaking (see Table S11 for complete solution). The factor loadings for line (β = .471, *SE* = .067 to β = .528, *SE* = .063) and circle MD (β = .523, *SE* = .065 to β = .558, *SE* = .062) improved slightly, and the residual correlation between spatial and temporal ratio slightly decreased from .632 (*SE* = .119) to .581 (*SE* = .118).

Taken together, the lack of changes in parameter estimates suggests that the reliability of the comparison ratios did not have an important effect on the final model, and therefore our conclusions remain unchanged. However, removing the problematic ratios did not solve the issue of low loadings for the spatial discrimination tasks either. This indicates that other factors, such as an insufficient number of trials, may be the source of low reliability and that results from the model must still be interpreted with caution.

Effect of Education and Music Experience

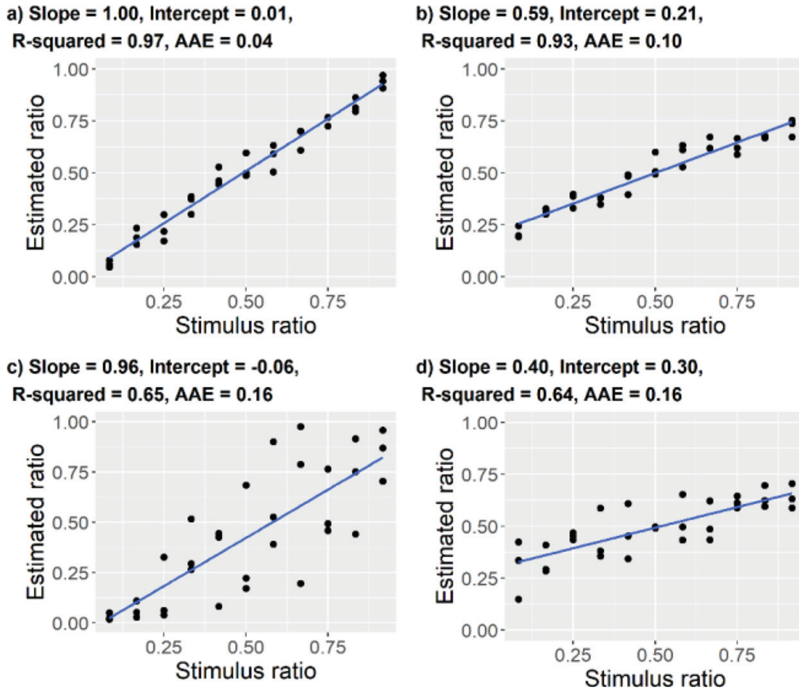
Finally, we tested whether performance was related to previous experience. For example, musicians have been shown to perform better on certain temporal tasks (Banai et al., 2012; Rammsayer & Altenmuller, 2006; Vibell et al., 2021). To examine the effect of prior experience, we analyzed the correlations between each task and years of education, years of music training, and years of music playing experience. None of the correlations were statistically significant after correcting for multiple comparisons using a Bonferroni correction (see online supplemental materials).

Discussion

The aim of the current study was to examine whether spatial and temporal ratios are processed by a common RPS. If ratios in space and time are processed by a common mechanism, then

Figure 11

Examples of Response Bias Influencing Average Absolute Error (AAE)



Note. (a) Perfect slope with high precision. (b) Low slope with high precision. (c) Near perfect slope with low precision. (d) Low slope with low precision. See the online article for the color version of this figure.

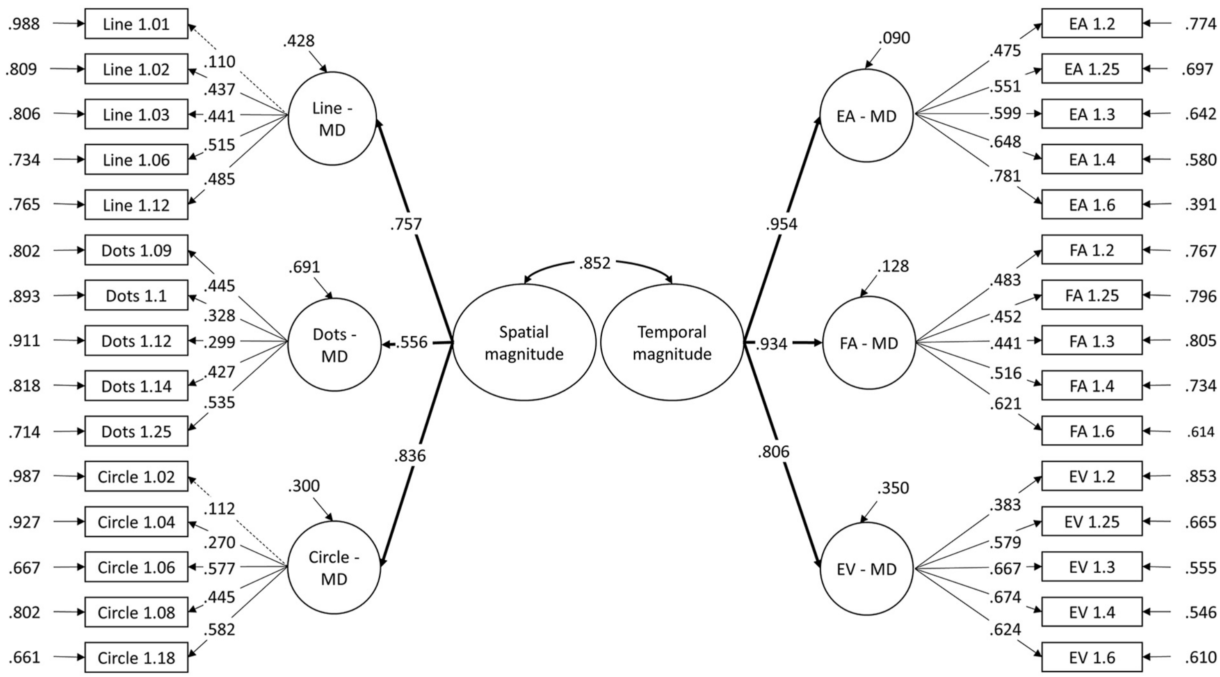
individuals' ability to process ratios in space (either in length, area, or numerosity) should correlate with their ability to process ratios in time. To test this hypothesis, we measured adult participants' ability to represent spatial and temporal ratios on a bounded line as well as discriminate temporal and spatial absolute magnitudes and modeled how these different abilities are related using SEM.

The single-factor model tested the hypothesis that all tasks could be explained by a single general factor. This model showed the worst fit out of the four models. Next, two two-factor models tested whether the data could be modeled using either a spatial and temporal factor or a magnitude and ratio factor. Although these two models fit the data better than the previous single-factor model, other model indices indicated that both two-factor models poorly fit the data. The final model was tested in the two steps. In the first step, the four-factor CFA model was fit to assess the measurement properties of the four factors. This model showed the best fit compared to the other CFA models according to all fit indices used. This indicates that the performance on the tasks can be modeled by four separable constructs: spatial magnitude, spatial ratio, temporal magnitude, and temporal ratio processing.

This measurement model showed that both temporal factors had adequate composite reliability indicating that the tasks were tapping into general timing ability, rather than modality-specific timing ability. This is consistent with previous literature showing that, although individuals tend to have a higher temporal resolution in the auditory modality than the visual modality, both use the same underlying timing mechanism when longer stimuli (near the 1-s range) are used (Rammsayer et al., 2015; Stauffer et al., 2012). Although the spatial ratio factor had adequate composite reliability, the spatial magnitude factor had low composite reliability, which could be explained by the individual tasks' low reliability, as opposed to the tasks sharing little common variance. Finally, all latent variables were shown to be empirically distinguishable, the only exception being the spatial magnitude and temporal magnitude latent variables, which showed moderate evidence of discriminant validity problems (Rönkkö & Cho, 2022). Although poor discriminant validity can indicate that two constructs are not empirically distinguishable, the nature of the tasks (spatial vs. temporal discrimination) was sufficiently different on a conceptual level to model them separately in the four-factor model.

In the second step, the four-factor CFA was respecified into an SEM including within- (e.g., spatial magnitude and spatial ratio) and between- (e.g., temporal magnitude and spatial ratio) domain regression paths between the ratio and magnitude latent variables. The between-domain paths (i.e., temporal magnitude and spatial ratio and spatial magnitude and temporal ratio) were initially included to control for leftover variance related to general cognitive ability. Conversely, we had little theoretical reason to believe that these paths would be significantly different from zero once we controlled for within-domain relationships between magnitude and ratio processing. Therefore, we fit this SEM model twice: once with the between-domain magnitude-ratio paths and once without. Although the latter model had two fewer paths than the first model, it was retained as it fit the data equally well to the former model and was the most parsimonious. This suggests that there is

Figure 12
Results for CFA Model on Comparison Ratios



Note. CFA = confirmatory factor analysis; EA = empty auditory interval; FA = filled auditory interval; EV = empty visual interval; MD = magnitude discrimination. Parameter estimates are fully standardized. Dashed lines indicate that the parameter estimate was not significantly different from zero ($p > .05$).

no evidence for additional mechanisms that are involved in linking magnitude and ratio factors across domains other than the common mechanisms related to absolute magnitude and ratio processing respective to each domain.

The resulting trimmed SEM model with correlated residuals (Figure 10a) revealed a significant relationship between people's ability to estimate spatial and temporal ratios even after controlling for people's ability to discriminate absolute magnitudes. This supports the hypothesis that ratio processing in the spatial and temporal domain are related, thereby supporting the idea of a common RPS. This RPS may be related to a neural mechanism that allows for amodal representation of ratio magnitudes and is independent of absolute magnitude processing. Whereas previous studies have shown that spatial ratios are processed by similar mechanisms across different symbolic formats (Jacob & Nieder, 2009a) as well as different spatial magnitudes, such as length and numerosity (Jacob & Nieder, 2009b; Matthews et al., 2016), this is the first study to show a relationship between ratio processing across spatial and temporal domains. The findings in the current study therefore support the existence of an RPS as proposed by Matthews et al. (2016) and significantly extend the theory beyond symbolic and nonsymbolic (spatial) ratios to accommodate temporal ratios as well.

In addition to finding a significant relationship between spatial and temporal ratio processing, absolute magnitude processing factors were also found to be significant predictors of ratio processing factors in both the spatial and temporal domains. For example, the spatial magnitude factor explained 60% of the variance in the spatial ratio processing factor, while the general ratio processing factor only explained 25% of the variance in the same factor in the higher order SEM model (see Figure 10b). Similarly, the temporal magnitude factor explained 66% of the variance in the temporal ratio processing factor, while the general ratio processing factor only explained 21% of the variance in the same factor. These results indirectly replicate the relationship found in the secondary analyses conducted on data from Park et al. (2021; see online supplemental materials) as well as extend it to magnitudes in the temporal domain, which is a novel finding. Though the replication is indirect because Park et al. (2021) used ratio discrimination tasks instead of estimation tasks, finding similar results to Park et al. (2021) supports the validity of the relationships between latent factors found in the current study, despite the poor measurement qualities of the spatial magnitude indicators. These findings also align with previous neuroimaging studies on ratio and absolute magnitude processing, which found that neural correlates associated with nonsymbolic (spatial) ratio processing are similar to the correlates associated with nonsymbolic absolute magnitude processing in the prefrontal and parietal cortices in both humans and monkeys (Jacob & Nieder, 2009b; Vallentin et al., 2012; Vallentin & Nieder, 2010).

Finally, we found a high correlation between spatial and temporal magnitude factors, which was unexpected given previous literature on the association between spatial and temporal magnitudes (Mendez et al., 2011). We can think of two plausible explanations for the magnitude of this relationship. The first is that the spatial magnitude factor, which was shown to have low composite reliability, is mostly measuring general processes (e.g., working memory and decision-making) needed to successfully complete the discrimination task rather than (or in addition to) spatial magnitude acuity. This explanation would align with findings by Marcos and Genovesio (2017) showing that discrimination for different types of magnitudes might share neuronal populations related to decision-making but not magnitude processing. However, to obtain a high correlation of .85 between spatial and temporal magnitude factors, the temporal magnitude factor would also have to be measuring general cognitive processes rather than temporal magnitude acuity. This seems unlikely given that other latent variables had a high composite reliability and that the correlations between domains (i.e., spatial magnitude with temporal ratio and temporal magnitude with spatial ratio) were lower than the correlations within domains (i.e., spatial magnitude with spatial ratio and temporal magnitude with temporal ratio), indicating that there is some domain specificity to the magnitude factors (i.e., they are not solely measuring domain-general processes).

The second possible explanation is that spatial and temporal magnitudes are processed by a common magnitude system, as suggested by the ATOM theory (Bueti & Walsh, 2009; Walsh, 2003). In contrast to the possibility outlined in the previous paragraph, this domain-general process would be specific to magnitude encoding (e.g., magnitude comparison) rather than unrelated to magnitude (e.g., decision-making). Although the low factor loadings indicate that the spatial magnitude tasks had low reliability, it may also indicate that absolute magnitude processing mechanisms only partially overlap across different types of spatial magnitudes. This explanation aligns with the results of the CFA conducted on the comparison ratios that showed lower loadings across the spatial tasks and, more specifically, the dots task, possibly because it is a discrete magnitude as opposed to continuous like in the line and circle tasks. It also aligns with results from single cell studies showing that some separate but overlapping neuronal populations encode information for different types of visual magnitudes, while other neuronal populations, possibly part of a larger fronto-parietal network responsible for general magnitude processing, encode magnitude information across different types of visual magnitudes (Eiselt & Nieder, 2013; Nieder et al., 2006; Tudusciuc & Nieder, 2007). Similarly, Cona et al. (2021) have shown in a meta-analysis that neuronal populations encoding spatial and temporal magnitudes might be organized in a gradient-like way in proximal brain areas, suggesting that different types of magnitudes may be encoded separately at lower levels of abstraction but processed by a common magnitude mechanism system at a higher level of processing (Cona et al., 2021). Therefore, the common magnitude system might be, as other authors have suggested, a higher order mechanism responsible for magnitude comparison that operates only after magnitudes have been encoded by neuronal populations tuned to those specific types of magnitude (Cohen Kadosh et al., 2008; Holloway & Ansari, 2008; Pinel et al., 2004). In the context of the current study, the high correlation between spatial and temporal magnitude factors could mean that these two factors represent the same or largely overlapping higher order mechanisms.

Last, the high correlation between spatial and temporal magnitude limits our ability to make conclusions about how domain specific these mechanisms are. For instance, one may argue that the strong relationship between within-domain magnitude and ratio factors is evidence for domain-specific processes. More specifically, individuals' performance on RE tasks is largely explained by absolute magnitude factors that are related to the magnitude type and thereby are specific to either the spatial or temporal domain. However, the high correlation between the two magnitude factors also indicates that the spatial and temporal magnitude factors are controlling for much of the same variance when regressed on the ratio factors. Therefore, we cannot say with certainty that the variance explained by the magnitude factors is purely domain specific. Both magnitude factors could, in fact, be indicators of a general magnitude system. In other words, these two factors could be controlling mostly for variance related to general magnitude processing, in addition to a small amount of variance related to within-domain magnitude processing (specific to temporal or spatial magnitudes).

Absolute Versus Relative Magnitude: How Different Are They?

The results above demonstrate distinguishable yet highly correlated constructs. The fact that absolute magnitude processing factors explain such a large part of the variance in the ratio factors (approximately 60%) may lead us to question the differentiability between absolute magnitude mechanisms, such as ATOM or the approximate number system, and relative magnitude mechanisms, such as the RPS. Although both types of mechanisms are said to process approximate, nonsymbolic magnitudes, ATOM hypothesizes a common mechanism that allows one to make "smaller than" or "larger than" judgments across various types of magnitudes (Bueti & Walsh, 2009). In contrast, the RPS is hypothesized to be the mechanism that allows for the automatic extraction of a ratio magnitude between two stimuli. Another way the RPS can be interpreted is as a mechanism that allows one to make "how much larger" or "how much smaller" judgments across various types of magnitudes. It might be that the main difference between the MD and RE tasks is the type of output required in each task but that the neural mechanisms used in both tasks are actually the same (or closely overlap).

More specifically, when shown two sets of dots in a discrimination task (such as the one used in this study), a participant may automatically extract the approximate ratio between these two sets of dots and base their response on the extracted relationship. As a concrete example, if a set of 100 dots is compared to 104 dots (exact ratio of 1:1.04), the participant may perceive it as an approximate ratio of 1 to a little over 1 (e.g., 1.04 $\pm$ .02; or an approximate ratio of 1:1 if they cannot discriminate between the two sets). Therefore, even if the measured response is binary (e.g., which of the two stimuli is larger), the mechanism itself may be continuous in nature. In that case, there would be little difference between the underlying mechanism allowing one to tell whether a stimulus magnitude is "larger than" another stimulus magnitude and the mechanism allowing one to tell "how much larger" a stimulus is compared to another. To perform the MD task, one could simply have the second (ratio) mechanism operating automatically (whether it is conscious or not) and then make a conscious decision based on that information without having a dedicated absolute magnitude mechanism.

Strengths and Limitations

The current study must be interpreted in light of its strengths but also its limitations. First, the low reliability associated to the spatial magnitude tasks limits interpretation as low reliability can lead to attenuated correlations. This result is surprising because the secondary analysis on data from Park et al. (2021), which included similar discrimination tasks, yielded factor loadings between .74 and .81, which is far greater than the loadings found in the current study. Though the tasks were closely modeled after the ones used in Park et al. (2021) as well as other studies by that research group (Matthews et al., 2016; Park & Matthews, 2020), the difference in factor loadings could be partially explained by differences in task parameters, either the comparison ratios chosen or the low number of trials, and sample characteristics (i.e., children vs. adults). However, despite the reliability of the tasks, the spatial magnitude factor still showed strong relationships with the other more reliable latent variables and replicated findings from a secondary analysis on similar tasks (Park et al., 2021). Given that the correlations between the latent variables seem reasonable, we think we are measuring spatial absolute magnitude processing to some degree, along with domain-general processes related to magnitude comparison.

Second, the use of RE tasks is considered a strength but also presents some limitations. The advantage of using a RE task with the same response line across all spatial and temporal tasks is that participants are doing the same operation in all tasks: They are transforming a ratio from one (spatial or temporal) format to a representation of that ratio on a line. This limits differences between domains that may be induced by a discrimination task. For example, a temporal discrimination task in which two divided intervals are presented sequentially is probably more demanding on working memory than a spatial ratio discrimination task in which two ratios are presented side by side simultaneously. However, this might also elevate correlations between the two domains because the response format is the same (i.e., the common method is contributing to the shared variance). We provide two counter arguments. First, though the common method may be problematic for studies in which the method is orthogonal to the subject of interest, the response format (externalizing the internal representation of a ratio) is directly related to the subject of interest in the current study (ratio processing). Second, one might argue that, even though the response format is directly related to the subject of interest, participants may still have response biases unrelated to their ratio processing ability (e.g., avoiding making estimations at either end of the response line). We acknowledge this limitation but simultaneously show that using another measure of estimation accuracy less susceptible to linear biases (R^2) leads to the similar results and conclusions.

Future Directions

The results of the current study provide evidence for a relationship between ratio processing across the spatial and temporal domain. However, it does not answer the question of how spatial and temporal ratio processing are related. This relationship may have two sources of common variance: the ability to perceive or extract the value of a given ratio and the ability to represent a given ratio. However, the current study does not allow us to distinguish between these explanations of the relationship observed. It also seems unlikely that results are solely a product of task demands (i.e., estimation) as the precision of RE is limited by the precision of ratio perception by principle of precedence. Given that there may be perceptual aspects of ratio processing that are specific to different types of magnitude, parsing out ratio perception from production may shed light on domain generality versus specificity. Additionally, future studies could extend the study of ratio processing to magnitudes beyond spatial and temporal domains, such as loudness or brightness.

Additionally, future studies could investigate the development of ratio processing across different domains in order to provide insight into the underlying mechanisms related to absolute and relative magnitude processing. For example, Park et al. (2021) found that performance on a given ratio task (e.g., line length ratio) was better by the performance on a ratio task with a different format (e.g., circle area ratio) than on the absolute magnitude task with the same format (e.g., line length discrimination). Given that all the factors in this study are highly correlated with each other, determining whether absolute and relative magnitude processing across different domains diverges (as suggested by Hamamouche & Cordes, 2019), converges, or develops in parallel throughout development would provide insight into the differences between these two mechanisms (Newcombe, 2014; Newcombe et al., 2015).

Finally, previous research has shown a link between performance on nonsymbolic (spatial) ratio processing tasks and symbolic fraction knowledge. Though this study addressed the relationship between two types of nonsymbolic magnitude (temporal and spatial), future research could investigate whether temporal ratio processing also predicts performance on symbolic fraction knowledge or whether this relationship between nonsymbolic and symbolic ratio processing applies only to spatial magnitudes.

Conclusion

From the progress bars on our screens to eighth notes in music, ratios are present everywhere: We cannot help but think in a relative way. The current study has important implications as it is the first, to our knowledge, to investigate how ratio processing in space and time are related, thereby extending the literature on ratio processing beyond the numerical and visuospatial domain. Results indicate that spatial and temporal ratio processing are related even when controlling for absolute magnitude processing ability. This supports the idea of a common ratio processing system operating across different domains. We also showed that absolute magnitude processing is a significant predictor of ratio processing in both spatial and temporal domains, indicating that both mechanisms might use related underlying mechanisms. Finally, we show that spatial MD is highly related to temporal MD, which supports the idea that absolute magnitudes might also be processed by a common magnitude system across domains.

References

Agrillo, C., Ranpura, A., & Butterworth, B. (2010). Time and numerosity estimation are independent: Behavioral evidence for two different systems using a conflict paradigm. *Cognitive Neuroscience*, *1*(2), 96–101. https://doi.org/10.1080/17588921003632537

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301. https://doi.org/10.1177/1094428112470848

Alards-Tomalin, D., Walker, A. C., Kravetz, A., & Leboe-McGowan, L. C. (2016). Numerical context and time perception: Contrast effects and the perceived duration of numbers. *Perception*, *45*(1–2), 222–245. https://doi.org/10.1177/0301006615594905

Anobile, G., Burr, D. C., Iaia, M., Marinelli, C. V., Angelelli, P., & Turi, M. (2018). Independent adaptation mechanisms for numerosity and size perception provide evidence against a common sense of magnitude. *Scientific Reports*, *8*(1), Article 13571. https://doi.org/10.1038/s41598-018-31893-6

Banai, K., Fisher, S., & Ganot, R. (2012). The effects of context and musical training on auditory temporal-interval discrimination. *Hearing Research*, *284*(1–2), 59–66. https://doi.org/10.1016/j.heares.2011.12.002

Basso, G., Nichelli, P., Frassinetti, F., & di Pellegrino, G. (1996). Time perception in a neglected space. *NeuroReport*, *7*(13), 2111–2114. https://doi.org/10.1097/00001756-199609020-00009

Bonn, C. D., & Cantlon, J. F. (2017). Spontaneous, modality-general abstraction of a ratio scale. *Cognition*, *169*, 36–45. https://doi.org/10.1016/j.cognition.2017.07.012

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1831–1840. https://doi.org/10.1098/rstb.2009.0028

Cai, Z. G., & Connell, L. (2015). Space-time interdependence: Evidence against asymmetric mapping between time and space. *Cognition*, *136*, 268–281. https://doi.org/10.1016/j.cognition.2014.11.039

Cai, Z. G., & Connell, L. (2016). On magnitudes in memory: An internal clock account of space-time interaction. *Acta Psychologica*, *168*, 1–11. https://doi.org/10.1016/j.actpsy.2016.04.003

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593. https://doi.org/10.1016/j.cognition.2007.03.004

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*(2), 132–147. https://doi.org/10.1016/j.pneurobio.2007.11.001

Cona, G., Wiener, M., & Scarpazza, C. (2021). From ATOM to GradiATOM: Cortical gradients support time and space processing as revealed by a meta-analysis of neuroimaging studies. *NeuroImage*, *224*, Article 117407. https://doi.org/10.1016/j.neuroimage.2020.117407

Dormal, V., Dormal, G., Joassin, F., & Pesenti, M. (2012). A common right fronto-parietal network for numerosity and duration processing: An fMRI study. *Human Brain Mapping*, *33*(6), 1490–1501. https://doi.org/10.1002/hbm.21300

Dormal, V., Grade, S., Mormont, E., & Pesenti, M. (2012). Dissociation between numerosity and duration processing in aging and early Parkinson's disease. *Neuropsychologia*, *50*(9), 2365–2370. https://doi.org/10.1016/j.neuropsychologia.2012.06.006

Eiselt, A. K., & Nieder, A. (2013). Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. *The Journal of Neuroscience*, *33*(17), 7526–7534. https://doi.org/10.1523/JNEUROSCI.5827-12.2013

Fabbri, M., Cancellieri, J., & Natale, V. (2012). The a theory of magnitude (ATOM) model in temporal perception and reproduction tasks. *Acta Psychologica*, *139*(1), 111–123. https://doi.org/10.1016/j.actpsy.2011.09.006

Gijssels, T., Bottini, R., Rueschemeyer, S.-A., & Casasanto, D. (2013). Space and time in the parietal cortex: FMRI evidence for a neural asymmetry. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 495–500). Cognitive Science Society.

Hamamouche, K., & Cordes, S. (2019). Number, time, and space are not singularly represented: Evidence against a common magnitude system beyond early childhood. *Psychonomic Bulletin & Review*, *26*(3), 833–854. https://doi.org/10.3758/s13423-018-1561-3

Hancock, G. R., & Mueller, R. (2008). Best practices in structural equation modeling. In J. Osborne (Ed.), *Best practices in quantitative methods* (pp. 488–510). SAGE Publications, Inc. https://doi.org/10.4135/9781412995627

Hayashi, M. J., Kanai, R., Tanabe, H. C., Yoshida, Y., Carlson, S., Walsh, V., & Sadato, N. (2013). Interaction of numerosity and time in prefrontal and parietal cortex. *The Journal of Neuroscience*, *33*(3), 883–893. https://doi.org/10.1523/JNEUROSCI.6257-11.2013

Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development of number comparison. *Developmental Science*, *11*(5), 644–649. https://doi.org/10.1111/j.1467-7687.2008.00712.x

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ishihara, M., Keller, P. E., Rossetti, Y., & Prinz, W. (2008). Horizontal spatial representations of time: Evidence for the STEARC effect. *Cortex*, *44*(4), 454–461. https://doi.org/10.1016/j.cortex.2007.08.010

Jacob, S. N., & Nieder, A. (2009a). Notation-independent representation of fractions in the human parietal cortex. *The Journal of Neuroscience*, *29*(14), 4652–4657. https://doi.org/10.1523/JNEUROSCI.0651-09.2009

Jacob, S. N., & Nieder, A. (2009b). Tuning to non-symbolic proportions in the human frontoparietal cortex. *European Journal of Neuroscience*, *30*(7), 1432–1442. https://doi.org/10.1111/j.1460-9568.2009.06932.x

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. Guildford Press.

Lewis, M. R., Matthews, P. G., & Hubbard, E. M. (2015). Neurocognitive architectures and the nonsymbolic foundations of fractions understanding. In D. B. Berch, D. C. Geary, & K. M. Koepke (Eds.), *Development of mathematical cognition: Neural substrates and genetic influences* (pp. 141–164). Academic Press. https://doi.org/10.1016/B978-0-12-801871-2.00006-X

Marcos, E., & Genovesio, A. (2017). Interference between space and time estimations: From behavior to neurons. *Frontiers in Neuroscience*, *11*, Article 631. https://doi.org/10.3389/fnins.2017.00631

Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science*, *27*(2), 191–202. https://doi.org/10.1177/0956797615617799

Meert, G., Grégoire, J., Seron, X., & Noël, M. P. (2012). The mental representation of the magnitude of symbolic and nonsymbolic ratios in adults. *The Quarterly Journal of Experimental Psychology*, *65*(4), 702–724. https://doi.org/10.1080/17470218.2011.632485

Mendez, J. C., Prado, L., Mendoza, G., & Merchant, H. (2011). Temporal and spatial categorization in human and non-human primates. *Frontiers in Integrative Neuroscience*, *5*, Article 50. https://doi.org/10.3389/fnint.2011.00050

Meng, R., Matthews, P. G., & Toomarian, E. Y. (2019). The relational SNARC: Spatial representation of nonsymbolic ratios. *Cognitive Science*, *43*(8), Article e12778. https://doi.org/10.1111/cogs.12778

Möhring, W., Newcombe, N. S., Levine, S. C., & Frick, A. (2016). Spatial proportional reasoning is associated with formal knowledge about fractions. *Journal of Cognition and Development*, *17*(1), 67–84. https://doi.org/10.1080/15248372.2014.996289

Nakajima, Y. (1987). A model of empty duration perception. *Perception*, *16*(4), 485–520. https://doi.org/10.1068/p160485

Nakajima, Y., Nishimura, S., & Teranishi, R. (1988). Ratio judgments of empty durations with numeric scales. *Perception*, *17*(1), 93–118. https://doi.org/10.1068/p170093

Newcombe, N. (2014). The origins and development of magnitude estimation. *Ecological Psychology*, *2*, 147–157.

Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *WIREs Cognitive Science*, *6*(6), 491–505. https://doi.org/10.1002/wcs.1369

Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science*, *313*(5792), 1431–1435. https://doi.org/10.1126/science.1130308

Odic, D., Lisboa, J. V., Eisinger, R., Olivera, M. G., Maiche, A., & Halberda, J. (2016). Approximate number and approximate time discrimination each correlate with school math abilities in young children. *Acta Psychologica*, *163*, 17–26. https://doi.org/10.1016/j.actpsy.2015.10.010

Park, Y., & Matthews, P. G. (2020, July 30). *Revisiting and refining relations between nonsymbolic ratio processing and symbolic math achievement*. Open Science Framework. osf.io/c75xy

Park, Y., Viegut, A. A., & Matthews, P. G. (2021). More than the sum of its parts: Exploring the development of ratio magnitude versus simple magnitude perception. *Developmental Science*, *24*(3), Article e13043. https://doi.org/10.1111/desc.13043

Pastore, M., & Altoe, G. (2018). influence.SEM: Case influence in structural equation models (R package Version 2.2) [Computer software]. https://CRAN.R-project.org/package=influence.SEM

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*(6), 983–993. https://doi.org/10.1016/S0896-6273(04)00107-2

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rammsayer, T. H., & Altenmuller, E. (2006). Temporal information processing in musicians and non-musicians. *Music Perception*, *24*(1), 37–48. https://doi.org/10.1525/mp.2006.24.1.37

Rammsayer, T. H., Borter, N., & Troche, S. J. (2015). Visual-auditory differences in duration discrimination of intervals in the subsecond and second range. *Frontiers in Psychology*, *6*, Article 1626. https://doi.org/10.3389/fpsyg.2015.01626

Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods*, *25*(1), 6–14. https://doi.org/10.1177/1094428120968614

Rosseel, Y. (2012). lavaan: An R package for structural equation. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Savalei, V. (2014). Understanding robust corrections in structural equation modeling. *Structural Equation Modeling*, *21*(1), 149–160. https://doi.org/10.1080/10705511.2013.824793

Savalei, V., & Falk, C. F. (2014). Robust two-stage approach outperforms robust full information maximum likelihood with incomplete nonnormal data. *Structural Equation Modeling*, *21*(2), 280–302. https://doi.org/10.1080/10705511.2014.882692

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–243. https://doi.org/10.1111/1467-9280.02438

Skagerlund, K., Karlsson, T., & Träff, U. (2016). Magnitude processing in the brain: An fMRI study of time, space, and numerosity as a shared cortical system. *Frontiers in Human Neuroscience*, *10*, Article 500. https://doi.org/10.3389/fnhum.2016.00500

Srinivasan, M., & Carey, S. (2010). The long and the short of it: On the nature and origin of functional overlap between representations of space and time. *Cognition*, *116*(2), 217–241. https://doi.org/10.1016/j.cognition.2010.05.005

Stauffer, C. C., Haldemann, J., Troche, S. J., & Rammsayer, T. H. (2012). Auditory and visual temporal sensitivity: Evidence for a hierarchical structure of modality-specific and modality-independent levels of temporal information processing. *Psychological Research*, *76*(1), 20–31. https://doi.org/10.1007/s00426-011-0333-8

Stevens, S. S. (1960). The Psychophysics of Sensory Function. *American Scientist*, *48*(2), 226–253.

Stevens, S. S., & Galanter, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, *54*(6), 377–411. https://doi.org/10.1037/h0043680

Tsouli, A., Dumoulin, S. O., Te Pas, S. F., & van der Smagt, M. J. (2019). Adaptation reveals unbalanced interaction between numerosity and time. *Cortex*, *114*, 5–16. https://doi.org/10.1016/j.cortex.2018.02.013

Tudusciuc, O., & Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(36), 14513–14518. https://doi.org/10.1073/pnas.0705495104

Vallentin, D., Jacob, S. N., & Nieder, A. (2012). Neuronal representation of number and proportion in the primate brain. *E-Neuroforum*, *18*(2), 27–33. https://doi.org/10.1007/s13295-012-0028-7

Vallentin, D., & Nieder, A. (2010). Representations of visual proportions in the primate posterior parietal and prefrontal cortices. *European Journal of Neuroscience*, *32*(8), 1380–1387. https://doi.org/10.1111/j.1460-9568.2010.07427.x

Van Opstal, F., & Verguts, T. (2013). Is there a generalized magnitude system in the brain? Behavioral, neuroimaging, and computational evidence. *Frontiers in Psychology*, *4*, Article 435. https://doi.org/10.3389/fpsyg.2013.00435

Vibell, J., Lim, A., & Sinnett, S. (2021). Temporal perception and attention in trained musicians. *Music Perception*, *38*(3), 293–312. https://doi.org/10.1525/mp.2021.38.3.293

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Yuan, K. H., Tong, X., & Zhang, Z. (2015). Bias and efficiency for SEM with missing data and auxiliary variables: Two-stage robust method versus two-stage ML. *Structural Equation Modeling*, *22*(2), 178–192. https://doi.org/10.1080/10705511.2014.935750

Received April 13, 2022
Revision received August 12, 2022
Accepted August 15, 2022 ■