

Aus dem Zentrum für
Neurodegenerative Erkrankungen e.V. Magdeburg
Forschungsgruppe Altern und Kognition
Prof. Dr. Thomas Wolbers

Accumulation processes of time, space and numerosity: an fMRI study

Dissertation
zur Erlangung des Doktorgrades
Dr. med.
(doctor medicinae)

an der Medizinischen Fakultät
der Otto-von-Guericke-Universität Magdeburg

vorgelegt von Johannes Achtzehn, M.Sc.
aus Magdeburg
Magdeburg August 2021



Nobody has ever noticed a place except at a time, or a time except at a place.
- Hermann Minkowski

Bibliographic description

Achtzehn, Johannes:

Accumulation processes of time, space and numerosity: an fMRI study. - 2021. - 67 P., 32 Fig., 10 Tab., 1 App.

Abstract

The underlying neuronal processes of the perception of time, space and other magnitudes are subject to intense scientific debate. Furthermore, mutual interactions between the perception of different dimensions are widely reported, but the origin of such cross-dimensional interference effects is still not well understood. This thesis investigates time, space and numerosity judgements as well as their cross-dimensional influence on a behavioural and neuronal level within a navigational context. Participants underwent fMRI imaging while being passively moved forward along a virtual path that was covered with randomly appearing dots. After each trial, participants were asked to make judgements about the trial duration, distance covered or numerosity of items on the floor, allowing for a simultaneous observation of behavioural and neuronal effects.

Behavioural results revealed a correlation between time and space judgement precision as well as a bidirectional interference only between these two dimensions, but not for numerosity. Analyses of the neuroimaging data identified the right IFG as a commonly activated area during the perception of all three dimensions and MT/V5 bilaterally as being uniquely activated during numerosity trials. Mirroring the behavioural results, activation patterns of time and space trials were indistinguishable by MVPA but could be decoded from numerosity trials, predominantly in MT/V5 as well as in the IPS bilaterally (time vs. numerosity).

The findings suggest that the interference between travel time and traveled distance could be mediated by neuronal computations of movement speed on the basis of visual optic flow information. Furthermore, the results add to the body of evidence that the cross-dimensional interference effects between time and space are based on overlapping neuronal representations.

Keywords

magnitude perception, time-space interaction, psychophysics, posterior parietal cortex, fMRI, MVPA

Contents

List of Abbreviated Terms	vii
I INTRODUCTION	1
1 Motivation	2
2 Outline	4
3 Psychophysics	5
3.1 Introduction	5
3.1.1 Weber-Fechner Law	5
3.1.2 Stevens's Power Law	6
3.2 Psychometric Functions	7
3.2.1 Definition	7
3.2.2 Parameters of Psychometric Functions	7
3.2.3 Function Types	9
3.2.4 Parameter Estimation	10
4 Magnitude Processing	15
4.1 Theories of Magnitude Processing	15
4.2 Behavioural Evidence in Magnitude Processing	15
4.2.1 Discrimination and Precision	16
4.2.2 Cross-Dimensional Interference	17
4.3 Neural Correlates of Magnitude Processing	18
4.3.1 Introduction	18
4.3.2 Spatial Representation in the Brain	18
4.3.3 Neural Code for Numerical Cognition	20
4.3.4 Keeping Track of Time	20
4.3.5 Neural Evidence for a Common Magnitude System	21
II METHODS AND MATERIALS	22
5 Experimental Design	23
5.1 Participants	23
5.2 Virtual Environment	23
5.3 Task and Stimulus Description	23
5.4 Trial Structure	25
5.5 Trial Order	26
5.6 Task Difficulty	26
6 Behavioural Data Analyses	27

6.1	Preprocessing	27
6.2	Reaction Time and Task Difficulty	27
6.3	Fitting Psychometric Functions	27
6.3.1	Input Data	27
6.3.2	Fitting Algorithm	28
6.3.3	Priors	29
6.3.4	Parameter Estimates	29
6.4	Magnitude Differentiation	30
6.5	Judgement Precision	31
6.6	Cross-Dimensional Interference	31
7	Neuroimaging Data Analyses	32
7.1	Data Acquisition	32
7.2	Data Quality Control	32
7.3	Preprocessing	33
7.4	Univariate Analysis	34
7.4.1	First Level Analysis	34
7.4.2	Second Level Analysis	35
7.5	Multivariate Pattern Analysis	36
7.5.1	Introduction	36
7.5.2	Sample Generation	37
7.5.3	Inter-Subject Pattern Analysis	38
III	RESULTS	41
8	Behavioural Results	42
8.1	Manipulation Checks	42
8.1.1	Reaction Times	42
8.1.2	Perceived Task Difficulty	43
8.1.3	Magnitude Differentiation	44
8.2	Inter-Dimensional Judgement Precision	45
8.3	Cross-Dimensional Interference	46
9	Neuroimaging Results	48
9.1	Quality Control	48
9.1.1	Spatial Information	48
9.1.2	Temporal Information	48
9.1.3	Artefact Detection	49
9.2	Second Level Univariate Analysis	51
9.2.1	Introduction	51
9.2.2	Activation During Accumulation Trials	51
9.2.3	Accumulation Trials vs. Control Trials	51
9.2.4	Dimension Specific Activation	53
9.3	Multivariate Pattern Analysis	55
9.3.1	Decoding of Accumulation Trials Against Control Trials	55
9.3.2	Crossdimensional Decoding	58
IV	CONCLUSION	60
10	Discussion	61

	vi
10.1 Paradigm Design	61
10.2 Behavioural Data	62
10.3 Functional Imaging Data	63
10.4 Concluding Remarks	65
11 Summary	67
12 Zusammenfassung	68
Bibliography	69
V Appendix	1
A Methods and Materials	2
B Behavioural Results	5
C Neuroimaging Results	6
C.1 Quality Control	6
C.2 Second Level Analysis	9
C.3 Multivariate Pattern Analysis	12

List of Abbreviated Terms

Abbreviaton	Explanation
AFC	Alternative forced-choice
ANOVA	Analysis of variance
ANS	Approximate number system
ATOM	A Theory of Magnitude
CND	Cumulative Normal Distribution
CV	Cross validation
EFC	Entropy-focus criterion
EPI	Echo-planar imaging
FD	Framewise displacement
FG	Fusiform gyrus
(f)MRI	(functional) magnetic resonance imaging
FOV	Field of view
FWE	Family-wise error
GLM	General linear model
GRE	Gradient echo
GSR	Ghost-to-signal ratio
HC	Hippocampus
HRF	Hemodynamic response function
IFG	Inferior frontal gyrus

ISI	Inter stimulus interval
ITG	Inferior temporal gyrus
ITI	Inter trial interval
IPS	Intraparietal sulcus
IQMs	Image quality metrics
IQR	Interquartile range
ISPA	Inter-subject pattern analysis
LOSO	Leave-one-subject-out
MAP	Maximum a posteriori
MOD	Minimal observable difference
MPRAGE	Magnetization Prepared Rapid Gradient Echo
MVPA	Multivariate pattern analysis
PCA	Principal component analysis
PF	Psychometric function
PE	Phase-encoding
PPC	Posterior parietal cortex
PSE	Point of subjective equality
ROI	Region-of-interest
SMS	Simultaneous multislice
SVM	Support vector machine
TE	Echo time
TMS	Transcranial magnetic stimulation
TR	Repetition time
TSE	Turbo spin-echo
tSNR	Temporal signal-to-noise ratio
VAS	Visual analogue scale

PART I:
INTRODUCTION

1. Motivation

When we are late for a meeting, anxiously awaiting the bus to finally arrive, five minutes can take excruciatingly long to pass. In contrast, four hours on a Sunday afternoon spend happily with our loved ones may seemingly fly by in a heartbeat and, after checking our watch, we ask ourselves "*how has it gotten so late so fast?*". Likewise, the last 500 meters of an exhausting hike along a mountain trail sometimes feel almost impossible to complete, while at other times, we find ourselves barely noticing the many kilometres we walked around a city during a day tour.

These two examples illustrate that the perception of time and space may vary greatly and is dependent on a multitude of extrinsic and intrinsic factors. Researchers have long been fascinated by this phenomenon and emerging technologies have allowed neuroscientists to tap into the underlying neuronal processes of not only time and space, but numerosity perception as well. Naturally, this led to the question of whether there are distinct neural substrates to quantify each dimension, or if a common magnitude system unifies the perception of time, space and number.

When viewed from a behavioural standpoint, a plethora of evidence hints at a common neural correlate, which is not surprising, given how intertwined the concepts of all three dimensions are in the human mind. For example, when we suggest somebody to focus on the future, we tell them to *look ahead*, while we think of the past as being located *behind us*. Likewise, in Western culture, consecutive events are ordered from *left to right* on a piece of paper, suggesting the existence of a mental time line that gives spatial structure to temporal information (Bonato et al. 2012). The same applies for numerical information, which is usually ordered along a mental number line from *left to right* in Western culture (Bonato et al. 2012). Another argument put forward in favour of a common locus is cross-dimensional influence. Should the size of a room have an impact on the perceived duration we spend sitting in it? Intuitively, we would reject this proposition. However, evidence exists that room size does in fact bias temporal judgements (Riemer et al. 2018). This is only one example, as behavioural evidence for any combination of cross-dimensional interference has been documented (Walsh 2003, Hubbard et al. 2005, Buetti et al. 2009, Dehaene et al. 2011). Despite the abundance of behavioural results hinting at a common magnitude system, a growing number of studies report dissociations between the processing of spatial, temporal and numerical information and challenge the theory of a common magnitude system (Hamamouche et al. 2019).

Functional imaging studies in humans have provided inconclusive results so far. While some propose the right intraparietal sulcus (IPS) and inferior frontal gyrus (IFG) as important

regions for the processing of magnitudes, evidence for a common perception of all three dimensions in one locus based on functional imaging is scarce (Skagerlund et al. 2016). An even greater uncertainty exists concerning a direct involvement of specific regions in cross-dimensional interaction, mainly because past studies have either focused on behavioural surrogates or did not use visually identical stimuli to test time, space and numerosity perception. This thesis set out to close this gap by carefully designing a functional magnetic resonance imaging (fMRI) paradigm that utilized visually identical stimuli for all dimensions. In addition, it was ensured that the magnitude of time, space and number could only be perceived accumulatively. Hence, this study enabled a direct comparison of the perception of all three dimensions, as well as their inter-dimensional interaction, both on the behavioural and neuronal level. The following list summarizes the main aims of this thesis:

Investigate whether there is evidence on the behavioural level for a shared magnitude processing system by measuring the precision in time, space and numerosity judgements as well their inter-dimensional interference effects.

During the accumulation phase of time, space and numerosity judgements, measure whole brain BOLD activation and analyse it in order to (i) identify areas commonly activated by multiple dimensions as well as (ii) characterize neuronal activation patterns with multivariate pattern analysis.

2. Outline

This thesis is divided into four parts. The first part starts of with an introduction into basic concepts of psychophysics, including the methodology of psychometric functions, as these are an important tool to quantify the behavioural results of the experiment. The second chapter reviews existing literature concerning key aspects of magnitude processing, both from a behavioural and a functional imaging perspective.

The second part is dedicated to the methods and materials used in this experiment. First, the experimental design is described in chapter 5, followed up by an overview of the statistical tools employed for the analyses of the behavioural data. Because fitting psychometric functions to the input data is an important aspect of the behavioural analyses, the Bayesian approach is explained in greater detail. Chapter 7 outlines the processing pipeline for the functional imaging data, starting with preprocessing steps and ending with the methodology of multivariate pattern analysis algorithms.

Part three contains the results obtained from analyzing the behavioural and BOLD imaging data. After a first manipulation check in section 8.1, results from the behavioural data are presented. Chapter 9 begins by presenting results from data quality control analyses. Subsequently, univariate and decoding results are presented in sections 9.2 and 9.3.

The last parts discusses and summarize the obtained results in the light of existing literature introduced in the first part.

3. Psychophysics

3.1 Introduction

3.1.1 Weber-Fechner Law

The field of psychophysics studies the quantitative relationship between a physical stimulus and the sensory and perceptual experience by a (human) observer of said stimulus. Hereby, a challenging aspect is that physical stimuli (e.g. brightness of a lightbulb) are accurately measurable, while the sensation and perception of such stimuli are a private, unobservable experience and thus, cannot be easily quantified. During the 19th century, scientists such as Hermann von Helmholtz or Ernst Heinrich Weber began to empirically study psychophysics and to formulate theoretical laws governing human perception. Most notably, Weber and his student Gustav Theodor Fechner first discovered that the minimal observable difference ΔI (MOD) of two stimuli is proportional to the magnitude I at which it is measured, now known as Weber's law (or Weber's ratio) (Fechner 1860):

$$\frac{\Delta I}{I} = k \quad (3.1)$$

The constant k is specific for the sensory system being measured. In practical terms, two important conclusions may be drawn from this formulation: (i) for small baseline intensities of a stimulus, small changes in magnitude are sufficient to evoke a perceived difference whereas greater changes in intensity are needed if the initial intensity is larger and (ii) the value of k gives insight into the sensitivity of a sensory systems when compared to others. Following this discovery, Fechner set out to formulate a more general law that relates perceived stimulus intensity to the changes in stimulus strength. He assumed that for a specific dimension, perceived differences between two stimuli separated by the minimal observable difference will be equal for all stimulus levels. Or in other terms: the MODs for smaller and larger initial values may differ significantly (according to Weber's ratio), but the perceived difference in intensity will be the same. Essentially, he describes a logarithmic relationship between perceived stimulus magnitude E and the physical intensity I (Fechner 1860):

$$E = k \cdot \ln \frac{I}{I_0} \quad (3.2)$$

The constant k is specific for each modality and I_0 denotes the minimal intensity at which a stimulus is perceived. This logarithmic relationship is also called *compression*, as sensory systems become less sensitive with increase stimulus intensities. Both Weber's ratio and Fechner's law are illustrated in Figure 3.1.

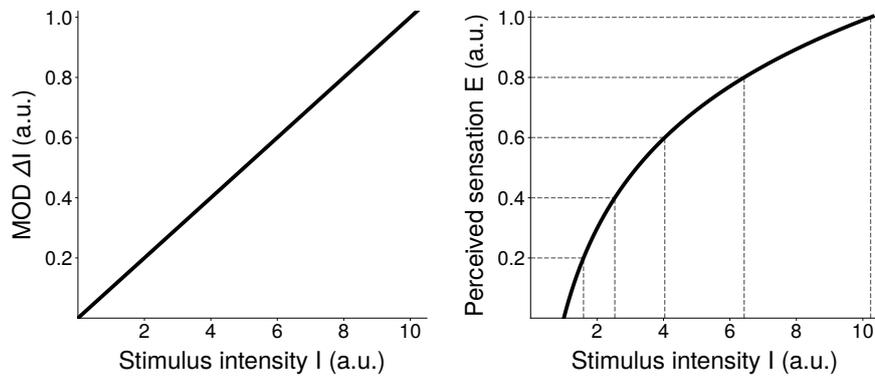


Fig. 3.1: Left: According to Weber's ratio, the MOD ΔI is linearly proportional to the magnitude I at which it is measured. Right: Fechner's law relates perceived stimulus intensity E to the stimulus intensity I . Note the logarithmic increase in physical stimulus intensity and linearly scaling of the magnitude of perceived sensation.

3.1.2 Stevens's Power Law

Around 100 years later, Stanley Smith Stevens challenged the formulations of Weber and Fechner (Stevens 1961). Indeed, if compared to knowledge available today, Weber's ratio is only a reasonable approximation for moderate intensities and fails short for both the low and the high end of the stimulus range. Stevens specifically refuted the assumption that there is a logarithmic relationship between stimulus and sensation intensity and instead proposed a power-law dependency:

$$E = c \cdot I^b \quad (3.3)$$

In Stevens's equation, c is a constant scaling factor that considers the units in which the stimulus is measured (e.g. grams, seconds) and b is a constant sensory specific exponent to which the intensity is raised. For most modalities, b takes on values smaller than 1 and the shape of the function in equation 3.3 becomes very similar to the logarithmic one of Fechner's law. However, there are modalities for which the opposite is observed: for example, participants become more sensitive to the pain induced by electrical shocks with increasing intensities. This phenomenon could not be explained by Fechner's law, but Stevens's power law describes these cases well if $b > 1$. For sensory systems that follow a near linear relationship (e.g.) apparent length, equation 3.3 takes on a linear shape if b is close to 1. Figure 3.2 illustrates the aforementioned cases.

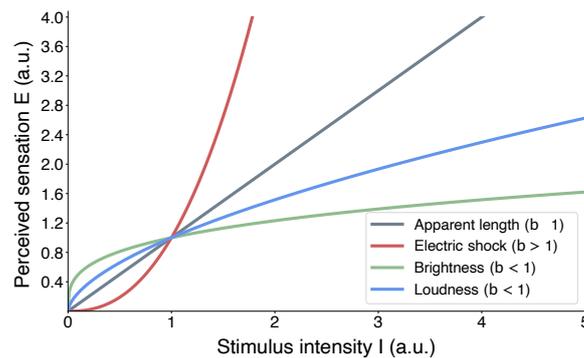


Fig. 3.2: Power functions relating stimulus intensity to perceived sensation for different dimensions. Many modalities take on a shape similar to a logarithmic function with $b < 1$. In contrast to Weber's and Fechner's theory, Stevens's power law may also describe cases in which there is an exponential relationship between stimulus intensity and perception, demonstrated here for the pain experienced by electrical shocks.

Nowadays, Stevens's power law is accepted as a valid representation among the scientific community. Following his publication, research has focused on refining his theories by taking into account e.g. observer bias, individual noise and other parameters that influence perception (Teghtsoonian et al. 1971, Robinson 1976). A wide array of diagnostic tools and statistical models enable the investigation of not only basic sensory systems (e.g. the auditory system), but also more complex perceptions, such as time or numerosity perception, as done in this study. Chapter 4 will give more insight into the more complex perception of time, space and number.

3.2 Psychometric Functions

3.2.1 Definition

In modern psychophysics, scientists model responses with a psychometric function (PF), which most commonly is a cumulative probability density function $\psi(x)$ relating the probability of a certain judgement on the ordinate to the intensity x of a physical stimulus on the abscissa. Typically, the judgement is either made in a yes-no fashion (detection of a stimulus) or by comparing two stimuli against each other (discrimination of two stimuli) at varying intensities or intensity differences¹. Figure 3.3 illustrates an example of a PF fitted to surrogate data of a two-alternative forced-choice (2-AFC) discrimination task.

3.2.2 Parameters of Psychometric Functions

During the fitting procedure, a number of parameters are estimated and will be discussed subsequently. Before each parameter is introduced, it is important to take into account that the true value of a parameter p cannot be exactly determined. Instead, by fitting a statistical

¹ There are a wide array of different approaches on how to vary stimulus intensities during the course of an experiment. To explain these is beyond the scope of this thesis, please refer to (Kingdom et al. 2016) for a detailed explanation.

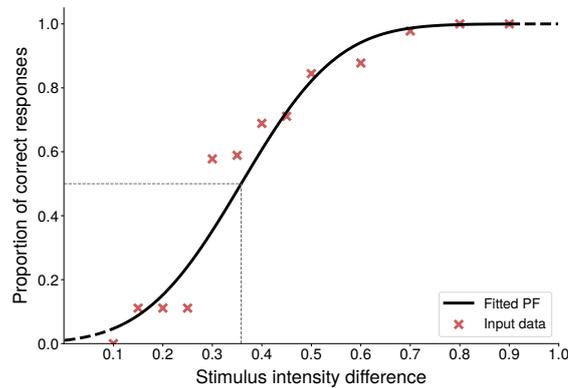


Fig. 3.3: Hypothetical data from a two-alternative forced-choice task in which observers had to indicate whether two stimuli are different from one another. The ordinate represents the probability of 'Yes' responses of n trials ($\psi(x)$) at each intensity difference (x) on the abscissa. The vertical dotted line on the x -axis indicates the threshold $\hat{\alpha}$ at which the participant perceives both stimuli as identical and thus, the performance is equal to the guessing rate $\hat{\gamma} = 0.5$.

model to the data, an approximation \hat{p} is calculated that, by nature of an approximation, has a certain margin of error. Consequently, it is important to assess the goodness-of-fit after computing a PF. Section 3.2.4 will provide more detail on the error assessment.

The first parameter, $\hat{\alpha}$ specifies the overall position of the fitted function along the x -axis. It simultaneously defines the threshold at which a desired performance is reached. For a stimulus detection task, this could mean that the stimulus is successfully detected when 75 % of trials are correct. For a task comparing two stimuli (as in this study), $\hat{\alpha}$ may also be interpreted as the *point of subjective equality* (PSE), at which the performance is 50 % and two stimuli are perceived as equal.

The second parameter needed to characterize a PF is denoted by $\hat{\beta}$ and controls the steepness of the function. Therefore, it is often called the slope of the PF, which is not entirely correct, given how it is defined mathematically. More recently, researches have argued that the width of the function is a more appropriate parameter, as it is (i) defined in the same units as the physical stimulus, (ii) independent of the employed sigmoid function and consequently, (iii) only a single set of priors has to be defined, irregardless of the sigmoid function chosen (Alcala-Quintana et al. 2004, Kuss et al. 2005). Because the PF is asymptotical towards 0 for small intensities and 1² for larger ones, the width should not be defined as the intensity at which $\psi(x)$ reaches both values. Instead, an arbitrary value (e.g. 0.05) is subtracted from the lower and upper asymptotes. The width is then defined as the difference in intensity at which $\psi(x)$ reaches those values (e.g. 0.05 and 0.95 for the lower and upper value, respectively).

Two more parameters, $\hat{\gamma}$ and $\hat{\lambda}$, are necessary to fully define the PF. In contrast to $\hat{\alpha}$ and $\hat{\beta}$, they are not dependent on the underlying sensory mechanism. The variable $\hat{\gamma}$ defines the guessing rate and is defined by the amount of choices participants have during their

² This is only true in the ideal scenario in which the participant does not make any mistakes (lapses). For the sake of simplicity, this is omitted here because the lapse rate λ is not yet introduced.

response. For an n-AFC, $\widehat{\gamma}$ is typically defined as $1/n$. As an example, consider the hypothetical data of Figure 3.3: participants had to choose between two alternatives and thus, $\widehat{\gamma}$ is defined as $1/2$ (0.5). The fourth parameter $\widehat{\lambda}$ is termed as the lapse rate. It accounts for the small amount of trials in which participants will respond independently of the stimulus level, perhaps because they hit the wrong response button or missed the cue presentation. Consequently, the upper and lower asymptotes of the PF will not be exactly 1 and 0, but rather $1 - \widehat{\gamma}$ and $0 + \widehat{\gamma}$, respectively.

3.2.3 Function Types

For a wide range of tasks, the PF assumes the shape of a sigmoidal function. There are several different functions available to describe the relationship between performance $\psi(x)$ by the participant and the stimulus intensity x . A general formulation of the PF is:

$$\psi(x; \alpha, \beta, \gamma, \lambda) = \gamma + (1 - \gamma - \lambda) \cdot F(x; \alpha, \beta) \quad (3.4)$$

In this section, only the three most popular options for $F(x; \alpha, \beta)$ are briefly described. Please refer to (Davis 2008) and (Kingdom et al. 2016) for a complete description of all possible functions.

Cumulative Normal Distribution

An important factor in fitting a PF is how noise is affecting the judgement probability (performance). The Cumulative Normal Distribution (CND) is a popular choice for $F(x; \alpha, \beta)$, as it assumes a normally distributed noise. This is in line with the Central Limit Theorem, which states that the linear combination of independent noise sources results in a normally distributed noise (Fischer 2011). The CND is defined as follows:

$$F(x; \alpha, \beta) = \frac{\beta}{2\pi} \int_{-\infty}^x \exp\left(-\frac{\beta^2(x-\alpha)^2}{2}\right) \quad (3.5)$$

with: $x, \alpha \in (-\infty, +\infty)$ and $\beta \in (0, +\infty)$

Parameter α denotes the threshold $F(x = \alpha; \alpha, \beta) = 0.5$. When β is kept constant, variations in α result in a rigid transformation along the x-axis. The slope β is determined by the reciprocal of the standard deviation of the normal distribution. For tasks in which $x = 0$ is equal to the absence of a stimulus and x is not log-transformed, the CND is not a valid solution (Kingdom et al. 2016).

Logistic Function

The CND includes an integral for which no analytical solution is known. In contrast, the Logistic function has a known closed-form integral and thus, does not need to include numerical methods to approximate the integral. It is specified as:

$$F(x; \alpha, \beta) = \frac{1}{1 + \exp(-\beta(x - \alpha))} \quad (3.6)$$

with: $x, \alpha \in (-\infty, +\infty)$ and $\beta \in (0, +\infty)$

Analog to the CND, α and β correspond to the threshold and slope of the PF, respectively. The same limitations regarding the absence of a stimulus for $x = 0$ for the CND apply to the Logistic function.

Weibull

The cumulative distribution function for the Weibull distribution is defined as:

$$F(x; \alpha, \beta) = 1 - \exp\left(-\left(\frac{x}{\alpha}\right)^\beta\right) \quad (3.7)$$

with: $x \in [0, +\infty)$ and $\alpha, \beta \in (0, +\infty)$

Because for $x < 0$, $F(x; \alpha, \beta) = 0$ for all α, β , the Weibull distribution should not be used when x is measured in logarithmic units. In the case of $x = \alpha$, $F(x = \alpha; \alpha, \beta) = 1 - \exp(-1) \approx 0.632$ for all values of β . An important difference between the Weibull distribution and the other two presented functions is that manipulating α will not result in a rigid transformation on the x-axis and thus, change β as well if not plotted against $\log(x)$.

In order to give a practical example of fitted PFs based on all three functions, a fit to the same hypothetical input data as in section 3.2.1 was done and plotted in Figure 3.4.

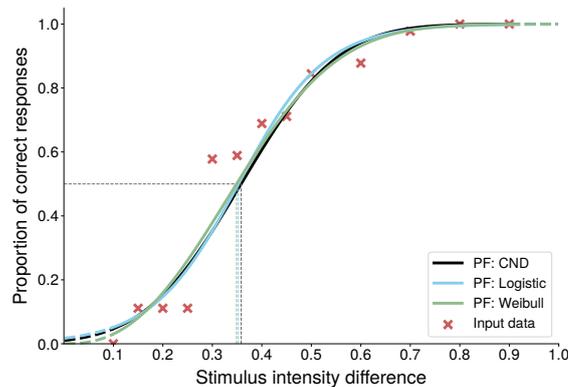


Fig. 3.4: For the same input data as in Figure 3.3, the PFs were fitted using the three different function types (i) CND, (ii) a Logistic function and (iii) the Weibull distribution. While their shapes exhibit small differences, the threshold at 50 % performance is nearly identical.

3.2.4 Parameter Estimation

Once a function type has been defined, the parameters introduced in section 3.2.2 have to be estimated in order to characterize the PF. Only α and β rely on the sensory mechanism. The guess rate γ is given by the experiment structure and the lapse rate λ is usually considered a nuisance value, as it may help to describe the PF but does not contain any information

about the underlying sensory mechanism. Hence, most studies are primarily interested in determining the values for α and β .

There are different methods to identify the best-fitting set of parameters. Among those, the two most common are (i) the maximum likelihood approach and (ii) Bayesian analyses. In this thesis, only the latter approach will be described in detail, as it will be used in the subsequent data analysis. Prior to that, the term likelihood will be defined, as it will help to relate and understand Bayesian statistics.

The Concept of Likelihood

Statistical problems often calculate probabilities for the outcome of an experiment, given a fixed set of parameters. In contrast, fitting a PF comprises the inverse problem: given the (already known) outcome and a model of interest (function type), the aim is to find a set of parameters (α , β) that is most likely to have produced the observed data. In order to solve this problem, a likelihood function $L(a|\mathbf{y})$ is defined, representing the likelihood of parameter a given the observed data \mathbf{y} (Boos et al. 2013).

Because experiments in Psychophysics typically consist of data obtained by responses to independent binary choices (i.e. a Binomial distribution), the likelihood function may be defined as:

$$L(a|\mathbf{y}) = \frac{N!}{m!(N-m)!} \prod_{i=1}^N p(y_i|a) \quad (0 \leq a \leq 1) \quad (3.8)$$

or: $L(a|\mathbf{y}) = \frac{N!}{N!(N-m)!} \cdot a^m(1-a)^{N-m} \quad (0 \leq a \leq 1)$

In this equation, N denotes the total amount of trials, m the number of successful trials and $p(y_i|a)$ the probability of observing the outcome \mathbf{y} on the specific trial i under the assumption of a value a . As a practical example, consider an experiment consisting of ten trials, in which a participant had to judge whether she or he saw a cat in the presented picture or not. Comparing the responses to the actual stimulus, the participant successfully classified six trials correct (S) and missed four (M). If we assume a success rate of $a = 0.4$ for a successful classification, equation 3.8 becomes (in its short form):

$$L(0.4|\mathbf{y}) = \frac{10!}{6!(10-6)!} \cdot 0.4^6(1-0.4)^4 \quad (3.9)$$

$$\approx 0.1115$$

It is important to note that obtained value for $L(0.4|\mathbf{y})$ is neither the probability of obtaining the specific result \mathbf{y} (the experiment is already over, so there is no uncertainty about the outcome), nor is it the probability of $a = 0.4$. Instead, one should think of $L(a|\mathbf{y})$ as a function in the parameter space that provides information about the likelihood of this particular parameter given a fixed data set. For the hypothetical experiment of identifying a cat, Figure 3.5 shows full likelihood function in the range of $0 \leq a \leq 1$. The global maximum of this

function is at the value $\hat{a} = 0.6$ and is therefore the maximum likelihood estimation for the true value of a .

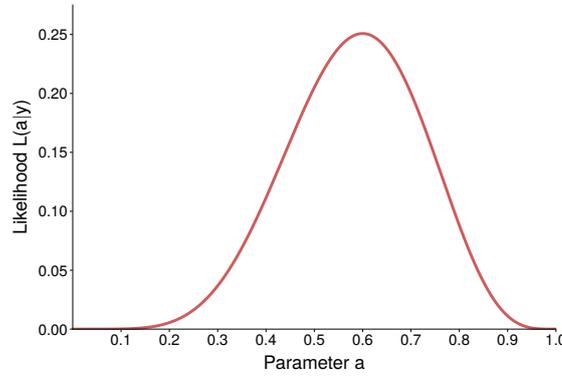


Fig. 3.5: The likelihood as a function of the parameter a , given the outcome data of the described hypothetical experiment.

As stated above, experimenters are usually concerned about two parameters of a Psychometric functions, α and β at certain stimulus intensity x_i . Equation 3.8 thus becomes:

$$L(\alpha, \beta | \mathbf{y}) = \frac{N!}{m!(N-m)!} \prod_{i=1}^N p(y_i | x_i; \alpha, \beta) \quad (3.10)$$

Similar to equation 3.8, $p(y_i | x_i; \alpha, \beta)$ represents the probability of a response y_i for a specific trial i , assuming the threshold α and the slope β of the PF. Since there are now two parameters present, the likelihood function assumes the form of a surface that lies within the two-dimensional parameter space. Naturally, the task of finding the global maximum for this function becomes much more complicated and an analytical solution is almost never found. Instead, researchers employ iterative search algorithms for this task. Since this is beyond the scope of the present thesis, interested readers are referred to (Myung 2003).

Bayesian Estimation

The Bayesian framework provides probability distributions over parameter values in order to describe beliefs and uncertainties about their values in relation to the obtained data (Kuss et al. 2005). A practical example is null-hypothesis testing, in which the obtained p-value describes the probability that the obtained results have occurred randomly under the assumption that the null-hypothesis was true. This can be formulated as $p(D|H)$, in which D is an observable outcome and H represents a true null-hypothesis. In the context of Psychophysics, this can be adapted to $p(D|\phi)$, whereas ϕ denotes modelled parameters, e.g. threshold and slope of the PF³. However, in order to obtain estimations for the parameter set ϕ , inferences about $p(\phi|D)$ need to be made. In the Bayesian framework, $p(\phi|D)$ is also

³ Keep in mind that the true values for these parameters are hidden within the participant and that Bayesian inference aims to minimize the uncertainty about their value.

called the *posterior* distribution and, with the help of Baye’s Theorem, is related to $p(D|\phi)$ by the following equation:

$$p(\phi|D) = \frac{p(D|\phi)p(\phi)}{p(D)} \tag{3.11}$$

Here, an important advantage of the Bayesian approach becomes evident: *a priori* beliefs about the true value of ϕ can be modelled and are represented by $p(\phi)$, the so called *prior* distribution. This equation may now be adapted to the task of estimating the posterior distribution $p(\alpha, \beta|\mathbf{y})$, given the likelihood function $L(\alpha, \beta|\mathbf{y})$ from equation 3.8 and the prior distributions $p(\alpha)$ and $p(\beta)$:

$$p(\alpha, \beta|\mathbf{y}) = \frac{L(\alpha, \beta|\mathbf{y})p(\alpha)p(\beta)}{\int L(\alpha, \beta|\mathbf{y})p(\alpha)p(\beta)\delta\alpha\delta\beta} \tag{3.12}$$

Ideally, priors should reflect beliefs of the true model parameters. A sensible approach here is to use a parametric family of distributions for each of the parameters to be estimated. As a practical example, prior distributions for the hypothetical data of Figure 3.3 are plotted in Figure 3.6. A uniform distribution over the range of the input data with a cosine fall off in both directions over half of the range was chosen as a prior for the threshold α , as it is believed that the threshold lies within the range of the input data with equal likeliness. For the width β , the prior has a uniform distribution in the range of twice the minimal difference between two intensities with a cosine falloff at three times the total range of the stimuli for higher intensities and at the minimal difference for intensities near zero.

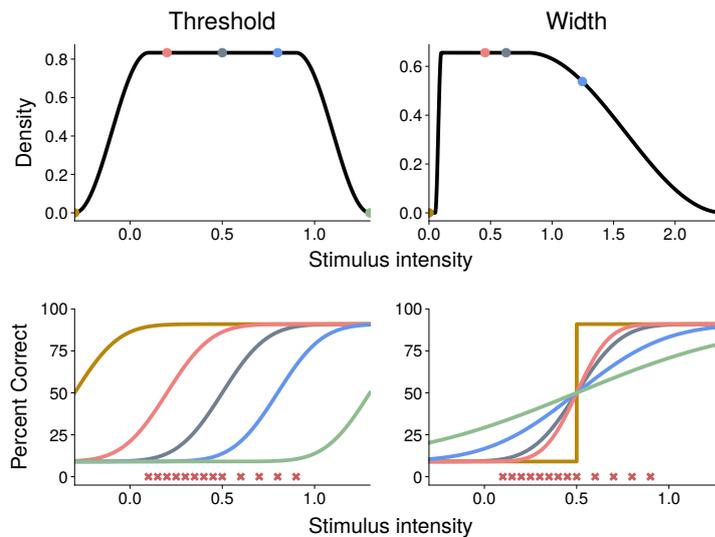


Fig. 3.6: Prior distributions for the threshold and the width for the hypothetical data used in Figure 3.3. The top row shows the prior densities and in the bottom row, the corresponding PFs for the 0 %, 25 %, 50 %, 75 % and 100 % quartiles are plotted. The sampled stimulus intensities are plotted with red crosses.

As stated above, the posterior distribution directly represents the uncertainty about parameters of the PF and could thus be utilized to make confidence judgements. Unfortunately, it is not straight forward to solve equation 3.12 analytically. Over the years, several approaches have been developed to obtain estimates and confidence intervals. Typically, either the mean, median or mode (also called maximum a posteriori (MAP)) of the posterior are estimated. Interested readers are referred to (Kuss et al. 2005) for an in-depth review. More recently, toolboxes that offer grid-based numerical solutions for the posterior have been published (Schütt et al. 2016) and are employed in this thesis. As an example, posterior distributions for the threshold and width of the hypothetical data from Figure 3.3 have been calculated and plotted in Figure 3.7.

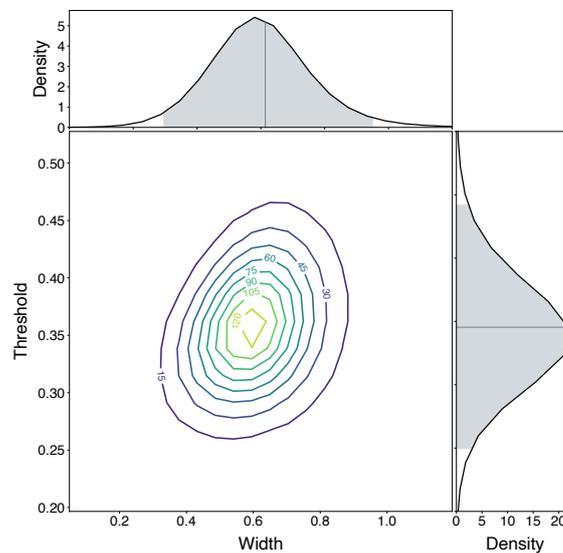


Fig. 3.7: Calculated posterior distributions for the threshold and width for the hypothetical data shown in Figure 3.3. The two plots on the side represent the posterior density functions for each parameter separately. Hereby, the vertical black line denotes the MAP value. Shaded gray areas fill the 95 % credible intervals.

4. Magnitude Processing

4.1 Theories of Magnitude Processing

Over the last decades, a multitude of theories have been proposed on how quantities are represented in the human and animal brain. Many of them have followed the concept of an innate common magnitude system integrating all modalities, while others have argued that distinct neural representations keep track of time, space and number, especially in developmental stages past infancy.

In 1983, a first hint of a shared cognition of time and space was discovered, when Meck et al. (Meck et al. 1983) reported that rats (i) can generalize rules learned in the time domain to the number domain and (ii) that the temporal and numerical discrimination characteristics follow Weber's law. In 2003, Vincent Walsh (Walsh 2003) published his impactful paper on *A Theory of Magnitude (ATOM)*, in which he argues that (i) not only time and number, but also spatial perception share a common metric for action, located in the inferior parietal cortex and (ii) that a common magnitude system is already present at birth.

Alternatively, in contrast to a common locus, individual modules could be responsible for magnitude processing of each dimension. Most prominently, this is formulated in the idea of the *Approximate Number System (ANS)* (Dehaene et al. 1998, Brannon 2006, Odic et al. 2018), which is thought to be innate and domain-specific for the processing of numerical magnitudes. Other groups have focused their research with regard to specialized representations of time (Buhusi et al. 2005, Buhusi et al. 2009) and space (Whitlock et al. 2008, Vasilyeva et al. 2010, Vasilyeva 2012).

A compromise between the two concepts was proposed by Newcombe (Newcombe 2014), called the *Development Divergence Model*, by arguing that a common magnitude system exists during infancy, which then, during later development, establishes distinct representations for each domain. This notion is backed up by the *Signal Clarity Theory* (Cantrell et al. 2013, Cantrell et al. 2015), suggesting that infants perceive all quantities as a single dimension. A paper by Hamamouche et al. (Hamamouche et al. 2019), reviewing the plethora of evidence for both a common and distinct magnitude processing, concludes that the current state of research favours the development divergence model.

4.2 Behavioural Evidence in Magnitude Processing

If a common magnitude system is responsible for processing information about time, space and number, a reflection of this should be measurable in behavioural experiments. Below,

a number such implications will be introduced. Evidence in the literature for the theorems important to the current thesis (namely cross-dimensional interference and correlation in performance) will be provided in further detail in the following sections. For a complete review, please refer to (Hamamouche et al. 2019).

Discrimination and precision There should be an inter-dimensional correlation in acuity within participants, i.e. if participants make highly precise judgements about time, they should also be equally exact when judging other magnitudes. If there is a change in precision during development, this change should be similar across dimensions. In addition, the discrimination precision of all three modalities should follow a common rule, for example Weber's law.

Cross-dimensional interference A common magnitude processing would implicate that there is a symmetric cross-dimensional influence of one domain to another, resulting in a biased response.

Cross-dimensional transfer Rules learned in one domain should transfer to other dimensions, either by instruction or spontaneously.

Contextual influence Quantity judgements should be influenced by external factors evenly, if they share a common processing. For example, the emotional state of a participant should be reflected in a similar bias in the responses of all three domains. The same should apply for other covariate factors, such as cognitive load.

4.2.1 Discrimination and Precision

There is ample evidence that the ability to discern time, space and number follow Weber's fraction in early development stages. According to Brannon et al. (Brannon et al. 2006), 6-month old infants need a 1:2 ratio to perceive a difference in area of visual stimuli and are unable to detect changes in area with a 2:3 ratio. Additionally, infants interest¹ in novel stimuli was modulated by the degree of area difference between stimuli. Similarly, research concerning change in number (Xu et al. 2000, Lipton et al. 2003) and duration (Vanmarle et al. 2006), discovered the same requirement for a 1:2 ratio of intensity difference to enable successful detection by 6-month old infants. At the age of 10 month, ratios of 2:3 can successfully be discriminated, whereas changes within a ratio of 3:4 are indistinguishable (Brannon et al. 2007, Cordes et al. 2008).

Research concerning children and adults does not draw such clear conclusions. In early childhood, spatial and numerical acuity seem to differ (Odic et al. 2012, Geary et al. 2016). When examining the discrimination precision of children between the age of two and twelve, an independent development for the sense of number was observed (Odic 2017) and thus, supports the existence of an ANS (Odic et al. 2018). Furthermore, adults seem to have a finer grained discrimination ability for area (9:10 ratio) when compared to number (8:9 ratio) (Odic et al. 2012). Droit-Volet et al. (Droit-Volet et al. 2008) obtained intriguing results when comparing discrimination precision between time, number and line length:

¹ The authors recorded looking behaviour of the infants and quantified the time they spend looking at the computer screen as a surrogate parameter for novelty.

in a first experiment, they presented numerical and spatial stimuli non-sequentially and found a reduced acuity for time compared to space and number. In a second experiment, sequential presentation was chosen for all stimuli and the differences in precision disappeared. These ambiguous results highlight the importance of presentation mode: time is inherently sequential (i.e. it can only be presented in an accumulative manner), while space and number may be presented sequentially (e.g. counting the number of dots that appear and disappear on a computer screen) or non-sequentially (e.g. judge the total amount of dots simultaneously present on the computer screen for a brief period of time).

A study comparing symbolic numerical abilities with non-symbolic numerical, temporal and spatial² abilities found a positive correlation between symbolic and non-symbolic numerical accuracy, but no correlation between non-symbolic numerical, temporal or spatial acuity (Agrillo et al. 2013). Others have reported cross-dimensional correlation in judgement precision, although not consistently for all modalities (Dewind et al. 2012, Jang et al. 2015). In their review, Hamamouche et al. (Hamamouche et al. 2019) call for caution when interpreting these results due to a low reliability between the tasks and different approaches on stimulus intensity scaling. Still, they come to the conclusion that, in summary, the incongruent findings undermine the support for a common magnitude system.

4.2.2 Cross-Dimensional Interference

Research investigating the interaction between the dimensions of time and space has revealed an asymmetric relationship: for basic visual stimuli, such as a line on a computer screen, time judgements seem to be biased by spatial properties, but not vice-versa (Xuan et al. 2007, Casasanto et al. 2008, Bottini et al. 2013). Analogue to studies mentioned in section 4.2.1, it is often criticized that the sequential vs. non-sequential presentation of time and space has to be taken into consideration. However, experiments presenting spatial and time-based stimuli both in a sequential manner (Lambrechts et al. 2013, Martin et al. 2017) came to a similar conclusion. Interestingly, Riemer et al. (Riemer et al. 2018) discovered a symmetric interaction between time and distance judgements in a static large-scale virtual environment (experiment 1), but an asymmetric interference when participants were passively moved along a path and had to judge travelled time and distance (experiment 2). Here, traversed distance strongly influenced travel time judgements, but not vice versa.

Regarding the cross-dimensional interaction between time and number, the results of a stroop experiment by Dormal et al. (Dormal et al. 2006) mirror those of time-space interference: time processing is affected by numerosity, but numerical judgements are unaffected by temporal cues. In stark contrast, Agrillo et al. (Agrillo et al. 2010) have found no inter-dimensional influence between time and number and argue against a common magnitude system.

Studies investigating all three modalities at once are scarce. Dormal et al. (Dormal et al. 2013) expanded their stroop-experiment to include all three dimensions and discovered

² Symbolic numerical abilities were assessed with mental calculations and mathematical reasoning. Non-symbolic judgement precision was quantified by evaluating line length, duration of a tone and numerosity of dots.

asymmetric interactions: numerosity and length (space) both interfere with duration processing separately and cumulatively, while time had no effect on neither of the other dimensions. Both number and space cross-influenced each other. These findings are opposed by the report of Lambrechts et al. (Lambrechts et al. 2013), who saw an influence of temporal cues on the judgement of number and space, but temporal judgement were unaffected by number and/or spatial properties.

4.3 Neural Correlates of Magnitude Processing

4.3.1 Introduction

So far, only behavioural correlates for the processing of time, number and space have been considered. Examination of neuronal activity may identify topographic organization underlying magnitude processing and could give insight into cross-dimensional interference effects found on the behavioural level. The following sections are structured by dimension and will mainly focus on neural activation measured by fMRI as well as the impact of lesions. An excellent in-depth review is provided in (Hamamouche et al. 2019). The perception of space in the context of spatial navigation has been a key focus of research, both by single unit recordings in animals as well as fMRI in humans. Results have provided compelling hints at a complex system of different cell types within the hippocampus (HC) that provide different aspects of spatial information (Eichenbaum 2017). Because the experimental paradigm employed in this thesis investigates magnitude processing within a navigational context, a separate paragraph is dedicated to the perception of space in the framework of navigation.

4.3.2 Spatial Representation in the Brain

Spatial Perception in the Context of Navigation

Among the three domains, spatial processing in the context of navigation was the first one for which scientists discovered a potential neural substrate with the discovery of *place cells* in the rat's HC, first described by John O'Keefe (O'Keefe 1976). He concluded that place cells inside the HC play an intricate part in facilitating the formation of a stable spatial map³ of the surrounding environment. Their research sparked an extensive investigation on place cells and their properties, not only in rats but also in other species, such as bats (Geva-Sagiv et al. 2015, Hoffman et al. 2016), birds (Sherry et al. 2017) and humans (Kim et al. 2017). Nowadays, it is widely accepted that place cells process information of various modalities (Dabaghian et al. 2014, França et al. 2019) and are only one part of a more complex system of different brain regions that support a multitude of spatial representations.

³ The idea of such a map dates back several decades earlier, when Tolman (Tolman 1948) proposed the existence of a cognitive map in order to explain the behavioural results he obtained by observing navigation in rats. With his theory, he opposed the generally accepted purely *stimulus-response* concept of navigational abilities at this time. He argued that behaviour such as taking shortcuts or latent learning could only be explained if the animal had a stable spatial representation of its surroundings.

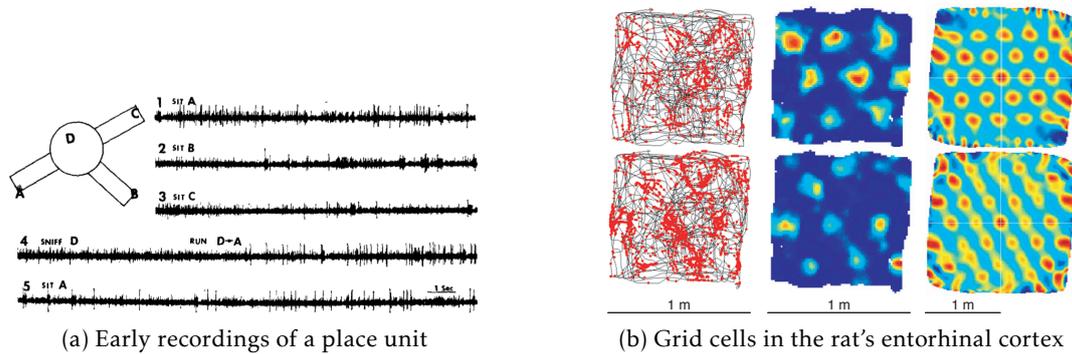


Fig. 4.1: (a): The cell represented here fired only at a high rate when the rat sat in the field marked with A. Note the frequent spikes in 1 during a phase of immobility at A, as well as in 4, when the animal moves from position D to A (O'Keefe 1976). (b): Grid cell recordings in the rat's entorhinal cortex. Left most plot shows the trajectory of the rat (black) and spike locations (red). Peak rate of those locations is drawn in the middle. Spatial autocorrelation for the rate map is depicted on the right (Hafting et al. 2005).

Other components of such a system are cells encoding allocentric orientation, termed *head-direction cells* (Taube et al. 1990, Wiener et al. 2005, Shine et al. 2016) and *border cells*, which are tuned to borders of an environment (Solstad et al. 2009, Shine et al. 2019). *Grid cells*, first described in rats by the team of May-Britt and Edward Moser (Hafting et al. 2005), are located in the entorhinal cortex⁴ and exhibit spatial-tuned firing pattern much like place cells, but have multiple firing fields across the environment, forming a periodic triangular lattice. An extensive review on different cell types facilitating spatial navigation in the hippocampal and parahippocampal formation is provided here (Ekstrom et al. 2003, Moser et al. 2017).

Neural Substrates for other Types of Spatial Perception

Research investigating other properties of spatial representation suggests that there are various regions involved apart from the hippocampal and parahippocampal formation. Mental rotation of 3D-objects seems to evoke activation in different regions, e.g. IPS, (pre-)motor areas and the inferior temporal gyrus (ITG) (Jordan et al. 2002). Involvement of the posterior parietal cortex (PPC) and IPS has also been found when participants judged the length of a line or the position of an array of dots (Dormal et al. 2009). This is further backed up by studies showing attenuated spatial processing performance when deactivating the IPS and PPC via TMS (Bjoertomt et al. 2002, Muggleton et al. 2006, Dormal et al. 2012a). In a study investigating topographic mapping of object size, Harvey et al. (Harvey et al. 2015) pinpointed the IPS in both hemispheres as key locations. The same areas have been identified for processing of numerosity, but with different properties of mapping. In a review about the functional organization of the IPS (Grefkes et al. 2005), the authors attribute visuospatial functions (e.g. spatial coordinates of objects) to the IPS.

⁴ Since their discovery, grid cells have been found in other brain areas, such as the pre- and parasubiculum (Boccarda et al. 2010).

4.3.3 Neural Code for Numerical Cognition

Similar to the perception of spatial cues, various studies have discovered the IPS as a key region for the representation of number. Eger et al. (Eger et al. 2003) presented participants with numbers, letters and colour in a visual and auditory manner while undergoing fMRI scanning. When contrasted against letters and colours, number singularly activated a horizontal region of the IPS bilaterally. A study by Piazza et al. (Piazza et al. 2004) confirmed these findings, as they found activation in the IPS in both hemispheres when participants were presented with sets of items with varying number. Furthermore, the role of the IPS in numerical processing is strengthened by a study comparing discrete numerical stimuli with analogue judgements in space or time (Castelli et al. 2006). Hereby, both the IPS and the parietal-occipital transition zone were significantly more active while processing discrete stimuli. In children with developmental dyscalculia, the right IPS was activated less when tasked with numerical judgements compared to healthy controls (Mussolin et al. 2009). Studies that employ TMS to selectively deactivate the left (Cappelletti et al. 2007) or the right (Dormal et al. 2012a) IPS could show an increased error rate when judging the amount of dots in an array.

4.3.4 Keeping Track of Time

In comparison with space and number, scientific evidence for neural substrates supporting temporal processing is even more ambiguous. This becomes evident on a very basic level: there is still no consensus of whether there are dedicated neural populations responsible to tell time or if timing is an intrinsic ability of various different neural circuits (Dehaene et al. 2011).

Experimental and theoretical studies could show an intrinsic sensory timing capability of neural circuitry, but fail to provide a general model that explains both temporal control of behaviour and is neurobiological feasible (Dehaene et al. 2011). In an attempt to provide a model for the timing of motor tasks, Buonomano et al. (Buonomano et al. 2010) proposed a population clock framework, in which the time-varying activity of individual neurons in a population support temporal judgements. In principal, population clocks could be formed by various different groups of neurons and thus, do not favour dedicated areas in the brain solely responsible for keeping track of time. In line with this notion, Buzsáki et al. (Buzsáki et al. 2018) argue that neuronal populations previously suggested to singularly compute time always contain other types of behaviourally relevant information in their firing characteristics.

Other studies have pointed out distinct neural populations which facilitate the processing of temporal information. One area of interest are the basal ganglia and their cortical connections (Matell et al. 2004, Meck et al. 2008, Buhusi et al. 2009). For example, Meck et al. (Meck et al. 2008) put forth the idea that striatal medium spiny neurons act as integrators for specific beat patterns of cortical oscillations. This is backed up by a review compiling data from functional neuroimaging, neuropsychological, psychopharmacological as well as lesion studies in humans (Coull et al. 2011). Here, firing rates of striatal areas interconnected with the frontal cortex are suggested as a neurophysiological substrate for processing

time. Other areas that have been proposed as possible loci for temporal perception include the right IPS and IFG (Hayashi et al. 2013), precentral, middle and superior frontal gyri (Dormal et al. 2012b) as well as the insula (Lewis et al. 2003).

4.3.5 Neural Evidence for a Common Magnitude System

As introduced in section 4.3.2, specialized cells inside the hippocampus facilitate the perception of space when navigating an environment. Since the discovery of place cells, research could show that they do not only encode spatial location, but also convey temporal information⁵ (Hampson et al. 1993, Pastalkova et al. 2008). In his review, Eichenbaum (Eichenbaum 2014) proposed the existence of *time cells*, which are identical in their properties to place cells and fire at specific moments in a temporally structured period. This is in line with the theory that the HC plays an intricate part in the spatial-temporal organization of episodic memory and emphasizes how closely time and space processing may be intertwined (Schiller et al. 2015, Eichenbaum 2017, Banquet et al. 2020). Buzsáki et al. (Buzsáki et al. 2018) have argued that the anatomical organization of the HC does not support the selective perception of time, space and other modalities. Instead, they theorize the HC to be a general, modality independent, theta-rhythm sequence generator which governs the computation of all incoming signals to the HC.

Outside the context of hippocampus-dependent perception of time, space and number, functional imaging experiments predominantly examined two dimensions simultaneously and point towards the IPS as a possible hub for a common magnitude system. Fias et al. (Fias et al. 2003) obtained significant activation in the left IPS while participants compared two digit numbers, line length and angles in a pairwise manner. The right superior parietal lobe, including the IPS, was conjointly active in adults when comparing non-symbolic number and spatial discrimination (Kaufmann et al. 2008). Results by Dormal et al. (Dormal et al. 2012b) further emphasize the role of the IPS, as they found shared activation in the right IPS as well as pre-central, middle, and superior frontal gyri when participants judged duration and numerosity of stimuli. In a paradigm investigating all three modalities, participants had to judge temporal intervals, numerosity and line length, Skagerlund et al. (Skagerlund et al. 2016) identified the right hemispheric IPS, insula, premotor cortex and IFG as key components of a common magnitude processing system. The authors hypothesize the IPS to be the locus for processing cardinal properties of magnitude, which then streams to the insula (salience detection and awareness), premotor cortex (action planning) and IFG (categorical decision-based representations).

⁵ Other experiments suggest that place cells might be tuned to information from various different sensory modalities, such as olfactory cues. For the sake of simplicity, only the time dimension is considered here. The interested reader is referred to (Dehaene et al. 2011) for an excellent review.

PART II:

METHODS AND MATERIALS

5. Experimental Design

5.1 Participants

Twenty-five right-handed participants were recruited from the local community (14 females, mean age 24.9 years, ranging from 20 to 41). Participants received monetary compensation and gave written informed consent to the experimental protocol, which was approved by the local ethics committee.

5.2 Virtual Environment

The virtual environment was comprised of a narrow concrete path enclosed by a grass texture to the left and right (see Figure 5.1). The coarse texture of both surfaces provided optic flow, but did not present any salient landmarks that could potentially be utilized to judge the travelled distance. The participant's viewport was placed in the middle of the path at a virtual height of 1.8 m, with a field-of-view of 60° and looking straight forward (yaw, pitch and roll all equaled 0°). At a distance of 1 m, grey fog with a linear increase in density started to obscure the view distance. Additionally, white dots appeared and disappeared constantly at random positions on the floor around the participant. Their individual lifetime was set to 100 ms.

Inside the MR scanner, the stimulus image was presented via a mirror (size 302 mm x 170 mm), placed at a distance of 350 mm from the participant's head. The image was projected onto the mirror with a resolution of 1920 x 1080 px and a refresh rate of 60 Hz. Due to visual obstruction by the head coil, the bottom 300 px of the screen were not visible, so the effective visible area was approximately 1920 x 780 px (1140 x 780 px binocularly visible).

5.3 Task and Stimulus Description

During fMRI scanning, participants performed a magnitude accumulation task in each trial with respect to either **time**, **space** and **numerosity**, while being passively moved through the virtual environment as described in 5.2. At the end of time and space trials, participants had to judge either for how long¹ or how far they were passively moved, respectively. After each numerosity trial, participants were asked to estimate the accumulative number of dots they saw on the floor². Importantly, the overall area in which the dots appeared remained

¹ Participants were instructed not to actively countdown seconds when making judgements about the duration.

² The randomly appearing and disappearing nature of the dots made it impossible to determine their exact number in order to force participants to make an approximate judgement.



Fig. 5.1: View from the participant's viewpoint during a trial. Evidently, the virtual environment does not provide any landmarks and the view distance was limited, in order to prevent the possibility to judge distance by referencing a fixed landmark. The coarse structure of the floor provided optic flow.

constant in size, as to not confound the number of dots by the physical space of the dot array.

As a control condition, **luminescence** was included as a fourth dimension. Hereby, participants did not perform a magnitude accumulation task but rather had to evaluate the greyscale value of a square that was overlaid in the middle of the screen 200 ms before the passive movement ended. Thus, this condition will be referred to as the **control** condition from here on.

Each dimension had two distinct and fixed magnitudes, which will be called *high* or *low* condition subsequently. The travelled distance was either 11.5 m or 19.7 m, the duration was either 2.8 s or 4.8 s, there were either 45 or 77 white dots present on the floor and the white square had either 16 pct or 28 pct white content. These values will be called the *reference or standard value* in the following sections. At the end of each trial, participants had to perform a two-alternative forced-choice task, indicating whether the experienced magnitude was higher or lower than a *comparison value*³ that was displayed during the response phase. Participants gave their responses with either the right index finger (less) or right middle finger (more) on a five-button MRI-compatible response box (LUMItouch, Photon Control Inc., Burnaby, Canada). The thumb button could also be pressed to indicate that the relevant dimension of the current trial was forgotten. Did the participant fail to make a response within 2 s, "*Please respond faster!*" was displayed in red letters for 500 ms on the screen. Otherwise, no feedback was provided throughout the whole experiment in order to minimize the influence of previous choices on the following trials.

It is important to note that the visual input for all trials was identical, regardless of the trial type of the current trial. So even when the participants were asked to pay attention to e.g. time, dots appeared on the floor and a white square appeared just before the end of the passive movement. Critically, this also allowed to vary the magnitude of the *irrelevant*

³ The comparison value was adjusted on a trial-by-trial bases. See section 5.6 for details on how its value was calculated.

dimensions to assess cross-dimensional influences. For example, a trial in which numerosity was the relevant dimension (participants had to judge the amount of dots on the floor), the distance traveled (irrelevant dimension) could either be long or short. The speed of the passive movement was adjusted in order to achieve times and distances that were equal to the high or low conditions.

Another critical point is that this experimental setup ensured a sequential presentation of all dimensions. The complete information needed to make a judgement was only accessible at the end of a trial and thus, all three dimensions are presented in a similar fashion (see sections 4.2.1 and 4.2.2).

5.4 Trial Structure

Each trial started with a 500 ms cue display, indicating the current relevant dimension. This was followed by an inter stimulus interval (ISI) with a randomized duration between 2 and 3 s. Afterwards, the participants were moved passively forward at a constant speed. 200 ms before the forward motion was terminated, the grey square for the control condition appeared in the middle of the screen. After another ISI with a fixed duration of 1 s, a comparison value (or rectangle in case of control trials) was displayed for 2 s (comparison phase). This was also the time during which the participants could make a response. The duration of the inter trial interval (ITI) was adjusted to equalize the total duration of the trial to either 10.8 s or 12.8 s. For null trials, a fixation cross was shown following the cue phase until the end of the response phase. The timeline of an exemplary trial is depicted on the left side of Figure 5.2.

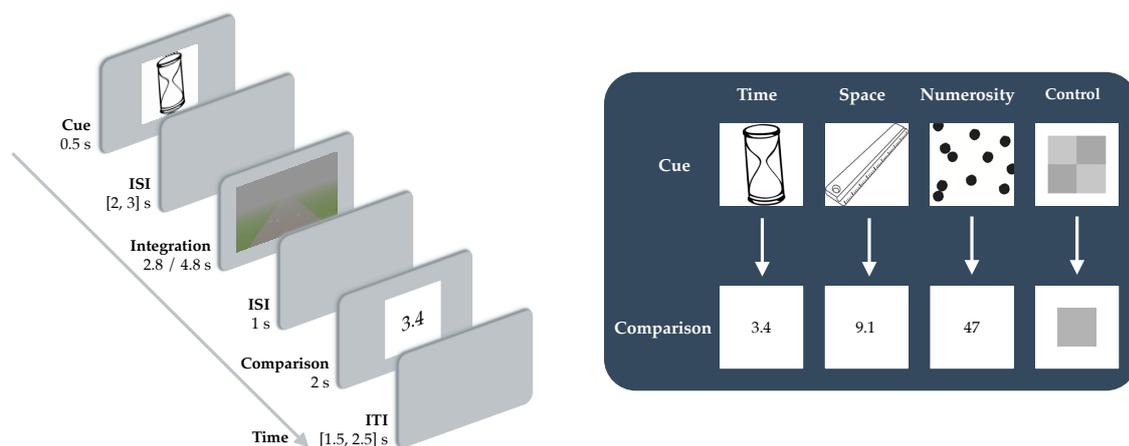


Fig. 5.2: Left: Timeline of an exemplary time trial. Please note that the onset of the grey square 200 ms before the end of the integration phase is not depicted. Right: Cue icons and corresponding exemplary comparison values for all four dimensions.

5.5 Trial Order

Trials had to be ordered in a sequence that allows for simultaneous estimation of (i) the mean difference in neural activity between the four dimensions and (ii) inter-stimulus effects of individual trials upon each other (carry-over effect) with a rapid-event design. To achieve this, a *de Bruijn* sequence, spanning 125 trials and balanced on the second order, within a continuous carry-over design (Aguirre 2007, Aguirre et al. 2011) was created, resulting in a trial order in which each of the five task types was preceded by every other task type (including itself) for an equal number of times. Additionally, each trial had a relevant and an irrelevant dimension, as described in section 5.3. Trials were counter-balanced so that the same amount of high and low conditions of irrelevant dimensions were included.

Finally, trials had to be balanced with respect to their total duration (either 10.8 s or 12.8 s). This was achieved by repeating the initial *de Bruijn* sequence four times. In the first repetition, short and long trials were inverted. During the second and third repetition, only even or odd trials were inverted, respectively.

The final result were four *de Bruijn* sequences with a length of 125 trials each, including 25 trials of each dimension (time, space, numerosity, control), as well as 25 null trials. The 500 trials were distributed among eight runs⁴ and conducted in two separate session on two different days that were not more than one month apart.

5.6 Task Difficulty

An adaptive strategy, depicted in equation 5.1, was utilized to calculate the comparison value for each trial. This was necessary to (i) equalize the task difficulty across all four dimensions, (ii) counter increased performance due to the learning effect and (iii) keep participants as engaged as possible throughout the experiment. Before each trial, the respective comparison value R_i (e.g. for time trials a value of 5.8 s) was randomly calculated according to the response (more or less) of the previous trial. The comparison value for the first trial R_0 was obtained by randomly choosing a number based upon the standard (true) value S for the current dimension.

$$\begin{aligned}
 \text{first trial: } R_0 &= \{x \in \mathbb{R} \mid S - 0.25 \cdot S \leq x \leq S + 0.25 \cdot S\} \\
 \text{response more: } R_i &= \{x \in \mathbb{R} \mid R_{i-1} \leq x \leq R_{i-1} + 0.25 \cdot R_{i-1}\} \\
 \text{response less: } R_i &= \{x \in \mathbb{R} \mid R_{i-1} - 0.25 \cdot R_{i-1} \leq x \leq R_{i-1}\}
 \end{aligned} \tag{5.1}$$

To monitor the perceived difficulty of each task type after each run, the participants had to judge the difficulty on a visual analogue scale (VAS, slider with 20 intermediate positions) ranging from easy to difficult with the help of their index and middle finger.

⁴ As this yields an unequal number of trials per run, null trials were added at the end of each run to equalize the total amount of trials per run.

6. Behavioural Data Analyses

6.1 Preprocessing

Trials with (i) a reaction time lower than 300 ms (0.18 %), (ii) in which participants indicated that they forgot the current trial type (0.28 %) or (iii) with a response later than 2 s (2.65 %) were excluded prior to any analysis. Additionally, behavioural and functional data of participant 15 was completely discarded, as he missed more than 20 % of his responses, casting overall doubt on alertness during the experiment. Participants 4, 5 and 6 confused the response buttons during control (luminescence) trials in session 1, hence this data was dropped from the analyses as well. This left a total number of 9430 trials of 24 participants for behavioural data analyses.

6.2 Reaction Time and Task Difficulty

Both reaction times and perceived task difficulty ratings for the four task types during the course of the eight runs were analysed with respective linear mixed effects models to assess whether (i) there are interdimensional differences or if (ii) there were changes during the course of the experiment. For both models, task type as well as the run number were fixed effects predictors and participant number was modelled as a random intercept (reaction times: $reaction\ time \sim dimension * run + 1 | subject$, task difficulty: $task\ difficulty \sim dimension * run + 1 | subject$). All analyses were conducted in R (R Core Team 2019), including the *nlme* package (Pinheiro et al. 2019).

6.3 Fitting Psychometric Functions

6.3.1 Input Data

Due to several reasons, fitting a PF to the behavioural data of this experiment was challenging. First, the amount of trials per stimulus intensity is usually $N \leq 10$. This is mainly due to the limited time a participant lays comfortably in the MRI scanner. Additionally, the functional data analysis strategy created additional constraints for the sequence and amount of trials (see section 5.5). Second, due to the adaptive adjustment of reference values to keep the difficulty constant, not all stimulus intensities were sampled with the same amount of trials. Some intensities, especially closer to the lower and upper boundaries, were sampled with only one trial. Erroneous responses from participants thus may have a substantial impact on the fitted PF, if the fitting algorithm is not robust against these outliers.

6.3.2 Fitting Algorithm

Due to the unique features of the behavioural data explained above, the chosen fitting algorithm must be robust against outliers, account for sparse and varying trial numbers per sampled intensity, should allow for a transparent adjustment of parameters and provide a reliable proxy for the goodness-of-fit. The Python package *psignifit 4* by (Schütt et al. 2016, Schütt 2016) is based on Bayesian inference and, compared to other alternatives¹, fulfilled all the aforementioned prerequisites best. Hence, a modified version was used to analyse the behavioural data of this experiment. Several key advantages of *psignifit* are summarized here:

1. Even for overdispersed data, *psignifit* is able to obtain credible interval estimations by integrating a *beta-binomial* model.
2. Instead of estimating the slope β of the PF, *psignifit 4* reports the width w . Due to advantages shown in section 3.2.2, this is the preferred parameter.
3. Posterior distributions are calculated by numerical integration within predefined limits, thus complicated approaches to solve the posterior equation are not needed.
4. The package provides predefined priors for each parameter, based on extensive research and experience. These priors can easily be adapted if needed.
5. Ease of use within the Python language environment.

Standard Bayesian approaches model the probability of a response only as a function of the stimulus intensity. Essentially, each trial is a Bernoulli trial and independent from the others. In reality however, this is rarely true. For example, participants might get tired over the course of the experiment or might be influenced by the choice she or he made the trial before. As stated above, participants did not get any feedback on any trial, so the influence of previous choices on the current response are minimized. Still, the assumption of a completely stable observer is violated, resulting in too narrow credible intervals for the posterior distribution (Schütt et al. 2016).

Instead of treating each trial as a Bernoulli trial with a binomial distribution, *psignifit 4* employs a beta-binomial model to account for that instability of the observer, causing overdispersion in the data. For a constant stimulus intensity x , the success probability is no longer fixed (as in binomial models), but a randomly drawn beta-distributed variable with mean $\psi(x)$. An additional parameter η ($0 \leq \eta \leq 1$) is introduced to the PF, defining the variance of success probability for a stimulus intensity x and number of trials N as follows:

$$\sigma^2 = \left(\eta^2 + \frac{1 - \eta^2}{N} \right) \psi(x)(1 - \psi(x)) \quad (6.1)$$

In the case of $\eta = 0$, equation 6.1 collapses to the variance of a standard binomial distribution $\sigma^2 = \frac{1}{n} \psi(x)(1 - \psi(x))$. For a maximally unstable observer with $\eta = 1$, the variance is independent of the trial number N with $\sigma^2 = \psi(x)(1 - \psi(x))$, i.e. increasing the number

¹ An alternative investigated was *BayesFit* (Slugocki 2019). However, it was found that its robustness to overdispersed data is not suitably high for the input data of this study.

of trials does not decrease uncertainty about the mean. For any value between $0 < \eta < 1$, this model allows to account for overdispersed data by adjusting the variance of the success probability to $1 + (N - 1)\eta^2$ times the one of the standard binomial model.

6.3.3 Priors

As described in 3.2.4, Bayesian inference requires a priori distributions for each parameters. The priors proposed by (Schütt et al. 2016) have been adjusted to reflect the estimated outcome of this experiment and are summarized in the following list. Figure 6.1 provides a graphical representation of the distributions.

- Threshold α : The prior was chosen as a uniform distribution across the range of the stimulus intensity with a cosine falloff to 0 over half of the range at both ends. This was identical to the proposition by (Schütt et al. 2016) and assumed that the threshold has an equal probability within the range of the sample stimuli and may be localized 50 % below or above the range with decreasing probability.
- Width w : The prior distribution was uniformly distributed between four times the minimal distance between two stimulus intensities (lower limit) and three times the total range of stimulus intensities (upper limit). The distribution had a cosine falloff to 0 at the lower limit and at 4 times the total range. This is different from the standard prior implemented in `psignifit` and accounted for larger width values due to the overdispersed and undersampled data of this experiment.
- The guess rate had a fixed value of $\gamma = 0$ for the two-alternative forced-choice task in this study.
- The lapse rate λ was fixed at a value of $\lambda = 0.01$. Preliminary analyses of the data showed that a variable lapse rate with a beta-distributed prior (as proposed by (Schütt et al. 2016)) provided unsatisfactory fitting results and overestimated lapse rates due to the low amount of trials at some intensities, especially near the lower and upper intensity boundaries.
- The prior for η , accounting for an unstable observer, was left unchanged and had a beta-distributed prior with $\alpha = 1$ and $\beta = 10$.

6.3.4 Parameter Estimates

For each participant, each dimension and magnitude (high or low), a separate PF was fitted to the experimental data. Once a PF had been fitted, `psignifit` provided posterior distributions, along with the MAP² and credible intervals for each parameter.

In this experiment, the threshold α is defined at the comparison value where $p(M) = 0.5$ ³

² (Schütt et al. 2016) argue that the MAP is the most appropriate estimator, as mean and median are heavily biased for the width, especially for data with a low amount of trials, such as in this study.

³ During the introduction of Psychophysics in chapter 3, values along the y-axis denoted the probability of a successful stimulus detection. In contrast, Psychometric functions in this thesis are fitted to the probability $p(M)$ of a participant correctly identifying a comparison value as being *more* than the just experienced value during the trial.

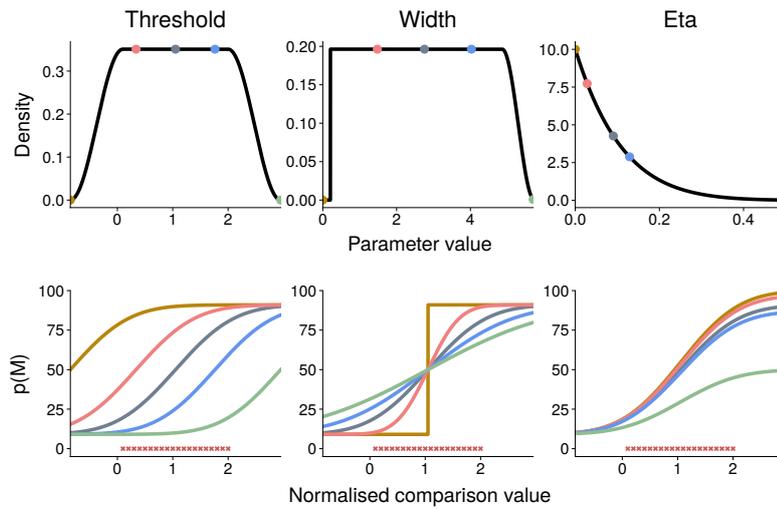


Fig. 6.1: Prior distributions for evenly spaced input data at comparison values in the range of $x = [0.1, 2]$ for the threshold, width and observer stability η . The guess and lapse rate priors are not shown, as these are fixed at their respective values. The top row shows the cumulative density functions for each prior. Below, Psychometric functions corresponding to 0 %, 25 %, 50 %, 75 % and 100 % quartiles of the prior are plotted. The label $p(M)$ denotes the probability that a participant perceives the comparison value (plotted on the x-axis) as greater than the actual value (e.g. the time travelled was longer than the comparison value presented).

and was identical to the point of subjective equality, i.e. participants perceived the shown comparison values during the response phase as equal to the reference value of the trial she or he just experienced. For comparison values above the threshold, participants become increasingly more likely to perceive the comparison value as greater (more) than the reference value and, thus, $p(M)$ converges asymptotically towards 1. Conversely, $p(M)$ asymptotically falls towards 0 for increasingly smaller comparison values.

The width is defined here as the difference in comparison values, at which the unscaled PF reaches 0.2 (lower limit) and 0.8 (upper limit). It is interpreted as a measure of how precise the participant was in her or his judgements (section 6.5).

For both threshold and width, results were classified as outliers if they exceeded $Q_{75} + 1.5 \cdot IQR$ or fell below $Q_{25} - 1.5 \cdot IQR$. Q_{25} and Q_{75} are the 25th and 75th percentile, respectively. IQR is the interquartile-range, defined as $IQR = Q_{75} - Q_{25}$.

6.4 Magnitude Differentiation

In a first sanity check of the behavioural data, it was tested if participants could distinguish between large and small magnitudes of each trial type by comparing the PSEs (thresholds) between large and small magnitudes. Before any statistical analysis, the data was checked for normality by a combined test of skew and kurtosis (K-S test) with Python's *SciPy* package (Virtanen et al. 2020). As all of the distributions were found to be not normally distributed, the results of one-tailed paired t-tests were verified by a Wilcoxon signed-rank test and corrected for multiple comparisons (FWE = 0.05).

Because the scales of physical stimuli differ substantially between dimension (e.g. time varied in the order of 2 - 10 seconds, the number of dots between 20 and 50), the comparison values of all trials are normalized to the mean of the two reference values for each dimension, in order to allow for simultaneous plotting of all dimensions. For example, the duration of a time trial was either 2.8 s or 4.8 s, so all comparison values for time trials were normalized to 3.8 s.

6.5 Judgement Precision

Participant's judgement precision was characterized by the width of the calculated PF as defined in 6.3.4. Participants that had a higher precision in their magnitude perception had a smaller width compared to participants with less precise judgements.

The data analysis assessed whether there were crossdimensional correlations in judgement precision between all possible pairs of dimensions, as this provided insight on how similar their perception might be. For this, Spearman rank correlation coefficients of the widths of all participants were obtained for each possible pair of dimensions.

In contrast to the magnitude differentiation sanity check, a different normalization was applied to the data: all trials were normalized to their respective standard values. This enabled a conjoint analysis of trials with a high and low standard value. For example, all space trials with a distance of 11.5 m or 19.7 m were normalized to 11.5 m or 19.7 m, respectively.

6.6 Cross-Dimensional Interference

In order to assess cross-dimensional influences, the behavioural data of one trial type was split into two subsets, based on the magnitude of the irrelevant dimension. This is possible, because the trial order was designed in a way to counter-balance the amount of trials in which the irrelevant dimension was either in a high or low condition (see sections 5.3 and 5.5). Practically, this meant that for e.g. time trials, separate PFs were fitted for trials in which the amount of dots on the floor was either high (77) or low (39), to assess whether the amount of dots on the floor influenced the perception of time. Vice versa, trials in which participants judged the amount of dots on the floor were split with respect to the duration of the passive movement (either 2.8 s or 4.8 s). For each possible pair of relevant-irrelevant dimension and each participant, the thresholds were obtained and their mean differences compared against 0 by a Wilcoxon signed-rank test with a correction for multiple comparisons (FWE = 0.05). Positive values indicate congruent interference (i.e., larger values of the irrelevant dimension lead to an overestimation of the relevant dimension).

Analogue to the interdimensional judgement precision analysis, comparison values were normalized in the same fashion to allow the analysis of cross-dimensional inference.

7. Neuroimaging Data Analyses

7.1 Data Acquisition

MRI data was acquired on a 3 T Magnetom Prisma-fit Syngo MR D13D Siemens Scanner (Siemens Healthcare GmbH, Erlangen, Germany) with a 64-channel phased array head coil. For anatomical images, both T1- and T2-weighted images were recorded. In each session, double-echo gradient-echo (GRE) field maps were acquired¹, allowing for susceptibility distortion correction of the functional images. Scanning parameters of the anatomical and functional sequences according to (Inglis 2015) are summarized in table A.1.

7.2 Data Quality Control

Prior to any analysis, both structural and functional imaging data were visually inspected. Additionally, image quality metrics (IQMs) for both were calculated using the software package *MRIQC* 0.14.2 (Esteban et al. 2017). This allowed for a quick assessment of various data quality parameters such as head movement, temporal SNR, ghosting or artefact detection. As of today, no comprehensive study has published guidelines on which combination of parameters are best suited to detect and exclude outliers. However, recent projects try correlate quality ratings by experts with IQMs (Esteban et al. 2019a). Shehzad et al. (Shehzad et al. 2015) investigated quality assessments of resting-state fMRI data from various sites and provides some advice on test-retest reliability of IQMs as well as correlation between those. Based upon these assessments, a set of IQMs was selected and they are briefly introduced in this section. Please refer to (Esteban et al. 2017, Esteban et al. 2019a) for a detailed description.

Analogous to the behavioural outlier criterion, quality parameter values for a specific participant above $Q_{75} + 1.5 \cdot IQR$ or below $Q_{25} - 1.5 \cdot IQR$ were considered to be suspect and triggered a detailed examination of the functional data. Furthermore, it was decided to exclude functional runs if the percentage of volumes exceeding a framewise displacement (FD) of 1 mm was above 10 %.

Spatial Information

- **EFC:** The entropy-focus criterion gives an indication of ghosting and blurring induced by head motion (Atkinson et al. 1997). In very basic terms, the entropy criterion favours alterations to the data which maximise the number of voxels containing no information

¹ Echo times were 4.92 ms and 7.38 ms.

and therefore have a value close to 0. Participant motion has the opposite effect, as neighbouring voxels at the borders will have increased values. EFC may therefore be used to detect and correct motion-induced image alterations. Lower values are better.

Temporal Information

- **tSNR:** Temporal SNR calculated by the average BOLD signal across time divided by the temporal standard deviation map. It gives insight into the temporal variability (Welvaert et al. 2013). Higher values are better.

Artefact Detection

- **GSR:** Ghost-to-signal ration estimates the mean signal in the areas prone to Nyquist $N/2$ ghosting artefacts along the possible phase-encoding directions relative to the mean signal within the voxels of the brain mask. Fast echo-planar imaging sequences are especially susceptible to Nyquist ghost artefacts, as the rapidly changing gradients may induce eddy currents in the RF coils and magnetic housing (Giannelli et al. 2010). It is therefore sensible to screen the data for this type of artefact. Lower values are better.
- **FD:** Framewise displacement is a measure for instantaneous head-motion from one volume to the next (Power et al. 2012). Typically, either the mean FD or the percentage of volumes above a certain threshold (1 mm in this analysis) are reported.

7.3 Preprocessing

Anatomical and functional images were preprocessed using *fMRIPrep* 1.1.1 (Esteban et al. 2019b). For the structural preprocessing, T1/T2 weighted (T1w/T2w) images were corrected for intensity non-uniformity with ANTs' `N4BiasFieldCorrection` v2.1.0 (Tustison et al. 2010). Afterwards, they were skull-stripped using ANTs' `antsBrainExtraction.sh` v2.1.0 (OASIS template, (Marcus et al. 2007)) to create an initial binary mask of the brain. After brain mask computation, FSL `FAST` v5.0.9 (Zhang et al. 2001) was used to perform brain tissue segmentation (CBF, white matter, gray matter). Brain surface reconstruction was performed by `recon-all` from `FreeSurfer` v6.0.1 (Fischl 2012). The initial brain mask was refined with a *fMRIPrep* custom variation of the method to reconcile ANTs-derived and `FreeSurfer`-derived segmentations of the cortical gray-matter (Klein et al. 2017). In the last step, spatial normalisation to the MNI152NLin2009cAsym template space (Fonov et al. 2009) was performed using `antsRegistration`. This entails a multiscale, mutual-information based, non-linear registration scheme. In order to obtain individual anatomical region-of-interest (ROI) masks, a cortical reconstruction was run with `FreeSurfer` 6.0. The resulting label files were converted to binary masks.

During the initial step of BOLD image preprocessing, a reference image was created in order to calculate a brain mask for the BOLD signal. This EPI reference image is then passed to FSL's `mcflirt` v5.0.9 (Jenkinson et al. 2002) in order to estimate head motion (6 parameters - three for rotation and three for translation) parameters. Slice time correction is performed using AFNI v16.2.07 `3dTShift` function (Cox 1996), realigning each slice to the middle of

each TR. Next, spatial distortion due to B0 field inhomogeneities were accounted for by susceptibility distortion correction (SDC) using an implementation of the TOPUP technique (Andersson et al. 2003) using function `3qwarp` in AFNI v16.2.07. Hereby, the field inhomogeneity can be mapped by measuring the phase evolution in time between two close GRE acquisitions (Hutton et al. 2002). Finally, the BOLD reference image is aligned to the T1w image using linear boundary-based registration with 9 DOF implemented in FSL `flirt`. To map the EPI image to the MNI152NLin2009cAsym template space, transform calculated upstream (head-motion estimation, SDC, EPI to T1w registration, T1w-to-MNI transform) workflows are concatenated. These transforms are applied all at once with one Lanczos interpolation step.

7.4 Univariate Analysis

7.4.1 First Level Analysis

Design and implementation of the statistical analysis was done using the *nipype* Python framework (Gorgolewski et al. 2011), which allows for a flexible and accessible implementation of various different neuroscientific software packages in one workflow. Figure 7.1a shows the corresponding workflow for the first level analysis. With few exceptions, SPM12 v7487 MATLAB 2016b (The MathWorks, Natick, 2016) was utilized for model generation and contrast estimation. The functional images were smoothed using a 6 mm FWHM gaussian filter and masked with an individual brain mask obtained from fMRIPrep. The data was high-pass filtered at 128 Hz.

During first level univariate analyses, the integration (passive movement) phase was modelled as delta functions convolved with the hemodynamic response function (HRF) as implemented in SPM12, excluding temporal or spatial derivatives. Additionally, (i) the cue onset as well as (ii) the comparison phase until a response was given were included as regressors of interest and convolved with the HRF (see Figure 5.2 for a reference of the trial structure). It was decided to leave out the actual response (button press), as this regressor showed high correlation with the one modelling the comparison phase. As nuisance regressors, six head motion parameters², framewise displacement (FD) as well as the PCA-based anatomical *aCompCor* (Behzadi et al. 2007, Muschelli et al. 2014) confounds were included. Figure 7.1b shows an exemplary design matrix for one run. Significance threshold for univariate analyses was set to a voxel-wise FWE ($p < 0.05$) and a cluster threshold of 10 voxels. Contrasts were defined based upon the aims of this study, introduced in section 1. Table 7.1 summarizes the conditions and contrasts. In addition to the independent contrasts, a conjunction analysis of the three accumulation trials (time, space, numerosity) contrasted against the control condition was performed in order to identify regions that become jointly significant.

² Three for translation and three for rotation.

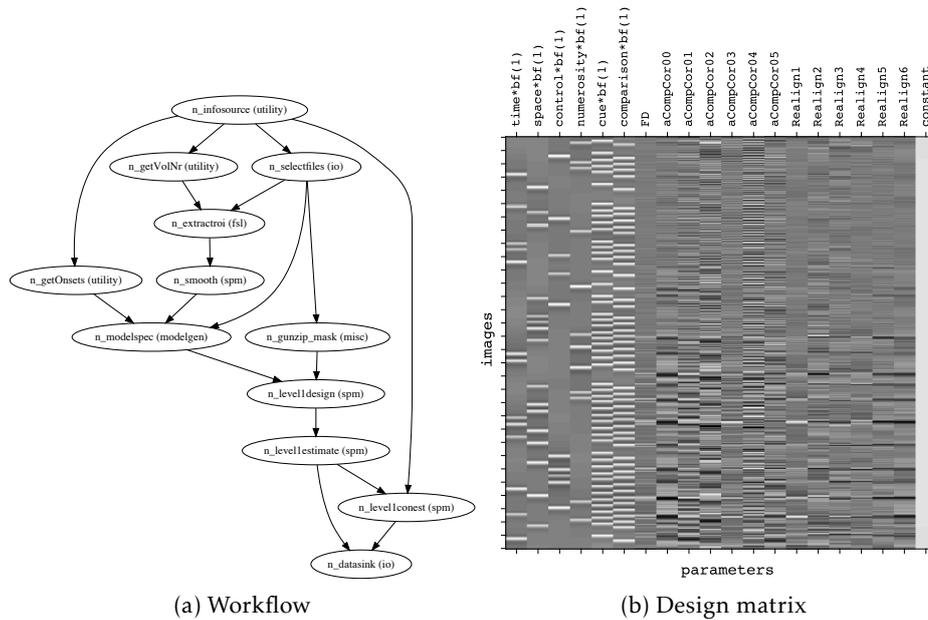


Fig. 7.1: (a): Simplified graph of the first level analysis workflow. Inputs and outputs are depicted with arrows. A detailed schematic for the workflow is provided in Figure A.1. (b): Exemplary first level design matrix for one run. The first six regressors are convolved with the HRF and represent the four trial types, cue presentation and response phase (until a response was given). Nuisance regressors are FD, six anatomical aCompCor confounds, calculated by fMRIPrep, as well as six motion parameters.

Table 7.1: List of T-contrasts estimated during the univariate analyses.

Nr.	Name	time	space	numerosity	control	cue	comparison
1	Accumulation	1/3	1/3	1/3	0	0	0
2	Accumulation > control	1/3	1/3	1/3	-1	0	0
3	Time > control	1	0	0	-1	0	0
4	Space > control	0	1	0	-1	0	0
5	Numerosity > control	0	0	1	-1	0	0
6	Conjunction	(Time > control) \cap (Space > control) \cap (Numerosity > control)					

7.4.2 Second Level Analysis

Analogue to the first level analysis, group level analyses were computed with SPM12 v7487 within the nipype Python framework, similar to the first level analysis. A classical inference on the second level with a one sample t-tests were performed on all contrasts. Thresholding was performed with a voxel-wise FWE ($p < 0.05$) and with a minimum cluster size of 10 voxels. As reported in (Eklund et al. 2016), this approach tends to be too conservative, but avoids the inflated false-positive of the clusterwise FWE. A graph of the nipype workflow is shown in the appendix (Figure A.2).

7.5 Multivariate Pattern Analysis

7.5.1 Introduction

Multivariate Pattern Analysis

The goal of this experiment was not only to gain insight which magnitude perception potentially activates a specific neuronal population (i.e. associating brain regions with functions), but also to decode distinct patterns of activity corresponding to processing time, space and number. Multivariate pattern analysis (MVPA), first employed by Haxby et al. (Haxby et al. 2001), is a technique to enable to classification of such activity patterns. This subsections will provide a superficial introduction into the methodology, for a more detailed introduction into the topic, please refer to (Pereira et al. 2009) and (Haxby et al. 2014).

In a first step during MVPA, the BOLD values of each voxel in a brain are vectorized into a high-dimensional representational space. The dimensionality of that space is defined by the amount of voxels and time-points (trials) of the functional imaging data. In the case of this study, the whole brain contains about 64290 voxels and 100 trials for each condition. Numerically, this is done by creating a $N \times M$ matrix, in which N denotes the number of observations (i.e. trials corresponding to certain conditions, also called labels) and M the amount of voxels. For a binary decision problem, in which e.g. time trials are classified against space trials, the matrix would contain 200 rows (100 trials for each condition) and 64290 columns. Each field of that matrix would contain the BOLD value of the specific voxel at a specific time point.

In a second step, the available data is now divided into a training and a testing dataset. Based on the training data, a classifier is now employed to find a decision surface that defines the boundaries between sectors in the representational space. The most popular option for a classifier is a support vector machine that optimizes the position by maximising the distances from the surface to the closest vector (i.e. the support vectors). In a binary classification problem such as in this study, the decision surface would take on the shape of a plane.

Once training is complete, the test data is now presented to the classifier and each trial is classified according to its position in regard to the decision plane. Classification performance is commonly assessed by counting the amount of correctly classified trials, but there are many other options for scoring parameters (see (Kelleher et al. 2015)).

Typically, splitting the available data into a training and testing set is not only done once (e.g. by splitting the data in half), but many times in a procedure called *cross-validation* (CV) with the aim of maximizing the amount of independent training and testing data. Hereby, only a fraction of the data is left out for testing (e.g. 20 %). The classification is then repeated e.g. 5 times (5-fold CV), each with a different 20% of the data serving as the test subset. The final result is obtained by calculating the average of the 5-fold CV classifier performances.

Searchlight Analyses

Because BOLD responses are mapped into a high-dimensional representational space in MVPA, their spatial and temporal information is disregarded. An approach how to re-

tain information about the anatomical localization of the representational space is called *information-based functional brain mapping* (Kriegeskorte et al. 2006, Chen et al. 2011). Among fMRI scientists, it is often called *searchlight decoding*, as its procedure is not unlike shining a searchlight through the whole brain: in succession, MVPA is performed based on multiple small subsets of voxels. The whole brain is traversed by iterating through every voxel and defining an ROI with the neighbouring voxels, based on a predefined radius. The output of such a searchlight analysis is a brain map in which each voxel's value is the cross-validated classification accuracy obtained by performing MVPA with only the subset of voxels surrounding the central voxel. In order to infer information about statistical significance, the accuracy maps have to be thresholded and corrected for multiple comparisons. The strategy employed in this thesis is described in section 7.5.3.

7.5.2 Sample Generation

Before running a MVPA, the raw BOLD signal has to be transformed into samples that are suited for the particular classification problem at hand. The most straightforward method is to use each volume as an individual sample for the classifier, but one could also e.g. average multiple volumes to obtain a less noisy sample. Approaches based on raw data have two main limitations: (i) they potentially ignore vital information as the HRF is not modelled and (ii) the BOLD signal of nearby trials inevitably overlaps, making the signal that is unique to one specific trial much harder to identify.

One solution to the first problem is to obtain beta estimates for each trial using one general linear model (GLM), in which each trial is modelled as a separate regressor (Rissman et al. 2004). Consequently, the design matrix has as many columns as trials. Mumford et al. (Mumford et al. 2012) called this the *least-squares-all* (LS-A) method. However, this approach is not suited for rapid event-related designs, as trial-specific regressors of closely timed trials may be highly correlated. To solve this, Mumford et al. (Mumford et al. 2012) propose the so called *least-squares-separate* (LS-S) method. Hereby, a separate GLM for each trial is created that predicts two timecourses: one for the event (trial) of interest and one for all other occurring events. The resulting design matrix has two columns plus additional nuisance regressors. A disadvantage of this method is a high computational load, as there have to be as many GLMs estimated as there are trials. In a follow-up paper (Turner et al. 2012), the authors modify the LS-S approach by modelling each type of event in a separate regressor instead of summarising them into one. This *LS2* method improves (i) overall classification performance and (ii) the ability to disentangle the signal of temporally adjacent events compared to LS-S and was therefore chosen in this thesis to create input samples for the MVPA. See Figure 7.2 for a graphical representation of the three methods. A detailed workflow map for the trialwise beta estimation is shown in the appendix in Figure A.3. In a last step, the trialwise beta maps were standardized (i.e. removal of the mean and scaled to unit variance).

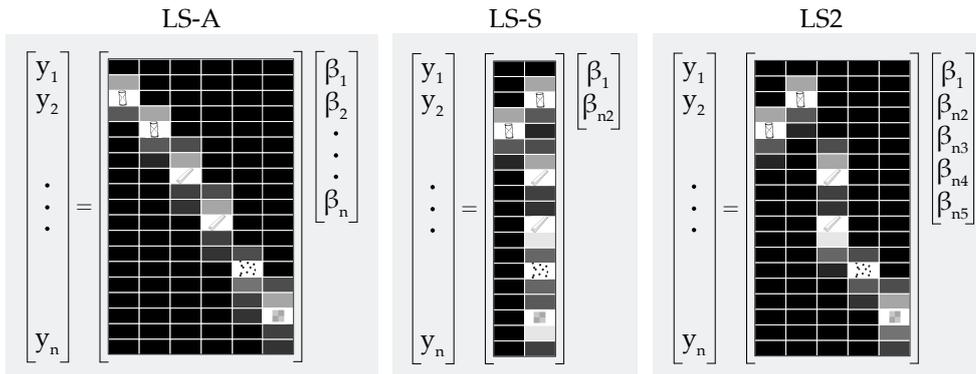


Fig. 7.2: Schematic representation of different beta-series estimation approaches. Note that additional nuisance regressors (e.g. motion parameters) are omitted. Left (LS-A): one GLM is modelled in which each trial is a separate regressor. Middle (LS-S): for each trial, a new GLM is created (in this example for the second trial). This trial is included as the first regressor and all other trials are summarized in a second nuisance regressor (β_{n2}). Right (LS2): similar to LS-S, each trial has a separate GLM. The first regressor estimates the timecourse of the current trial. All other occurrences of the other trial types (including the trial type of the current trial) are modelled in separate regressors ($\beta_{n2} - \beta_{n5}$). Adapted after (Turner et al. 2012).

7.5.3 Inter-Subject Pattern Analysis

Statistical inference on the group level in whole brain MVPA commonly employs t-tests on participant's individual decoding accuracies. Warranted criticism has been raised against this approach, mainly due to (i) low number of observations and non-uniformity of statistical distribution of classification accuracies (Olivetti et al. 2012, Stelzer et al. 2013, Allefeld et al. 2016), (ii) the non-directional nature of the identified group information (Gilron et al. 2016) and (iii) biases in the results due to confounds, e.g. time on task or individual differences (Todd et al. 2013). In their publication, Allefeld et al. (Allefeld et al. 2016) bring forward another important point: because the classification performance of a participant can never be below chance level, the t-test tests a null-hypothesis that there is no information in any subject on the group level. Consequently, the results may show that there are participants with an effect, but not whether this generalises to the population. Crucially, Stelzer et al. (Stelzer et al. 2013) could show that t-tests lead to high levels of false-positivity. In a recent paper, Wang et al. (Wang et al. 2020) reviewed the *inter-subject pattern analysis* (ISPA) approach. Instead of splitting the data of one subject into a train and test dataset, the data is split on the group level in training and testing subjects. Cross-validation is done via a *leave-one-subject-out* (LOSO) algorithm, in which the accuracy is repeatedly tested on the left out participant. As a results, there as many folds as there are participants and each fold produces a whole-brain accuracy map. In a second step, the proposed approach uses non-parametric permutations tests to asses significance on the group-level (T. E. Nichols et al. 2002). When compared to the traditional group level MVPA methods, ISPA provides various key advantages and was therefore chosen as the analysis method for this thesis:

- ISPA offers straight forward interpretation of the results. A significant effect detected by ISPA may directly be translated to the population level (Kragel et al. 2018).
- Since no permutations of labels on the participant level are employed for significance testing, computational requirements remain reasonable.
- Compared to decoding on the participant level, the amount of samples available for training is larger when ISPA decoding is utilized³. This in turn allows ISPA to detect smaller multivariate effects (Wang et al. 2020).

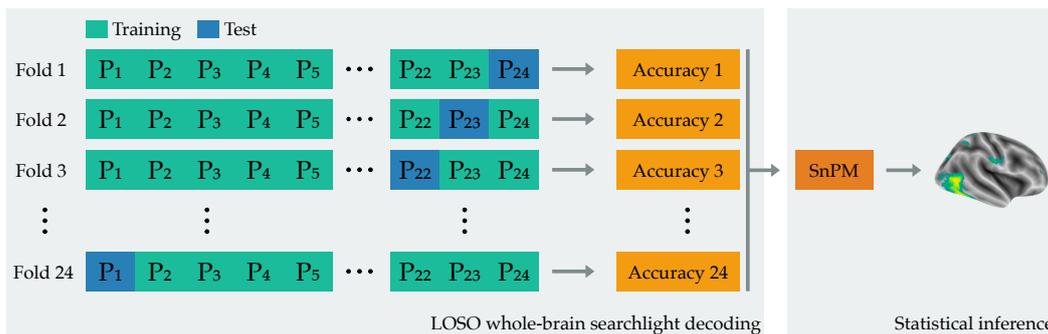


Fig. 7.3: Schematic representation of the ISPA analysis algorithm for one pair of conditions (e.g. time vs. space). In a 24-fold LOSO cross-validation, trial-wise beta maps of 23 participants are combined into a training set. The classifier is then tested on the data from the 24th participant. Each fold results in a whole-brain accuracy map, in which each voxel carries the decoding accuracy of the specific searchlight run. The 24 accuracy maps are then analysed for their significance via non-parametric permutations, as implemented in SnPM.

Figure 7.3 depicts a schematic representation of the ISPA decoding algorithm. Pairwise whole brain searchlight decoding with a LOSO cross-validation and a radius of 6 mm was implemented in Python using the packages *nilearn* 0.5.2 (Abraham et al. 2014) and *scikit-learn* 0.21.3 (Pedregosa et al. 2011). As an estimator, a logistic regression classifier was chosen due to (i) its computational efficiency and (ii) readily interpretable results (Ryali et al. 2010). Table 7.2 summarizes the relevant parameters. The optimization problem was solved by using the *lbfgs-solver*⁴, as it is memory conservative and relatively fast for the amount of data points in this analysis.

As described in section 5.3, participants could indicate that they forgot the dimension of interest for the current trial with the thumb button during the response phase. These trials were excluded from the classification samples. To prevent an inflated performance estimates for the classifier due to different amount of samples, a *balanced accuracy* scoring was employed, weighting samples according to the inverse prevalence of its true class. If the participant did not press the thumb button for any trials and the sample numbers are thus balanced, the reported classification score is equal to the conventional accuracy (i.e. the

³ On the participant level, the number of samples (trials) is usually in the order of dozens to a few hundred. In ISPA, this number is multiplied by the amount of participants in the study, greatly increasing the total amount of samples available for training.

⁴ Limited-memory Broyden-Fletcher-Goldfarb-Shanno approximates the second derivative matrix updates with gradient evaluations (Zhu et al. 1997).

number of correctly classified samples divided by the total number of predictions). Other parameters have been left at their default value.

In a second step, the significance at the group level of 24 accuracy maps was tested with the implementation of SnPM toolbox (T. Nichols et al. 2001) with 1000 permutations, a voxel-wise significance thresholding of $p < 0.05$ (FWE corrected) and a minimum cluster size of 20 voxels. In a last step, the obtained statistical maps were converted to binary images and the average classification accuracy maps from 24 folds were masked using these maps. The final result is a map of average classification accuracy that shows only voxels that remain after significance testing on the group level.

Table 7.2: Parameters used for the logistic regression classifier.

Parameter	Setting
penalty	l_2
C	0.1
solver	lbfgs
scoring	balanced_accuracy
tol	$1e^{-4}$
class_weight	none
decision_function_shape	one-versus-rest (ovr)

PART III: RESULTS

8. Behavioural Results

8.1 Manipulation Checks

8.1.1 Reaction Times

The average reaction times across eight runs for the four dimensions are depicted in Figure 8.1. ANOVA of the fitted linear mixed effects model revealed two main effects: (i) participants responded faster during the course of experiment and (ii) response times differed between dimensions significantly. No interaction was found between run and dimension, indicating that the change in response times across eight runs did not differ between task dimensions. Table 8.1 summarizes the results.

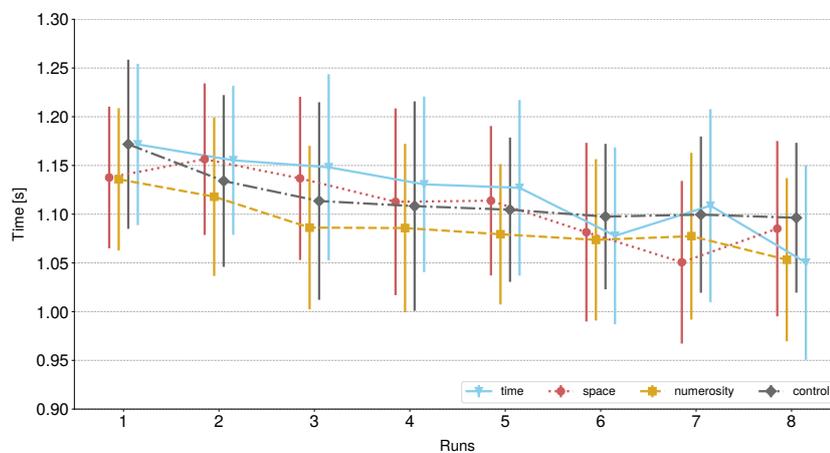


Fig. 8.1: Reaction times in seconds across all eight runs. The markers are jittered along the x-axis for better visibility. Error bars represent 95 % confidence intervals across participants.

In order to test for which dimensions the reaction times differed significantly, a post-hoc test was done using *Tukey's HSD* to control for family-wise error rate at $p = 0.05$. Hereby, only the reaction times for numerosity and time showed a significant difference ($\bar{x} = 0.03$ s, $p = 0.005$).

Table 8.1: Results of the ANOVA of the fitted linear mixed effects models for reaction times. Task dimension and run number were included as fixed predictors, subject number was modelled as a random intercept ($reaction\ time \sim dimension * run + 1 | subject$).

	DF	denDF	F-value	p-value
intercept	1	725	835.40	<0.001
dimension	3	725	3.6175	0.013
run	1	725	50.72	<0.001
dimension:run	3	725	0.5442	0.652

8.1.2 Perceived Task Difficulty

As described in section 5.6, the comparison values were adjusted by an adaptive algorithm with the intention to maintain an equal difficulty level across runs and across different dimensions. In order to assess task difficulty, participants had to judge the difficulty of the four task types at the end of each run. Figure 8.2 shows the results.

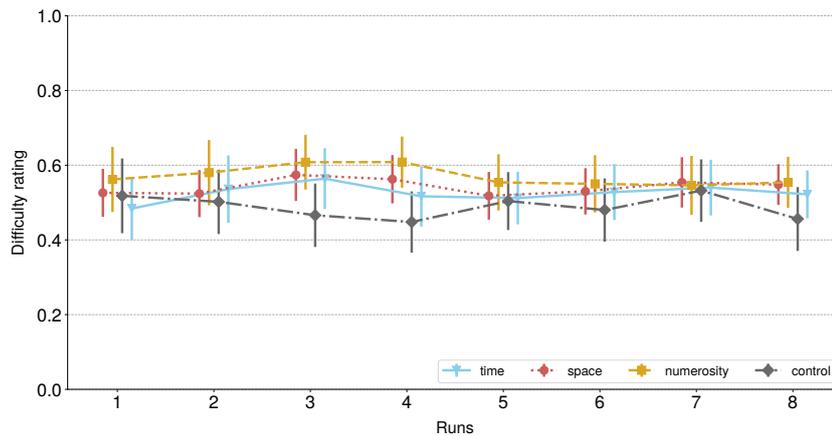


Fig. 8.2: Perceived task difficulty for each dimension and each run. Participants made their response on a VAS with 20 intermediate positions ranging from easy (0) to difficult (1). The markers are jittered along the x-axis for better visibility. Error bars represent 95 % confidence intervals across participants.

Analogue to the reaction times, an ANOVA of the fitted linear mixed effects model was performed. The results show that perceived task difficulty did not change during the course of the experiment, as intended by the adaptive algorithm employed (see Table 8.2). However, there was a main effect for dimension, indicating that difficulty among the different trial types was significantly different. A pairwise post-hoc *Tukey's HSD* test revealed a significant difference in perceived difficulty between space and control ($\bar{x} = 0.07, p = 0.003$), numerosity and control ($\bar{x} = 0.09, p < 0.001$) as well as time and control ($\bar{x} = 0.05, p = 0.03$). Difficulty among the three accumulation dimensions did not differ significantly.

Table 8.2: Results of the ANOVA of the fitted linear mixed effects models for perceived task difficulty. Task dimension and run number were included as fixed predictors, subject number was modelled as a random intercept ($task\ difficulty \sim dimension * run + 1 | subject$).

	DF	denDF	F-value	p-value
intercept	1	721	1112.67	<0.001
dimension	3	721	8.71	<0.001
run	1	721	0.0022	0.96
dimension:run	3	721	0.49	0.69

8.1.3 Magnitude Differentiation

Wilcoxon signed-rank tests confirmed that the thresholds for large and small magnitude of time ($W_{23} = 252, p < 0.001$), space ($W_{23} = 210, p < 0.001$), numerosity ($W_{23} = 276, p < 0.001$) and control ($W_{23} = 276, p < 0.001$) differed significantly, confirming that participants were able to distinguish between the low and high standard values. For the statistical analysis, outliers were excluded from the data using Tukey's criterion ($1.5 \cdot IQR$), as described in section 6.3.4. However, regardless of the exclusion or inclusion of outliers, significance levels remained the same. A graphical representation of the results is shown in Figure 8.3.

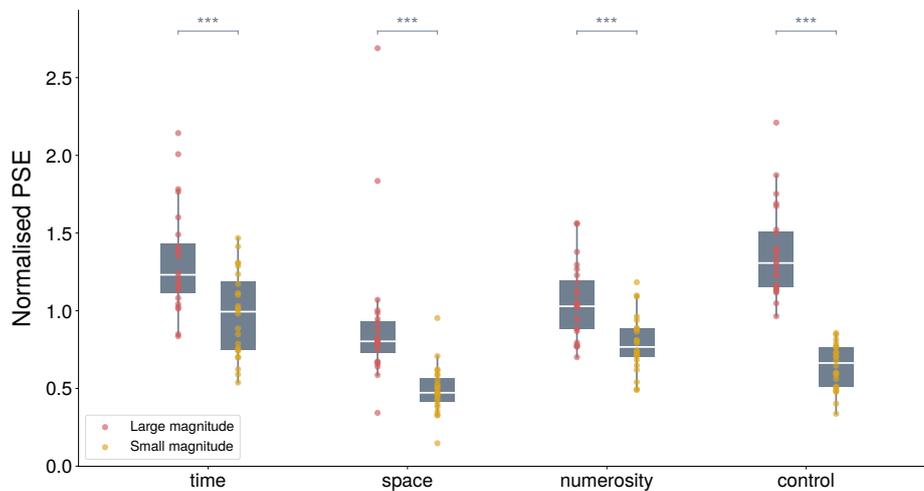


Fig. 8.3: Boxplot of the calculated thresholds (PSEs) of each participant for the four dimensions and two magnitudes, including outliers. Upper and lower boundaries of the boxes represent 25 % and 75 % quartiles. Whiskers extend to 1.5 times the interquartile range. Threshold values are normalized to the respective mean of reference values, as described in section 6.1.

8.2 Inter-Dimensional Judgement Precision

One of the main aims of this experiment was to investigate behavioural evidence for a shared magnitude processing system by comparing judgement precisions of the different dimensions. If the perception of two dimensions is grounded on the same neuronal mechanism, judgement precision should be correlated between these two dimensions. As an example, figure 8.4 illustrates two PFs with different widths from different participants to provide a graphical illustration how judgement precisions is reflected in the shape of the PF.

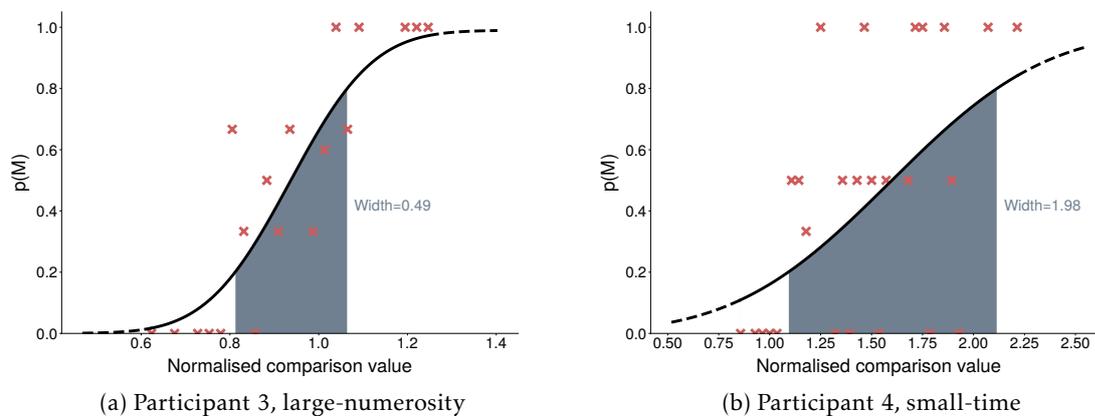


Fig. 8.4: Behavioural data (red crosses) and calculated PFs for two different participants and different task types as well as magnitudes to illustrate how the shape of the PFs differ with higher and lower precision. In this case, the judgements made by participant 3 in large-numerosity trials were more accurate than the ones of participant 4 in small-time trials. The y-axis denotes the probability of a participant responding with *more* when asked to compare the displayed value with the perceived value during the response phase.

Between each possible pair of dimensions, Spearman rank correlation coefficients were calculated and are summarized in Table 8.3. Analogue to the magnitude differentiation analysis above, outliers were identified and excluded from the statistical analysis. Again, this had no impact on the overall outcome. Results show that participants who were relatively accurate in their judgements of time were also more accurate in their perception of space. For the other dimensions, no significant correlations were found. Spearman rank plots for all dimensions, including the control condition, can be found in Figure B.1 of the appendix.

Table 8.3: Inter-dimensional precision correlations between all possible combinations of modalities. Reported are Spearman rank correlation coefficients and their corrected p-values (FWE = 0.05).

Dimension	ρ	p-value
Time & space	0.678	0.002
Time & numerosity	0.284	0.443
Space & numerosity	0.271	0.443
Time & control	0.237	0.768
Space & control	0.054	0.913
Numerosity & control	0.090	0.913

8.3 Cross-Dimensional Interference

As described in section 6.6, the last behavioural analysis examined the inter-dimensional bias on magnitude perception, meaning the impact an irrelevant dimension had on the judgement of the relevant dimension. A practical example is displayed in Figure 8.5, which shows the influence of distance as the irrelevant dimension on (i) time (Figure 8.5a) and (ii) numerosity (Figure 8.5b) judgements in one participant. The trials of the relevant dimension were split into trials with short or long distances and separate PFs were calculated for both conditions. In this example, distance biased time judgements in a positive direction whereas no influence of the trial distance on numerosity perception was visible.

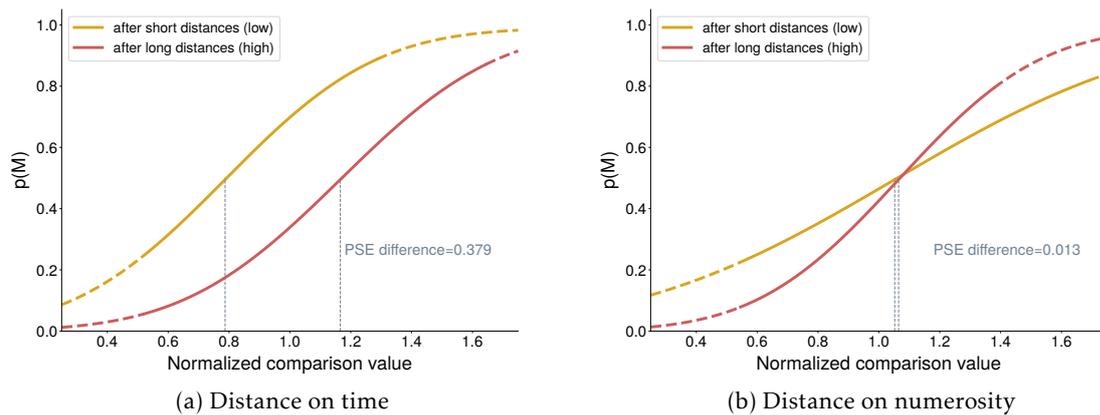


Fig. 8.5: Exemplary, the effect of trial distance as the irrelevant dimension on (a) time and (b) numerosity judgements are shown for participant 2. For each dimension, separate PFs were calculated after splitting all trials according to the distance travelled during the trial (either 11.5 m or 19.7 m). For time, a positive bias was observed. In contrast, numerosity judgements were not affected by distance.

Figure 8.6 visualizes the results obtained from all 24 participants. Out of all possible combinations, only time and space interfered in both directions with each other. Specifically, a Wilcoxon signed-rank test revealed a positive influence of space on time ($\bar{x} = 0.177$, $W_{23} = 36$, $p = 0.006$),

meaning that time was perceived longer when the distance travelled was also longer compared to shorter travel lengths, even though the actual time of the passive movement was the same in each trial. On the other hand, time as the irrelevant dimension influenced space negatively ($\bar{x} = -0.139$, $W_{23} = 45$, $p = 0.013$): for trials in which the participant travelled the exact same length and had to judge distance, the PSE was smaller when the duration of the movement (time) was 4.8 s compared to trials in which the movement only lasted 2.8 s. Among the other possible combinations of dimensions, no significant interference was found. Table B.1 shows the results for all possible combinations, including the control condition.

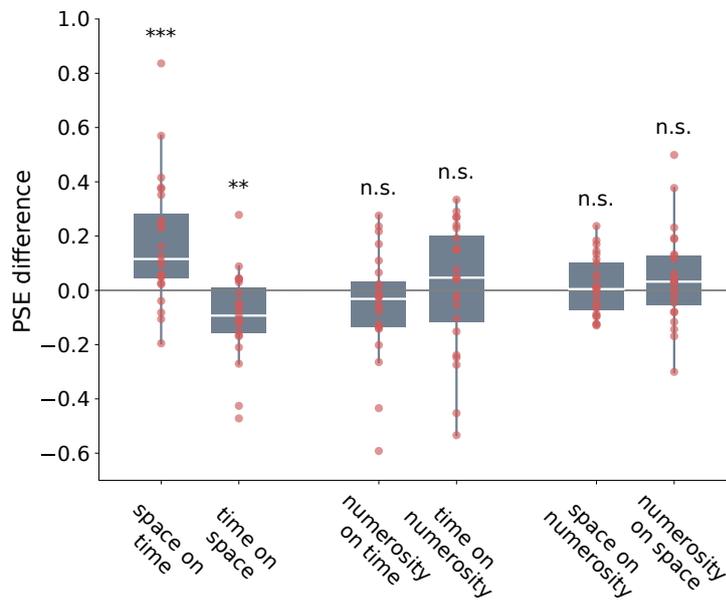


Fig. 8.6: Results of cross-dimensional interference on magnitude perception plotted with box-whisker plots. Individual data points are plotted as red dots. A positive PSE difference indicates higher PSE values if the irrelevant dimension was in high condition and vice versa. Significance was tested with a Wilcoxon signed-rank test and corrected for multiple comparisons (FWE = 0.05).

9. Neuroimaging Results

9.1 Quality Control

In this section, the results for the quality metrics introduced in section 7.2 are presented. Figures for the other parameters that were not selected may be found in section C.1 of the appendix.

9.1.1 Spatial Information

Entropy-focus criterion (EFC) provides a proxy for participant-movement induced blurriness and ghosting in MRI data (Atkinson et al. 1997), as introduced in section 7.2. Figure 9.1 shows the EFC for each participant separately. While there are noticeable differences in the EFC between participants (e.g. participant 4 vs. 16), none of the participants had values that were outside the inclusion area, so no detailed examination based on EFC was triggered.

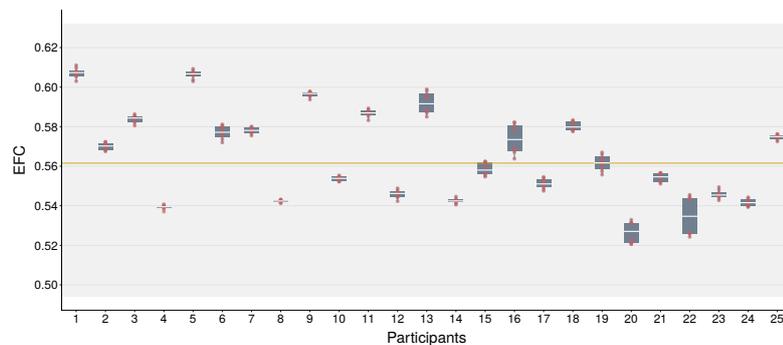


Fig. 9.1: Box-whisker plot of the entropy-focus criterion (EFC) of all participants. Individual runs are plotted as red dots. The group median is shown by the orange horizontal line. The area shaded in light grey visualizes the data range in which data points are not considered outliers, as defined in section 7.2. Lower values are better.

9.1.2 Temporal Information

Temporal SNR gives insight into the temporal variability of BOLD images. A low tSNR reflects higher fluctuations in the signal, mainly due to thermal noise and physiological signal alterations. In Figure 9.2, tSNR values of each participant and run are plotted. Similar to the EFC, no participant had a value outside the outlier range. Most of the participants have

varying tSNR values across the eight runs, with participant 5 being a noticeable exception, as her or his values are consistently low.

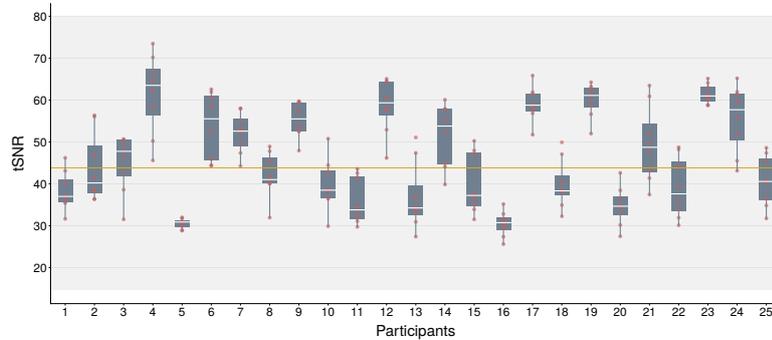


Fig. 9.2: Box-whisker plot of the whole-brain tSNR values of all participants. Individual runs are plotted as red dots. The group median is shown by the orange horizontal line. The area shaded in light grey is visualizes the data range in which data points are not considered outliers, as defined in section 7.2. Higher values are better.

9.1.3 Artefact Detection

Ghosting

Figure 9.3 shows the Ghost-to-signal ration (GSR) values along the phase encoding y-axis. Higher values indicate a significant presence of ghosting in the anterior and posterior sections of the image. Image data of participant 13 and 18 lie outside of the inclusion area and indicate the presence of ghosting. Hence, the corresponding images were manually examined.

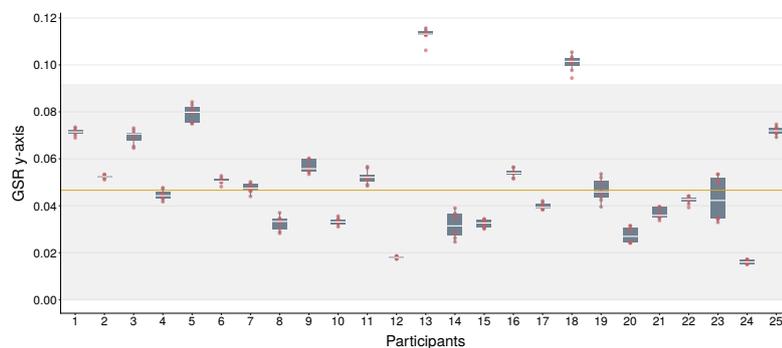


Fig. 9.3: Box-whisker plot of the GSR values along the phase-encoding y-axis of all participants. Individual runs are plotted as red dots. The group median is shown by the orange horizontal line. The area shaded in light grey is visualizes the data range in which data points are not considered outliers, as defined in section 7.2. Lower values are better.

Upon manual inspection, ghosting was found in all runs for participant 13 and 18. As an example, Figure 9.4a shows a volume of participant 13, session 1, run 4 in which ghost arte-

facts are visible at the top row of the image. Ghosting artefacts of participant 18 were similar in nature. After preprocessing with fMRIPrep, the ghosting was successfully removed in all volumes. For both participants, the brain masks were visually inspected as well, to make sure that ghosting did not have an adverse effect on this crucial preprocessing step (Figure 9.4c).

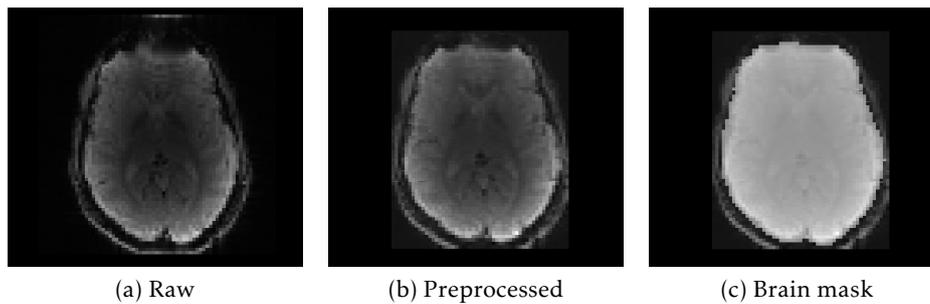


Fig. 9.4: (a): Raw data of an EPI volume of participant 13, session 1, run 4. Ghosting is visible in the top part of the image. After preprocessing, ghosting was successfully corrected (b). Additionally, the brain mask (c) calculation was not affected by the ghosting.

Movement Induced Artefacts

The percentage of volumes that exhibited a fieldwise displacement above 1 mm for all participants and runs is plotted on Figure 9.5. A median percentage of all participants of 0.54 % indicates an overall low level of head motion. However, some runs exceeded the inclusion range. Most strikingly, run 4 in session 1 from participant 20 has 10.45 % of volumes that exceeded the threshold of 1 mm FD, resulting in the exclusion of this particular run in subsequent analyses.

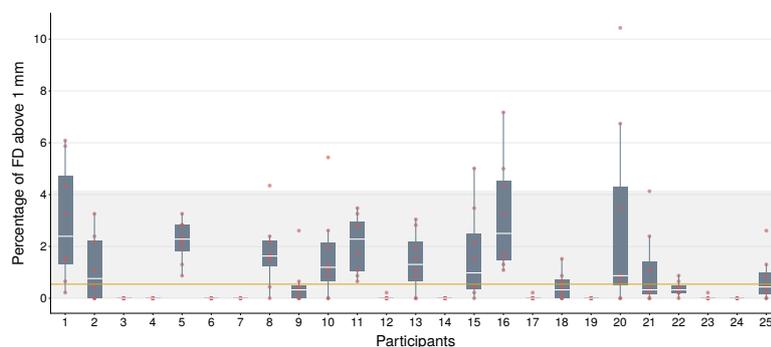


Fig. 9.5: Box-whisker plot of the percentage of volumes that are above a threshold of 1 mm for the FD. Individual runs are plotted as red dots. The group median is shown by the orange horizontal line. The area shaded in light grey is visualizes the data range in which data points are not considered outliers, as defined in section 7.2. Lower values are better.

9.2 Second Level Univariate Analysis

9.2.1 Introduction

In addition to behavioural data analyses, another main aim of this experiment was to measure BOLD activation in order to identify brain regions exclusively or commonly activated during the processing of time, space or numerosity processing. In this chapter, the results from the group level activation analyses are presented.

9.2.2 Activation During Accumulation Trials

The first contrast was calculated in order to identify neural structures that were commonly activated during the accumulation phase of all three dimensions (time, space or numerosity). As shown in Figure 9.6 and Table C.1, nine activation clusters could be identified, mainly located in (i) IFG bilaterally, (ii) the IPS bilaterally, (iii) the fusiform gyrus bilaterally, although predominantly in the right hemisphere, (iv) right and left posterior medial frontal gyrus as well as (v) a cluster spanning the right and left primary visual cortex.

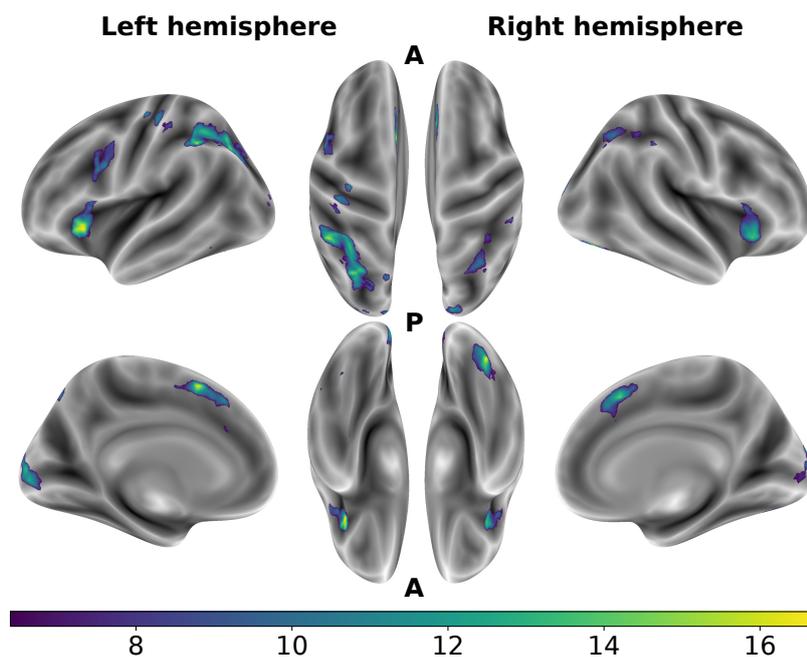


Fig. 9.6: Surface t-map of the average activation during time, space or numerosity trials (contrast 1 in Table 7.1). Glass brain plot shown in Figure C.7 of the appendix.

9.2.3 Accumulation Trials vs. Control Trials

The next contrast was designed to uncover regions that had a higher BOLD response during the accumulation trials compared to the control condition, regardless of the specific dimension. Here, four clusters remained after thresholding and are shown in Figure 9.7. Both in the right (78 voxels) and left (44 voxels) hemispheres, clusters in the pars orbitalis

of the respective IFG were identified. On the right hemisphere, two clusters in the inferior occipital gyrus as well as fusiform gyrus showed significant activation for this contrast.

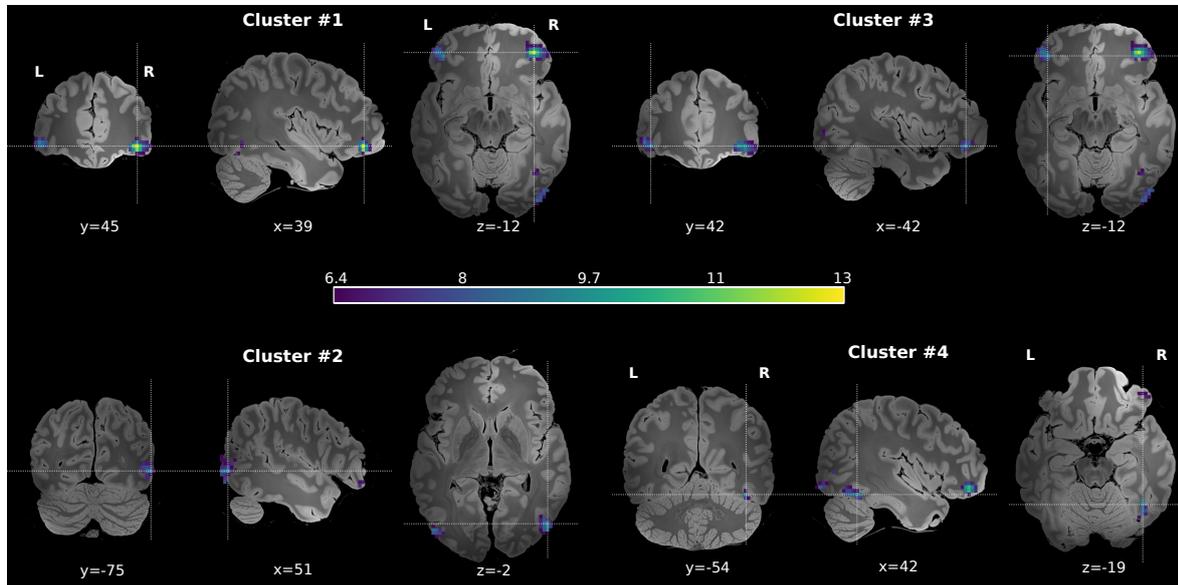


Fig. 9.7: T-map of voxels that show a significantly higher BOLD response during accumulation trials compared to control trials (contrast 2 in Table 7.1). Four clusters could be identified and slices at their local maxima are shown in coronal plane. Glass brain plot shown in Figure C.7 of the appendix.

Table 9.1: Clusters identified for contrast 2: activation during accumulation trials vs. control trials. The value of cluster size indicates the number of voxels. Coordinates represent peak level activation.

Nr.	Anatomical region	Cluster size	T	p	MNI coordinates
1	R inferior frontal gyrus	78	12.49	<0.001	(39, 45, -12)
2	R inferior occipital gyrus (hO4cla)	76	9.44	<0.001	(51, -75, -2)
3	L inferior frontal gyrus	44	9.41	<0.001	(-42, 42, -12)
4	R fusiform gyrus	35	8.84	<0.001	(-51, -81, -5)

9.2.4 Dimension Specific Activation

The next three contrasts (plotted in Figure 9.8 and Table 9.2) were calculated in order to identify regions that had a dimension-specific significant higher BOLD response compared to the control condition (contrasts 3-5 in Table 7.1). For time > control, bilateral clusters in the pars orbitalis of the IFG (left: 81 voxels, right: 68 voxels) as well as pars triangularis of the IFG (12 voxels) on the left hemisphere showed significant activation.

For the contrast space > control, three clusters located within the right hemisphere remain after thresholding: (i) in the pars orbitalis of the IFG, (ii) in the lingual gyrus and (iii) in the inferior parietal lobe (PGp).

In trials where numerosity was to be judged, six clusters were discovered. First, the group level analysis revealed a large cluster including 320 voxels on the right hemisphere with its maximum located in in the fusiform gyrus. Furthermore, this cluster expands to the lateral and inferior gyri and sulci occipitotemporalis. A similar, smaller cluster with 91 voxels is present in the middle occipital gyrus contralaterally. Additionally, four more clusters within the right hemisphere could be identified, notably including the IPS and cuneus.

As a last univariate analysis, a conjunction analysis of a all three accumulation conditions revealed only the right IFG as an area that exhibited higher BOLD activation during time, space and numerosity trials than during control trials.

Table 9.2: Clusters identified for contrasts 3-6 and the conjunction analysis. The value of cluster size indicates the number of voxels. Coordinates represent peak level activation.

Nr.	Anatomical region	Cluster size	T	p	MNI coordinates
Time > control					
1	L inferior frontal gyrus	81	11.56	<0.001	(-45, 45, -12)
2	R inferior frontal gyrus	68	9.28	<0.001	(36, 45, -12)
3	L inferior frontal gyrus	12	7.75	0.002	(-54, 27, 1)
Space > control					
1	R inferior frontal gyrus	38	11.00	<0.001	(36, 45, -12)
2	R lingual gyrus	21	7.54	0.004	(18, -66, -2)
3	R Inferior parietal lobe (PGp)	10	7.33	0.006	(45, -81, 24)
Numerosity > control					
1	R fusiform gyrus	320	13.44	<0.001	(42, -51, -19)
	R middle temporal gyrus (V5)		11.47	<0.001	(45, -66, 1)
	R inferior occipital gyrus (V4)		11.24	<0.001	(45, -84, -12)
2	L inferior occipital gyrus (hO4cla)	91	10.97	<0.001	(-54, -78, -2)
3	R middle orbital gyrus	33	8.92	0.001	(39, 48, -15)
4	R intraparietal sulcus	28	8.21	0.001	(48, -36, 54)
5	R middle occipital gyrus	19	7.97	0.001	(30, -72, 34)
6	R cuneus	10	7.60	0.003	(24, -84, 44)
(Time > control) \cap (Space > control) \cap (Numerosity > control)					
1	R inferior frontal gyrus	29	6.52	0.01	(42, 45, -12)

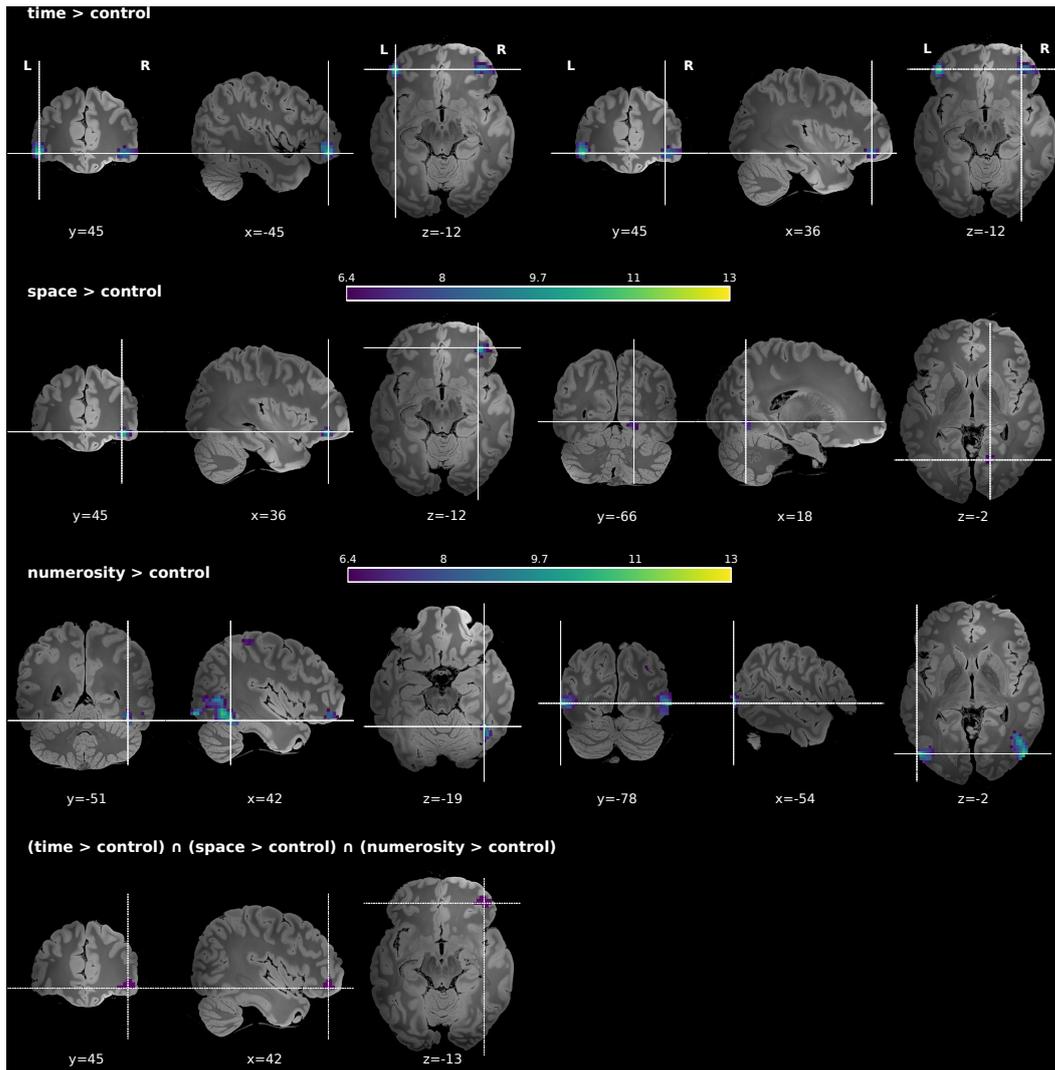


Fig. 9.8: T-maps of two exemplary clusters that showed highest activation for each dimension compared to the control condition (a-c) and the conjunction analysis (d). Glass brain plots of each contrast are shown in Figure C.7 and C.8 of the appendix.

9.3 Multivariate Pattern Analysis

9.3.1 Decoding of Accumulation Trials Against Control Trials

As described in section 7.5.3, MVPA was performed with an inter-subject pattern analysis approach. First, the three trial types in which the participants had to perform accumulation of either time, space or numerosity were decoded pairwise against trials in which the participants had to judge the luminance of a square (control condition). This analysis was done in order to gain insight into the activation patterns generally associated to the processing of time, space or numerosity and whether these areas are specific for one dimension or not.

To provide a first overview, glass brain plots are shown in Figure 9.9. Common for all three dimensions was a cluster located in the left and right thalamic area that remained after significance testing, but showed an overall low classification accuracy when compared to other regions. Another area in which classification between accumulation and control condition was possible for all tested dimensions was the pars orbitalis of the right IFG (depicted exemplary for time in Figure 9.10B). Decoding of both space or numerosity versus control trials was possible in a common cluster that is located in the left and right precuneus (area 7M, 7P and 7A), shown in Figure 9.10 as well as C.10 of the appendix. Although both time and space can be decoded against control trials within the lingual gyrus, the cluster for time trials is located anterior to the one for space trials vs. control trials.

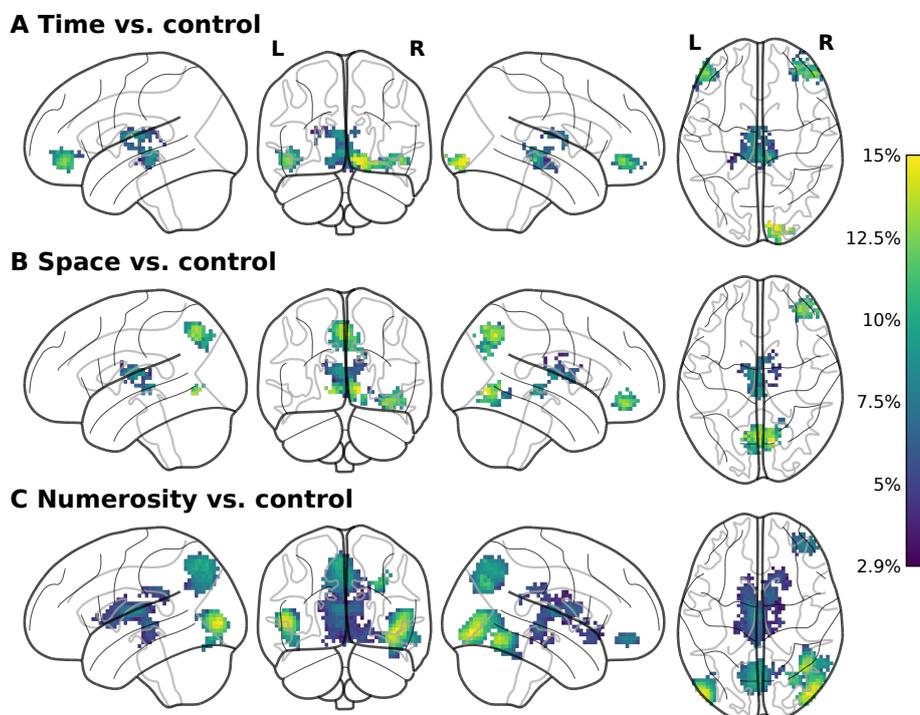


Fig. 9.9: Glass brain plots of average decoding accuracy above chance level (50 %) of 24 inter-subject folds for all three accumulation dimensions against the control dimension. Colormap values are fixed at the lowest (2.9 %) and the highest (15 %) accuracy values.

One unique cluster remains after significance testing on the group level for time versus control: pars orbitalis of the left IFG (98 voxels). Decoding of numerosity condition versus control condition is uniquely possible bilaterally in a large cluster with 473 voxels (right) and 181 voxels (left), spanning along the fusiform gyrus on the right side as well as bilaterally the posterior temporal cortex and anterior inferior parts of the occipital cortex (area hOc4la, hOc4lp and hOc5). Exemplary, the cluster in the right hemisphere is depicted in Figure 9.10D.

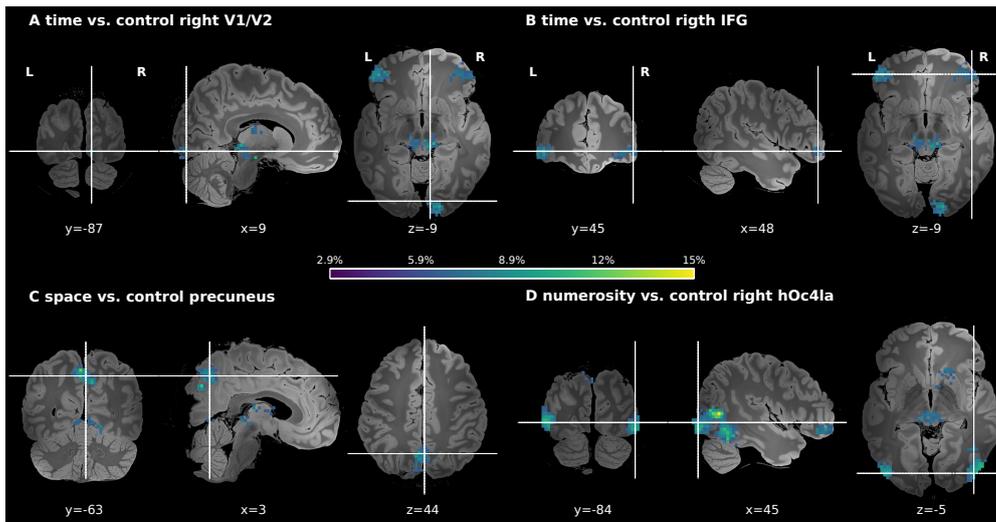


Fig. 9.10: Accuracy maps of exemplary clusters that show significant decoding accuracy against the control condition. Values are percentages above chance level of 50 %.

Table 9.3: Clusters which show significant decoding accuracies when decoding each dimension vs. control trials. The value of cluster size indicates the number of voxels. Accuracy values are in percent above chance (50 %). Coordinates represent peak level activation.

Nr.	Anatomical region	Cluster size	Accuracy	MNI coordinates
Time vs. control				
1	R lingual gyrus (V2)	64	10.06	(9, -87, -9)
	R lingual gyrus (V1)		9.33	(18, -99, -12)
	R inferior occipital gyrus (V3v)		9.29	(27, -93, -12)
2	R inferior frontal gyrus	75	8.95	(48, 45, -9)
3	L inferior frontal gyrus	98	8.84	(-48, 45, -9)
4	R + L mesencephalon	230	8.46	9, -30, -9)
	L thalamus		6.98	(-3, -15, +14)
	R thalamus		6.97	(9, -27, -2)
Space vs. control				
1	R + L lingual gyrus (V2)	63	9.71	(12, -72, -2)
2	R + L precuneus	184	9.50	(3, -63, 44)
3	R middle orbital gyrus	77	8.84	(36, 48, -15)
4	R + L thalamus	155	6.81	(0, -30, -5)
Numerosity vs. control				
1	R inferior occipital gyrus (hOc4la)	473	14.65	(45, -84, -5)
	R fusiforme gyrus		13.74	(39, -60, -12)
	R middle temporal gyrus (V5)		12.73	(51, -72, 8)
2	L inferior occipital gyrus (hOc4la)	181	14.60	(-48, -87, 1)
3	R superior occipital gyrus	145	11.43	(27, -69, 34)
4	R + L precuneus	310	10.51	(0, -72, 54)
5	R + L thalamus	509	8.29	(-3, -18, 14)
6	R inferior frontal gyrus	88	8.22	(42, 48, -12)
7	L posterior cingulate gyrus	67	7.12	(-3, -39, 28)

9.3.2 Crossdimensional Decoding

In a next step, pairwise whole brain ISPA decoding was performed for every combination of accumulation conditions (time, space or numerosity). One might speculate that a significant decoding accuracy between two dimensions is an indicator for different neural processing of these dimensions within these areas. On the contrary, failure to decode dimensions hints at a similar processing of these dimensions. As an overview, Figure 9.11 shows glass brain plots for the decoding results. Figure 9.12 show slices of specific clusters for a better localization.

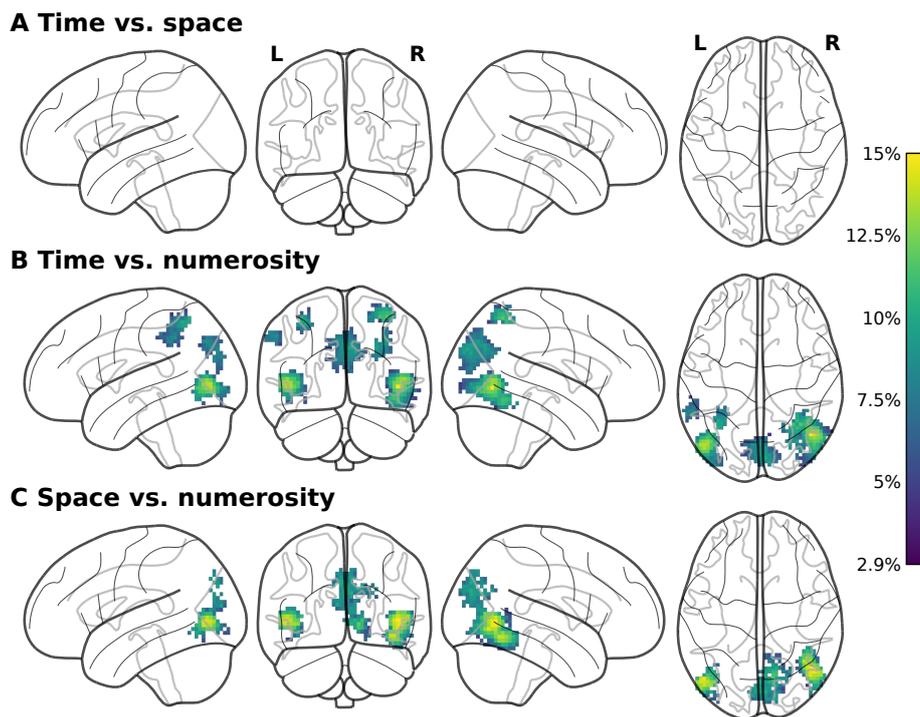


Fig. 9.11: Glass brain plots of average decoding accuracy above chance level (50 %) of 24 inter-subject folds for all three accumulation dimensions against each other. For time vs. space, no significant voxels remain after significance thresholding.

Interestingly, the decoding algorithm could not identify any clusters in which a separation between time and space trials is possible, hence Figure 9.11A shows an empty glassbrain plot.

For the pairwise decoding of space or time trials vs. numerosity, similar areas seem to contain significant decoding information, as visual inspections of the patterns of Figure 9.11B and C reveal. First, large clusters in both hemispheres along the fusiform gyrus, inferior, middle as well as posterior temporal areas (Figure 9.12A) showed neuronal activity that enabled decoding of time or space trials vs. numerosity trials. The same clusters also (i) exhibited unique firing patterns when decoding numerosity vs. control trials and (ii) showed a unique activation contrast for numerosity > control in the activation analysis (see section 9.2.4). Another cluster in which successful decoding of time or space versus numerosity was possible was located bilaterally at the border between the occipital and parietal cortex (cuneus and precuneus, see Figure 9.12B).

Two unique clusters could be identified for the decoding of time trials versus numerosity trials: (i) bilaterally, clusters located at the IPS (Figure 9.12C) and (ii) within the inferior parietal lobe in the left hemisphere.

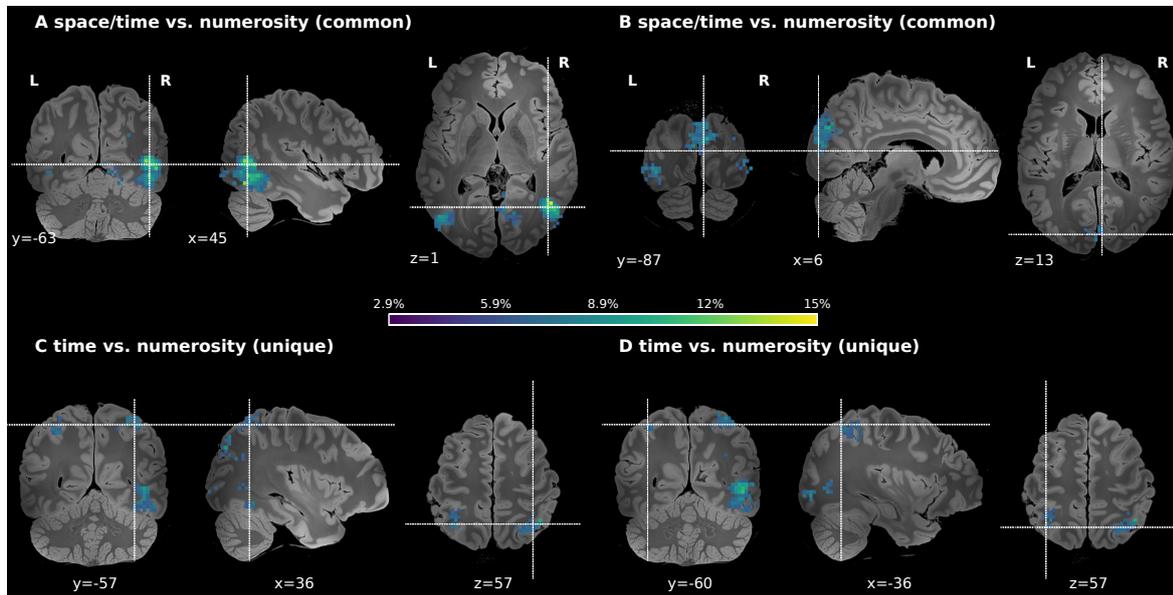


Fig. 9.12: Accuracy maps of exemplary clusters exhibiting significant decoding accuracy between dimensions. A and B show clusters that are common when decoding time or space trials vs. numerosity trials. The bottom row (C and D) show clusters that are unique for the decoding pair time vs. numerosity. Values are percentages above chance level of 50 %.

Table 9.4: Clusters with significant decoding accuracies when performing crossdimensional decoding. The value of cluster size indicates the number of voxels. Accuracy values are in percent above chance (50 %). Coordinates represent peak level activation.

Nr.	Anatomical region	Cluster size	Accuracy	MNI coordinates
Time vs. numerosity				
1	R middle temporal gyrus (V5)	343	12.99	(45, -66, -2)
	R fusiforme gyrus		7.69	(42, -45, -15)
2	L middle temporal gyrus (V5)	220	11.96	(-48, -78, 1)
3	R intraparietal sulcus	67	10.36	(36, -57, +57)
4	L intraparietal sulcus	51	9.12	(-36, -60, 57)
5	R + L cuneus	263	8.87	(6, -87, 13)
6	L inferior parietal lobe	41	8.04	(-57, -48, +41)
Space vs. numerosity				
1	R middle temporal gyrus (V5)	335	11.21	(45, -63, 1)
	R inferior temporal gyrus (FG1)		10.56	(42, -60, -12)
2	L middle occipital gyrus (hOc4la)	135	10.99	(-48, -72, 1)
3	R + L cuneus	204	8.02	(0, -84, 28)

**PART IV:
CONCLUSION**

10. Discussion

10.1 Paradigm Design

This thesis examined magnitude perception of time, space and numerosity on a behavioural and neuronal level. Creating paradigms which allow tapping into these higher order mental representations is not trivial, as multiple, potentially confounding, aspects have to be considered.

First, distance and numerosity are both dimensions that may be presented in a sequential or instantaneous format. The same is not true for time, as it can only be experienced in an accumulative (i.e. sequential) manner. The importance of an identical presentation format has been highlighted in a recent review by (Hamamouche et al. 2019). By using passive movement inside a virtual environment, with dots appearing at random locations and with a random lifetime, participants were forced to make an accumulative judgement for space and numerosity. Furthermore, by presenting the control condition (luminescence of a white square) in an instantaneous manner, the paradigm allowed to study precisely the accumulation phase of time, space or numerosity perception in contrast to an instant judgement of a fourth dimension. A potential difference in perception of time trials vs. all other dimensions remained, as participants could have mentally counted down the seconds during movement in order to judge elapsed time, even though there were instructed not to do so (see section 5.3). However, given that reaction times were highest for time trials (see section 8.1), this seems unlikely, as counting down would have reduced the difficulty of time trials substantially.

Second, identical visual input for all trials was an important aspect during the design of this paradigm. Only then could variances in BOLD activation or shape of activation patterns be attributed to different neuronal magnitude perception mechanisms. A potential limitation of this study is that visual attention to different elements of the virtual environment during the trials was not fully controlled. For example, in numerosity trials, participant were likely focusing the ground to a higher degree than in other trial types, thus increasing the sensation of optical flow for that specific trial type. In future studies, the inclusion of eye-tracking could help to elicit neuronal processes that are numerosity-specific by including the point of gaze as a confound.

Third, extensive pilot testings were conducted in order to fine-tune the standard values in high and low condition as well as the initial comparison values. This was done in order to achieve similar judgement difficulties between magnitude judgements of all dimensions. Furthermore, post-hoc manipulation checks were performed before any further analyses of

the behavioural data were conducted. To this end, the reaction times and the reported task difficulty were analyzed. Hereby, all three trial types of interest were subjectively perceived as identical in their difficulty. While reaction times differed significantly between numerosity and time trials, the change in reaction times over the course of the experiment was identical between all task types, i.e. participants answered faster as the experiment progressed, but did so equally for all dimensions. In a third preliminary analysis, it was confirmed that participants could distinguish the two different magnitudes (low and high condition) of each trial type by comparing their normalized PSE values.

Finally, trial ordering proved to be a challenging aspect of paradigm design, as the order had to satisfy multiple, partially competing, requirements. As an illustrative example, one can envision that the time a participant comfortably spends inside the scanner is limited. However, classification performance by MVPA usually improves with the amount of training data (Haxby et al. 2014). Based on results from pilot scans and other fMRI studies, a target for approximately 100 trials per condition was set. Furthermore, the amount of training data was drastically improved by using ISPA, thus pooling trials for training and testing across subjects. In order to minimize the effect of neural habituation and maximize effect sizes, counterbalancing of trials was achieved by creating *de Bruijn* sequences (Aguirre 2007, Aguirre et al. 2011).

10.2 Behavioural Data

In order to quantify magnitude perception across time, space and numerosity, psychometric functions were fitted to the magnitude judgements of each participant and each dimension. This proved to be highly challenging, as restrictions on the amount of trials, imposed by limited scan times, resulted in overdispersed data and a lower than usual number of trials per stimulus intensity. To overcome this challenge, *psignifit 4* (Schütt et al. 2016), a Bayesian framework to estimate PFs, was used. After modification of the priors to the characteristics of the behavioural data, this approach was able to partially compensate for overdispersed data and provide a good proxy for credible intervals for the majority of behavioural data. In future studies, the reliability of the behavioural data could be improved by reducing constraints on the trial ordering and thus, allowing for more trials per stimulus interval. Additionally, an adaptive algorithm that monitors the amount of trials for each stimulus intensity during the experiment could be beneficial and help reduce uncertainty in fitting the PFs. Another methodological improvement to the reliability of the behavioural results could be achieved by using a multi-level Bayesian model to incorporate the size of credible intervals on the group level. In the present study, PFs were only fitted on the individual level with a Bayesian approach. Results on the group level were then obtained by computing frequentist statistical methods.

Behavioural data analysis focused on two aspects: (i) comparison of judgement precisions between the different dimensions and (ii) the cross-dimensional bias afflicted by one dimension on the other.

Judgement precision (see section 8.2), characterized by the width of the fitted PF, showed significant correlation between the acuity for time and space perception, but not for any other combination of dimensions. As iterated in section 4.2.1, research regarding judgement precision of different dimensions remains inconclusive at the point of this thesis (Hamamouche et al. 2019). Furthermore, a comparison with other studies is challenging, as few studies have investigated time, space and number perception during self-motion and while presenting the stimuli in a sequential manner. A study by (Droit-Volet et al. 2008) included both sequential and non-sequential presentation modes for space and numerosity. In contrast to results of this study, the authors found no differences in judgement performance between time, space and number when all three dimensions were presented sequentially. Similarly to judgement precision, a cross-dimensional bias could only be found between time and space (section 8.3). It was found that time judgements were positively influenced by distance, i.e. the duration of time trials was perceived as longer when the distance travelled during the trials was longer, even though the actual duration was the same. Counterintuitively, the duration of a trial biased distance judgements negatively: when the duration of a distance trial was shorter, the distance travelled was judged as farther compared to longer lasting trials with the exact same distance. This effect has been reported before (Lambrechts et al. 2013) and stands in apparent contrast to other studies which reported a positive influence of time on space perception (Riemer et al. 2018, Cai et al. 2015). However, the current findings may be explained by considering the additional factor of speed. When travel time is kept constant but traversed distance increases, so does speed. Conversely, in trials with identical travelled distance and increased travel time, speed decreases. Therefore, a positive effect of trial duration on perceived distance and a negative bias of trial distance on perceived duration is expected, if one assumes that interference between time and space is predominately driven by the factor of speed. In line with this interpretation, studies using static stimuli, and thus excluding the additional factor of speed, reported a positive bias of time on space ((exp. 1 Riemer et al. 2018) or (Cai et al. 2015)) whereas studies using dynamic stimuli reported a negative (e.g. Martin et al. 2017) or the absence of ((exp. 2 Riemer et al. 2018) or (Casasanto et al. 2008)) an interference effect. It is important to note that movement speed was counterbalanced within the trial order in order to avoid an asymmetric bias.

10.3 Functional Imaging Data

The quality of the functional and anatomical image data from 24 participants was first thoroughly checked by calculating various IQMs with the excellent toolset available in *MRIQC* (Esteban et al. 2017). Overall, a high quality of the dataset could be assessed. One run of a participant had to be excluded due to excessive movement. Otherwise, the amount of movement was within tolerable limits. Ghosting artefacts detected in the data of two participants were successfully removed during preprocessing and were therefore included in the analysis.

The obtained functional image data were analysed with two different methodologies: (i) a

univariate GLM analysis to identify regions that are activated during magnitude processing of each dimension and (ii) an inter-subject MVPA (ISPA) analyses with the aim to decode unique activation patterns evoked while perceiving each dimension.

Univariate activation analysis revealed the right IFG as a commonly activated ROI during the accumulation of all three dimensions in contrast to the control condition. This was evident both in individual contrasts against control trials as well as in a conjunction analysis. Uniquely, the left IFG showed higher activation only during time trials. In numerosity trials, the fusiform gyrus/V5 bilaterally and the cuneus bilaterally were ROI with dimension-specific activation.

According to (Hayashi et al. 2013), the IFG is specifically involved in decision processes of magnitude estimation. Our findings support this assumption, as the decision between *less* or *more* is equal for time, space and numerosity trials. Additionally, the IFG has been hypothesized as a key structure within a fronto-basal-ganglia network for neural and behavioural of inhibition (Jahanshahi et al. 2015). According to this model, the activation of the IFG during the accumulation phase in this study could be due to the action inhibition until a decision had to be made in the comparison phase. A study which also conducted fMRI measurements and magnitude judgements (Skagerlund et al. 2016) identified more regions commonly activated during time, space and numerosity processing, including the insula and right IPS. However, comparison of these findings remains difficult, as the author also reported considerable differences between the difficulty levels for the respective tasks. This holds also true for other studies attributing neuronal activation patterns to the perception of specific dimensions, as it was reported that these were confounded by differences in task difficulty ((e.g. Tregellas et al. 2006) or (Livesey et al. 2007)).

Specifically for numerosity trials, areas MT/V5 showed higher activation when contrasted against control trials in the univariate analysis. A possible explanation could be the heightened focus of the participants on the moving virtual floor in front of them, which would lead to an increase in the processing of optic flow and the speed of directional movement. These findings are in line with studies linking the MT/V5 area to (i) the detection of translational change (Martinez-Trujillo et al. 2007), (ii) processing of speed gradients (Martinez-Trujillo et al. 2005) and (iii) registering uni-directional motion (Zeki 2015).

ISPA was always performed pairwise. First, each dimension was decoded against control trials. Interestingly, a major portion of these results mirror the ones from the univariate analysis: areas that evoked a higher BOLD response in accumulation trials also gave rise to activation patterns that can be successfully decoded from control trials. For example, for all dimensions, the right IFG showed a unique activation pattern when decoded from control trials. Specific for time trials, activation in the left IFG was significantly different from control trials in the left IFG. Additionally, bilateral posterior temporal regions and anterior inferior parts of the occipital cortex had unique activation patterns in numerosity trials when compared to control trials. This suggests that these areas do not only show a heightened BOLD activation during perception of the respective dimension, but also spe-

cific activation patterns in comparison to the control condition. An area that does not show increased activation in univariate analysis but in which decoding of time, space and numerosity against control trials is possible is the thalamus bilaterally. Similar to the IFG, the thalamus has been associated with inhibition of action (Jahanshahi et al. 2015) within the same cortico-basal-ganglia network and again, could be involved in withholding a decision during the accumulation phase. However, there is also recent evidence that the thalamus is a key area in switching the cortex into a highly excitable, asynchronous state in situations with a high uncertainty (Kosciessa et al. 2021).

Remarkably, when decoding accumulation trials in a pairwise manner against each other, no ROI could be identified that differentiates between time and space trials. In contrast, both time and space trials could be successfully decoded bilaterally in posterior temporal regions as well as at the junction of parietal and occipital cortices (cuneus and precuneus). Selective for time vs. numerosity trials, both right and left IPS evoked unique activation patterns for this pair of dimensions.

10.4 Concluding Remarks

Two main theories are currently discussed as the origin of cross-dimensional interferences. First, a potential origin for cross-dimensional influences are similar neuronal encodings (Srinivasan et al. 2010). Following this hypothesis, a bi-directional influence of time and space could be originating from similar brain mechanisms for the perception of these two dimensions (Bonato et al. 2012). The second theory, called metaphoric structuring, postulates that imposition of the concept of one dimensions upon the other dimension causes interference (Boroditsky 2000). For example, as illustrated in section 1, humans in Western cultures frequently use a mental timeline from left to right to order sequential events. The theory of metaphoric structuring would imply an asymmetrical interference between space and time, due to the dominance of spatial perception (Lambrechts et al. 2013, Riemer 2015). Thus, one would expect time to be more influenced by space and vice versa (Casasanto et al. 2008).

Taking into account both the behavioural and the functional imaging results, the present thesis is more in line with the former hypothesis of overlapping neuronal representations. On the behavioural level, the data of this thesis is explainable by both theories. However, the inability to decode time from space trial activation patterns, while each one alone is distinguishable from numerosity trials, hints more towards overlapping neuronal representations of time and space, while numerosity perception could be computed by a distinct mechanism, such as the ANS (Dehaene et al. 1998, Brannon 2006, Odic et al. 2018).

As mentioned above, Area MT/V5 is critically involved in processing visual optic flow and serves as important input for the computation of movement speed (Martinez-Trujillo et al. 2005, Martinez-Trujillo et al. 2007, Zeki 2015). As the relation between travel time and travelled distance is fully determined by speed, neuronal processes in area MT/V5 are a potential source for the perceptual influence between time and space.

In order to further elicit the role of each brain region and potentially unmask a neuronal network responsible for magnitude processing, future work could employ a normative connectome or individualized structural connectivity analysis which e.g. correlate individual inter-dimensional influences on the behavioural level with connectivity strengths within different ROI. Similarly, a functional resting-state fMRI dataset could be obtained alongside this experiment. Again, behavioural data could then be correlated with functional connectivity profiles within predefined networks.

11. Summary

The present thesis aimed at shedding light on the behavioural and neural perception of time, space and numerosity in humans within a spatial navigation context. An fMRI imaging paradigm was carefully designed in which participants were passively moved forward along a virtual path while randomly placed dots appeared on the floor. Each trial, participants had to either judge trial duration, distance covered or numerosity of items on the floor.

At the behavioural level, the acuity of judgements showed significant correlation only between time and space trials. Identically, a bi-directional bias was only evident between time and space with opposing directions: time judgements were positively influenced by distance, but conversely, trial duration biased distance judgements negatively.

Univariate activation analysis on the second level revealed higher activation of the right IFG during the accumulation of all three dimensions when contrasted against control trials. Activation clusters that were unique to one dimension were only found in the left IFG for time trials as well as in (i) the fusiform gyrus/V5 bilaterally and (ii) the cuneus for numerosity trials.

Using ISPA, each dimension was first decoded against control trials in a pairwise manner. In general, areas that evoked a higher BOLD response in accumulation trials also evoked activation patterns that could be successfully decoded from control trials. For example, for all dimensions, the right IFG showed a unique activation pattern when decoded from control trials. Unique for time trials, activation in the left IFG was significantly different from control trials in the left IFG. Additionally, bilateral posterior temporal regions and anterior inferior parts of the occipital cortex had unique activation patterns in numerosity trials when compared to control trials.

Interestingly, decoding of accumulation trials in a pairwise manner was not possible for time and space trials. Conversely, successful classification of neural activation patterns of time and space trials against numerosity trials was possible bilaterally in posterior temporal regions as well as at the junction of parietal and occipital cortices (cuneus and precuneus). Given the ongoing scientific debate about the existence of a common magnitude system in adults, these findings add to the body of evidence that time and space are indeed based on overlapping neuronal representations, while the judgement of numerosity could be grounded in a separate system.

12. Zusammenfassung

Die zugrunde liegenden neuronalen Prozesse der Wahrnehmung von Zeit, Raum und anderen Größenordnungen sind Gegenstand intensiver Forschung. Darüber hinaus wird häufig über gegenseitige Wechselwirkungen zwischen der Wahrnehmung verschiedener Dimensionen berichtet, aber der Ursprung dieser Interferenzeffekte ist noch nicht gut verstanden. Diese Arbeit untersuchte Zeit-, Raum- und Numerositätsurteile sowie deren gegenseitige Beeinflussung auf verhaltens- und neuronaler Ebene mittels der funktionellen-MRT Bildgebung. Hierzu wurden die Versuchsteilnehmenden passiv auf einem virtuellen Pfad vorwärts bewegt, der mit zufällig erscheinenden Punkten bedeckt war. Nach jedem Versuch wurden die Teilnehmenden gebeten, (i) die Dauer der Bewegung, (ii) die zurückgelegte Entfernung oder (iii) die Anzahl der Punkte auf dem Boden zu beurteilen.

Die Verhaltensergebnisse zeigten eine Korrelation zwischen der Genauigkeit zeitlicher und räumlicher Beurteilungen sowie eine bidirektionale Interferenz zwischen ausschließlich diesen beiden Dimensionen. Analysen der bildgebenden Daten identifizierten den rechten IFG als ein gemeinsam aktives Areal während der Wahrnehmung aller Dimensionen. Im Gegensatz dazu wurde der Bereich MT/V5 bilateral einzig während Numerositätsversuchen aktiviert. Aktivierungsmuster von Zeit- und Raumversuchen waren mittels multivariater Musteranalyse nicht zu unterscheiden, konnten jedoch von Numerositätswahrnehmungsversuchen im MT/V5 und IPS bilateral dekodiert werden.

Die Ergebnisse legen nahe, dass die Interferenz zwischen Zeit und zurückgelegter Entfernung durch neuronale Berechnungen der Bewegungsgeschwindigkeit auf der Grundlage visueller optischer Flußinformationen vermittelt werden könnte. Darüber hinaus ergänzen die Ergebnisse die Beweise dafür, dass die interdimensionalen Interferenzeffekte zwischen Zeit und Raum auf überlappenden neuronalen Darstellungen beruhen.

Bibliography

1. Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G.: Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*. 8. 14 (2014)
2. Agrillo, C., Piffer, L., Adriano, A.: Individual differences in non-symbolic numerical abilities predict mathematical achievements but contradict ATOM. *Behavioral and brain functions: BBF*. 9. 26 (2013)
3. Agrillo, C., Ranpura, A., Butterworth, B.: Time and numerosity estimation are independent: Behavioral evidence for two different systems using a conflict paradigm. *Cognitive Neuroscience*. 1. 96–101 (2010)
4. Aguirre, G. K.: Continuous carry-over designs for fMRI. *NeuroImage*. 35. 1480–1494 (2007)
5. Aguirre, G. K., Mattar, M. G., Magis-Weinberg, L.: de Bruijn cycles for neural decoding. *NeuroImage*. 56. 1293–1300 (2011)
6. Alcala-Quintana, R., Garcia-Perez, M.: The Role of Parametric Assumptions in Adaptive Bayesian Estimation. *Psychological methods*. 9. 250–271 (2004)
7. Allefeld, C., Goergen, K., Haynes, J.-D.: Valid population inference for information-based imaging: Information prevalence inference. *NeuroImage*. 141. 378–392 (2016)
8. Andersson, J., Skare, S., Ashburner, J.: How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*. 20. 870–888 (2003)
9. Atkinson, D., Hill, D. L., Stoye, P. N., Summers, P. E., Keevil, S. F.: Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion. *IEEE Transactions on Medical Imaging*. 16. 903–910 (1997)
10. Banquet, J., Gaussier, P., Cuperlier, N., Hok, V., Save, E., Poucet, B., Quoy, M., Wiener, S.: Time as the fourth dimension in the hippocampus. *Progress in Neurobiology*. 101920 (2020)
11. Behzadi, Y., Restom, K., Liau, J., Liu, T. T.: A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*. 37. 90–101 (2007)

12. Bjoertomt, O., Cowey, A., Walsh, V.: Spatial neglect in near and far space investigated by repetitive transcranial magnetic stimulation. *Brain : a journal of neurology*. 125. 2012–2022 (2002)
13. Boccara, C., Sargolini, F., Thoresen, V., Solstad, T., Witter, M., Moser, E., Moser, M.-B.: Grid cells in pre- and parasubiculum. *Nature Neuroscience*. 13. 987–994 (2010)
14. Bonato, M., Zorzi, M., Umiltà, C.: When time is space: Evidence for a mental time line. *Neuroscience and Biobehavioral Reviews*. 36. 2257–2273 (2012)
15. Boos, D., Stefanski, L.: *Likelihood Construction and Estimation*. pp. 27–124. Springer. 2013.
16. Boroditsky, L.: Metaphoric Structuring: Understanding Time through Spatial Metaphors. *Cognition*. 75. 1–28 (2000)
17. Bottini, R., Casasanto, D.: Space and Time in the Child’s Mind: Metaphoric or ATOMIC? *Frontiers in psychology*. 4. 803 (2013)
18. Brannon, E.: The representation of numerical magnitude. *Current Opinion in Neurobiology*. 16. 222–229 (2006)
19. Brannon, E., Lutz, D., Cordes, S.: The development of area discrimination and its implications for number representation in infancy. *Developmental science*. 9. 59–64 (2006)
20. Brannon, E., Suanda, S., Libertus, K.: Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Developmental science*. 10. 770–777 (2007)
21. Buetti, D., Walsh, V.: The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 364. 1831–1840 (2009)
22. Buhusi, C., Meck, W.: Relativity Theory and Time Perception: Single or Multiple Clocks? *PLoS ONE*. 4. 62–68 (2009)
23. Buhusi, C., Meck, W.: What makes us tick? Functional and neural mechanisms of interval timing. *Nature reviews. Neuroscience*. 6. 755–765 (2005)
24. Buonomano, D., Laje, R.: Population Clocks: Motor Timing with Neural Dynamics. *Trends in cognitive sciences*. 14. 520–527 (2010)
25. Buzsáki, G., Tingley, D.: Space and Time: The Hippocampus as a Sequence Generator. *Trends in Cognitive Sciences*. 22. 853–869 (2018)
26. Cai, Z. G., Connell, L.: Space-time interdependence: Evidence against asymmetric mapping between time and space. *Cognition*. 136. 268–281 (2015)

27. Cantrell, L., Boyer, T., Cordes, S., Smith, L.: Signal clarity: An account of the variability in infant quantity discrimination tasks. *Developmental Science*. 18. 877–893 (2015)
28. Cantrell, L., Smith, L.: Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition*. 128. 331–352 (2013)
29. Cappelletti, M., Barth, H., Fregni, F., Spelke, E., Pascual-Leone, A.: rTMS over the intraparietal sulcus disrupts numerosity processing. *Experimental brain research*. 179. 631–642 (2007)
30. Casasanto, D., Boroditsky, L.: Time in the mind: using space to think about time. *Cognition*. 106. 579–593 (2008)
31. Castelli, F., Glaser, D. E., Butterworth, B.: Discrete and analogue quantity processing in the parietal lobe: a functional MRI study. *Proceedings of the National Academy of Sciences of the United States of America*. 103. 4693–4698 (2006)
32. Chen, Y., Namburi, P., Elliott, L., Heinzle, J., Soon, C., Chee, M., Haynes, J.-D.: Cortical surface-based searchlight decoding. *NeuroImage*. 56. 582–592 (2011)
33. Cordes, S., Brannon, E.: The Difficulties of Representing Continuous Extent in Infancy: Using Number Is Just Easier. *Child development*. 79. 476–489 (2008)
34. Coull, J. T., Cheng, R.-K., Meck, W. H.: Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*. 36. 3–25 (2011)
35. Cox, R.: AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Computers and biomedical research, an international journal*. 29. 162–173 (1996)
36. Dabaghian, Y., Brandt, V. L., Frank, L. M.: Reconceiving the hippocampal map as a topological template. *eLife*. 3. e03476 (2014)
37. Davis, S. F.: *Handbook of Research Methods in Experimental Psychology*. pp. 441–469. John Wiley & Sons. 2008.
38. Dehaene, S., Brannon, E.: *Space, Time and Number in the Brain*. pp. 37–186. Academic Press. 2011.
39. Dehaene, S., Dehaene-Lambertz, G., Cohen, L.: Abstract representations of numbers in the animal and human brain. *Trends in neurosciences*. 21. 355–361 (1998)
40. Dewind, N., Brannon, E.: Malleability of the approximate number system: Effects of feedback and training. *Frontiers in human neuroscience*. 6. 68 (2012)
41. Dormal, V., Andres, M., Pesenti, M.: Contribution of the right intraparietal sulcus to numerosity and length processing: An fMRI-guided TMS study. *Cortex; a journal devoted to the study of the nervous system and behavior*. 48. 623–629 (2012)

42. Dormal, V., Dormal, G., Joassin, F., Pesenti, M.: A common right fronto-parietal network for numerosity and duration processing: An fMRI study. *Human brain mapping*. 33. 1490–1501 (2012)
43. Dormal, V., Pesenti, M.: Common and Specific Contributions of the Intraparietal Sulci to Numerosity and Length Processing. *Human brain mapping*. 30. 2466–2476 (2009)
44. Dormal, V., Pesenti, M.: Processing numerosity, length and duration in a three-dimensional Stroop-like task: Towards a gradient of processing automaticity? *Psychological research*. 77. 116–127 (2013)
45. Dormal, V., Seron, X., Pesenti, M.: Numerosity-duration interference: A Stroop experiment. *Acta psychologica*. 121. 109–124 (2006)
46. Droit-Volet, S., Clement, A., Michel, F.: Time, number and length: Similarities and differences in discrimination in adults and children. *Quarterly journal of experimental psychology* (2006). 61. 1827–1846 (2008)
47. Eger, E., Sterzer, P., Russ, M., Giraud, A.-L., Kleinschmidt, A.: A Supramodal Number Representation in Human Intraparietal Cortex. *Neuron*. 37. 719–725 (2003)
48. Eichenbaum, H.: On the Integration of Space, Time, and Memory. *Neuron*. 95. 1007–1018 (2017)
49. Eichenbaum, H.: Time cells in the hippocampus: a new dimension for mapping memories. *Nature reviews. Neuroscience*. 15. 732–744 (2014)
50. Eklund, A., Nichols, T. E., Knutsson, H.: Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*. 201602413 (2016)
51. Ekstrom, A., Kahana, M., Caplan, J., Fields, T., Isham, E., Newman, E., Fried, I.: Cellular Networks underlying human spatial navigation. *Nature*. 425. 184–188 (2003)
52. Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., Gorgolewski, K. J.: MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PloS one*. 12. e0184661 (2017)
53. Esteban, O., Blair, R. W., Nielson, D. M., Varada, J. C., Marrett, S., Thomas, A. G., Poldrack, R. A., Gorgolewski, K. J.: Crowdsourced MRI quality metrics and expert quality annotations for training of humans and machines. *Scientific data*. 6. 30–37 (2019)
54. Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., Gorgolewski, K. J.: fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature methods*. 16. 111–116 (2019)
55. Fechner, G. T.: *Elemente der Psychophysik*. pp. 9–27. Breitkopf und Haertel. 1860.

56. Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., Orban, G.: Parietal Representation of Symbolic and Nonsymbolic Magnitude. *Journal of cognitive neuroscience*. 15. 47–56 (2003)
57. Ficher, H.: *A History of the Central Limit Theorem*. pp. 218–220. Springer. 2011.
58. Fischl, B.: FreeSurfer. *NeuroImage*. 62. 774–781 (2012)
59. Fonov, V., Evans, A., Mckinstry, R., Almlí, C., Collins, L.: Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage*. 47. 313–327 (2009)
60. França, T. F. A., Monserrat, J. M.: Hippocampal place cells are topographically organized, but physical space has nothing to do with it. *Brain Structure and Function*. 224. 3019–3029 (2019)
61. Geary, D., Vanmarle, K.: Young Children’s Core Symbolic and Nonsymbolic Quantitative Knowledge in the Prediction of Later Mathematics Achievement. *Developmental Psychology*. 52. 2130–2144 (2016)
62. Geva-Sagiv, M., Las, L., Yovel, Y., Ulanovsky, N.: Spatial cognition in bats and rats: From sensory acquisition to multiscale maps and navigation. *Nature reviews. Neuroscience*. 16. 94–108 (2015)
63. Giannelli, M., Diciotti, S., Tessa, C., Mascalchi, M.: Characterization of Nyquist ghost in EPI-fMRI acquisition sequences implemented on two clinical 1.5 T MR scanner systems: Effect of readout bandwidth and echo spacing. *Journal of applied clinical medical physics / American College of Medical Physics*. 11. 3237 (2010)
64. Gilron, R., Rosenblatt, J., Koyejo, O., Poldrack, R., Mukamel, R.: What’s in a pattern? Examining the Type of Signal Multivariate Analysis Uncovers At the Group Level. *NeuroImage*. 146. 113–120 (2016)
65. Gorgolewski, K., Burns, C. D., in, C. M. F.: Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*. 5. 13 (2011)
66. Grefkes, C., Fink, G.: REVIEW: The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of Anatomy*. 207. 3–17 (2005)
67. Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., Moser, E. I.: Microstructure of a spatial map in the entorhinal cortex. *Nature*. 436. 801–806 (2005)
68. Hamamouche, K., Cordes, S.: Number, time, and space are not singularly represented: Evidence against a common magnitude system beyond early childhood. *Psychonomic Bulletin and Review*. 26. 833–854 (2019)
69. Hampson, R., Heyser, C., Deadwyler, S.: Hippocampal Cell Firing Correlates of Delayed-Match-to-Sample Performance in the Rat. *Behavioral neuroscience*. 107. 715–739 (1993)

70. Harvey, B. M., Fracasso, A., Petridou, N., Dumoulin, S. O.: Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*. 112. 13525–13530 (2015)
71. Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., Pietrini, P.: Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science (New York, N.Y.)* 293. 2425–2430 (2001)
72. Haxby, J. V., Connolly, A. C., Guntupalli, J. S.: Decoding neural representational spaces using multivariate pattern analysis. *Annual review of neuroscience*. 37. 435–456 (2014)
73. Hayashi, M. J., Kanai, R., Tanabe, H. C., Yoshida, Y., Carlson, S., Walsh, V., Sadato, N.: Interaction of numerosity and time in prefrontal and parietal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 33. 883–893 (2013)
74. Hoffman, K., Babichev, A., Dabaghian, Y.: A model of topological mapping of space in bat hippocampus. *Hippocampus*. 26. 1345–1353 (2016)
75. Hubbard, E. M., Piazza, M., Pinel, P., Dehaene, S.: Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*. 6. 435–448 (2005)
76. Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., Turner, R.: Image Distortion Correction in fMRI: A Quantitative Evaluation. *NeuroImage*. 16. 217–240 (2002)
77. Inglis, B.: A checklist for fMRI acquisition methods reporting in the literature. *The Winnower*. e143191 (2015)
78. Jahanshahi, M., Obeso, I., Rothwell, J., Obeso, J.: A fronto-striato-subthalamic-pallidal network for goal-directed and habitual inhibition. *Nature reviews. Neuroscience*. 16. 719–732 (2015)
79. Jang, S., Cho, S.: The Acuity for Numerosity (but Not Continuous Magnitude) Discrimination Correlates with Quantitative Problem Solving but Not Routinized Arithmetic. *Current Psychology*. 35. 1–13 (2015)
80. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*. 17. 825–841 (2002)
81. Jordan, K., Wüstenberg, T., Heinze, H. J., Peters, M., Jäncke, L.: Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia*. 40. 2397–2408 (2002)
82. Kaufmann, L., Vogel, S., Wood, G., Kremser, C., Schocke, M., Zimmerhackl, L.-B., Koten, J.: A developmental fMRI study of nonsymbolic numerical and spatial processing. *Cortex; a journal devoted to the study of the nervous system and behavior*. 44. 376–385 (2008)

83. Kelleher, J., Mac Namee, B., D'Arcy, A.: *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. pp. 397–463. MIT Press. 2015.
84. Kim, M., Jeffery, K. J., Maguire, E. A.: Multivoxel pattern analysis reveals 3D place information in the human hippocampus. *The Journal of neuroscience*. 2703–2716 (2017)
85. Kingdom, F. A. A., Prins, N.: *Psychophysics: A Practical Introduction: Second Edition*. pp. 55–117. Elsevier. 2016.
86. Klein, A., Ghosh, S., Bao, F., Giard, J., Haeme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., Keshavan, A.: Mindboggling morphometry of human brains. *PLOS Computational Biology*. 13. e1005350 (2017)
87. Kosciessa, J., Lindenberger, U., Garrett, D.: Thalamocortical excitability modulation guides human perception under uncertainty. *Nature Communications*. 12. 24430 (2021)
88. Kragel, P. A., Koban, L., Barrett, L. F., Wager, T. D.: Representation, Pattern Information, and Brain Signatures: From Neurons to Neuroimaging. *Neuron*. 99. 257–273 (2018)
89. Kriegeskorte, N., Goebel, R., Bandettini, P.: Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*. 103. 3863–3868 (2006)
90. Kuss, M., Jäkel, F., Wichmann, F. A.: Bayesian inference for psychometric functions. *Journal of Vision*. 5. 478–492 (2005)
91. Lambrechts, A., Walsh, V., Wassenhove, V. van: Evidence accumulation in the magnitude system. *PloS one*. 8. e82122 (2013)
92. Lewis, P., Miall, R. C.: Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. *Current opinion in neurobiology*. 13. 250–255 (2003)
93. Lipton, J., Spelke, E.: Origins of Number Sense Large-Number Discrimination in Human Infants. *Psychological science*. 14. 396–401 (2003)
94. Livesey, A., Wall, M., Smith, A.: Time perception: Manipulation of task difficulty dissociates clock functions from other cognitive demands. *Neuropsychologia*. 45. 321–331 (2007)
95. Marcus, D., Wang, T., Parker, J., Csernansky, J., Morris, J., Buckner, R.: Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of cognitive neuroscience*. 19. 1498–1507 (2007)
96. Martin, B., Wiener, M., Wassenhove, V. van: A Bayesian Perspective on Accumulation in the Magnitude System. *Scientific Reports*. 7. 630 (2017)

97. Martinez-Trujillo, J., Cheyne, D., Gaetz, W., Simine, E., Tsotsos, J.: Activation of Area MT/V5 and the Right Inferior Parietal Cortex during the Discrimination of Transient Direction Changes in Translational Motion. *Cerebral cortex*. 17. 1733–1739 (2007)
98. Martinez-Trujillo, J., Tsotsos, J., Simine, E., Pomplun, M., Wildes, R., Treue, S., Heinze, H.-J., Hopf, J.-M.: Selectivity for speed gradients in human area MT/V5. *Neuroreport*. 16. 435–438 (2005)
99. Matell, M., Meck, W.: Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research*. 21. 139–170 (2004)
100. Meck, W., Church, R.: A mode control model of counting and timing processes. *Journal of Experimental Psychology. Animal Behavior Processes*. 9. 320–334 (1983)
101. Meck, W., Penney, T., Pouthas, V.: Cortico-striatal representation of time in animals and humans. *Current Opinion in Neurobiology*. 18. 145–152 (2008)
102. Moser, E. I., Moser, M.-B., McNaughton, B. L.: Spatial representation in the hippocampal formation: a history. *Nature Publishing Group*. 20. 1448–1464 (2017)
103. Muggleton, N., Postma, P., Moutsopoulou, K., Nimmo-Smith, I., Marcel, A., Walsh, V.: TMS over right posterior parietal cortex induces neglect in a scene-based frame of reference. *Neuropsychologia*. 44. 1222–1229 (2006)
104. Mumford, J. A., Turner, B. O., Ashby, F. G., Poldrack, R. A.: Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*. 59. 2636–2643 (2012)
105. Muschelli, J., Nebel, M. B., Caffo, B. S., Barber, A. D., Pekar, J. J., Mostofsky, S. H.: Reduction of motion-related artifacts in resting state fMRI using aCompCor. *NeuroImage*. 96. 22–35 (2014)
106. Mussolin, C., De Volder, A., Grandin, C., Schloegel, X., Nassogne, M.-C., Noel, M.-P.: Neural Correlates of Symbolic Number Comparison in Developmental Dyscalculia. *Journal of cognitive neuroscience*. 22. 860–874 (2009)
107. Myung, I. J.: Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*. 47. 90–100 (2003)
108. Newcombe, N.: The Origins and Development of Magnitude Estimation. *Ecological Psychology*. 26. 147–157 (2014)
109. Nichols, T., Holmes, A.: SnPM. Version 8c40ad237ccd12ed074124a29b1bf2095f872841. url: <https://github.com/SnPM-toolbox/SnPM-devel>
110. Nichols, T. E., Holmes, A. P.: Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*. 15. 1–25 (2002)
111. O'Keefe, J.: Place units in the hippocampus of the freely moving rat. *Experimental neurology*. 4. 1363 (1976)

112. Odic, D.: Children's intuitive sense of number develops independently of their perception of area, density, length, and time. *Developmental science*. 21. e12533 (2017)
113. Odic, D., Libertus, M., Feigenson, L., Halberda, J.: Developmental Change in the Acuity of Approximate Number and Area Representations. *Developmental psychology*. 49. 1103–1112 (2012)
114. Odic, D., Starr, A.: An Introduction to the Approximate Number System. *Child development perspectives*. 12. 223–229 (2018)
115. Olivetti, E., Veeramachaneni, S., Nowakowska, E.: Bayesian hypothesis testing for pattern discrimination in brain decoding. *Pattern Recognition*. 45. 2075–2084 (2012)
116. Pastalkova, E., Itskov, V., Amarasingham, A., Buzsaki, G.: Internally Generated Cell Assembly Sequences in the Rat Hippocampus. *Science (New York, N.Y.)* 321. 1322–1327 (2008)
117. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É.: Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 12. 2825–2830 (2011)
118. Pereira, F., Mitchell, T., Botvinick, M.: Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage*. 45. 199–209 (2009)
119. Piazza, M., Izard, V., Pinel, P., Le Bihan, D., Dehaene, S.: Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. *Neuron*. 44. 547–555 (2004)
120. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team: nlme: Linear and Nonlinear Mixed Effects Models. Version 3.1-142. url: <https://CRAN.R-project.org/package=nlme>
121. Power, J., Barnes, K., Snyder, A., Schlaggar, B., Petersen, S.: Spurious but systematic conditions in functional connectivity MRI networks arise from subject motion. *Neuroimage*. 59. 2141–2154 (2012)
122. R Core Team: R: A Language and Environment for Statistical Computing. Version 3.6.2. url: <https://www.R-project.org/>
123. Riemer, M.: Psychophysics and the anisotropy of time. *Consciousness and cognition*. 38. 191–197 (2015)
124. Riemer, M., Shine, J. P., Wolbers, T.: On the (a)symmetry between the perception of time and space in large-scale environments. *Hippocampus*. 28. 539–548 (2018)
125. Rissman, J., Gazzaley, A., D'Esposito, M.: Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*. 23. 752–763 (2004)
126. Robinson, G.: Biasing power law exponents by magnitude estimation instructions. *Perception and Psychophysics*. 19. 80–84 (1976)

127. Ryali, S., Supekar, K., Abrams, D., Menon, V.: Sparse logistic regression for whole-brain classification of fMRI data. *NeuroImage*. 51. 752–764 (2010)
128. Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., Ranganath, C.: Memory and Space: Towards an Understanding of the Cognitive Map. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 35. 13904–13911 (2015)
129. Schütt, H.: psignifit 4. Version 3013da0. url: <https://github.com/wichmann-lab/psignifit>
130. Schütt, H., Harmeling, S., Macke, J. H., Wichmann, F. A.: Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision research*. 122. 105–123 (2016)
131. Shehzad, Z., Giavasis, S., Li, Q., Benhajali, Y., Yan, C., Yang, Z., Milham, M., Bellec, P., Craddock, C.: The Preprocessed Connectomes Project Quality Assessment Protocol - a resource for measuring the quality of MRI data. *Frontiers in Neuroscience*. 1 (2015)
132. Sherry, D., Grella, S., Guigueno, M., White, D., Marrone, D.: Are There Place Cells in the Avian Hippocampus? *Brain Behavior and Evolution*. 90. 73–80 (2017)
133. Shine, J., Valdes-Herrera, J., Hegarty, M., Wolbers, T.: The Human Retrosplenial Cortex and Thalamus Code Head Direction in a Global Reference Frame. *Journal of Neuroscience*. 36. 6371–6381 (2016)
134. Shine, J., Valdes-Herrera, J., Tempelmann, C., Wolbers, T.: Evidence for allocentric boundary and goal direction information in the human entorhinal cortex and subiculum. *Nature Communications*. 10. 4004 (2019)
135. Skagerlund, K., Karlsson, T., Träff, U.: Magnitude processing in the brain: An fMRI study of time, space, and numerosity as a shared cortical system. *Frontiers in human neuroscience*. 10. 500 (2016)
136. Slugocki, M.: BayesFit: A tool for modeling psychophysical data using Bayesian inference. *Journal of open research software*. 7. 2 (2019)
137. Solstad, T., Boccara, C., Kropff, E., Moser, M.-B., Moser, E.: Representation of Geometric Borders in the Entorhinal Cortex. *Science (New York, N.Y.)* 322. 1865–1868 (2009)
138. Srinivasan, M., Carey, S.: The long and the short of it: On the nature and origin of functional overlap between representations of space and time. *Cognition*. 116. 217–241 (2010)
139. Stelzer, J., Chen, Y., Turner, R.: Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *NeuroImage*. 65. 69–82 (2013)
140. Stevens, S. S.: To Honor Fechner and Repeal His Law. *Science*. 133. 80–86 (1961)

141. Taube, J., Muller, R., Ranck Jr, J.: Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 10. 420–435 (1990)
142. Teghtsoonian, M., Teghtsoonian, R.: How repeatable are Stevens's power law exponents for individual subjects? *Perception and Psychophysics*. 10. 147–149 (1971)
143. Todd, M., Nystrom, L., Cohen, J.: Confounds in Multivariate Pattern Analysis: Theory and Rule Representation Case Study. *NeuroImage*. 77. 157–165 (2013)
144. Tolman, E. C.: Cognitive maps in rats and men. *Psychological review*. 55. 189–208 (1948)
145. Tregellas, J., Davalos, D., Rojas, D.: Effect of task difficulty on the functional anatomy of temporal processing. *NeuroImage*. 32. 307–315 (2006)
146. Turner, B. O., Mumford, J. A., Poldrack, R. A., Ashby, F. G.: Spatiotemporal activity estimation for multivoxel pattern analysis with rapid event-related designs. *NeuroImage*. 62. 1429–1438 (2012)
147. Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J.: N4ITK: improved N3 bias correction. *Medical Imaging, IEEE Transactions on*. 29. 1310–1320 (2010)
148. Vanmarle, K., Wynn, K.: Six-month-old infants use analog magnitudes to represent duration. *Developmental science*. 9. 41–49 (2006)
149. Vasilyeva, M.: Development of spatial cognition. *WIREs Cognitive Science*. 349–362 (2012)
150. Vasilyeva, M., Lourenco, S. F.: Spatial Development. In: Lerner, R., Overton, W., Freund, A., Lam, M. (eds.) *The Handbook of Lifespan Development*. Vol. 1. pp. 720–753. American Cancer Society 2010.
151. Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, I., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P.: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*. 261–272 (2020)
152. Walsh, V.: A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences*. 7. 483–488 (2003)
153. Wang, Q., Cagna, B., Chaminade, T., Takerkart, S.: Inter-subject pattern analysis: A straightforward and powerful scheme for group-level MVPA. *NeuroImage*. 204. 116205 (2020)
154. Welvaert, M., Rosseel, Y.: On the Definition of Signal-To-Noise Ratio and Contrast-To-Noise Ratio for fMRI Data. *PloS one*. 8. 1–10 (2013)

155. Whitlock, J. R., Sutherland, R. J., Witter, M. P., Moser, M.-B., Moser, E. I.: Navigating from hippocampus to parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*. 105. 14755–14762 (2008)
156. Wiener, S., Taube, J.: *Head Direction Cells and the Neural Mechanisms of Spatial Orientation*. pp. 7–248. MIT Press. 2005.
157. Xu, F., Spelke, E.: Large number discrimination in 6-month olds. *Cognition*. 74. 1–11 (2000)
158. Xuan, B., Zhang, D., He, S., Chen, X.: Larger Stimuli Are Judged to Last Longer. *Journal of vision*. 7. 1–5 (2007)
159. Zeki, S.: Area V5-a microcosm of the visual brain. *Frontiers in Integrative Neuroscience*. 9. 21 (2015)
160. Zhang, Y., Brady, M., Smith, S.: Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the Expectation Maximization Algorithm. *IEEE transactions on medical imaging*. 20. 45–57 (2001)
161. Zhu, C., Byrd, R., Lu, P., Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization. *ACM Trans. Math. Softw.* 23. 550–560 (1997)

Ehrenerklärung

Ich erkläre, dass ich die der Medizinischen Fakultät der Otto-von-Guericke-Universität zur Promotion eingereichte Dissertation mit dem Titel

Accumulation processes of time, space and numerosity: an fMRI study

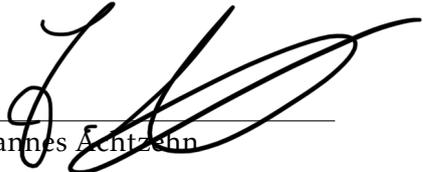
im Zentrum für Neurodegenerative Erkrankungen e.V. Magdeburg,
mit Unterstützung durch Prof. Dr. Thomas Wolbers und Dr. Martin Riemer,

ohne sonstige Hilfe durchgeführt und bei der Abfassung der Dissertation keine anderen als die dort aufgeführten Hilfsmittel benutzt habe. Bei der Abfassung der Dissertation sind Rechte Dritter nicht verletzt worden.

Ich habe diese Dissertation bisher an keiner in- oder ausländischen Hochschule zur Promotion eingereicht. Ich übertrage der Medizinischen Fakultät das Recht, weitere Kopien meiner Dissertation herzustellen und zu vertreiben.

Berlin, 30.06.2022

Ort, Datum

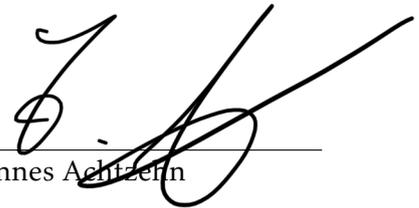

Johannes Achtzehn

Erklärung zur strafrechtlichen Verurteilung

Ich erkläre hiermit, nicht wegen einer Straftat verurteilt worden zu sein, die wissenschaftsbezug hat.

Berlin, 30.06.2022

Ort, Datum

A handwritten signature in black ink, consisting of a stylized 'J' and 'A' with a long horizontal stroke extending to the right.

Johannes Achitzern

Danksagung

Die Danksagung ist in dieser Version aus Datenschutzgründen nicht enthalten.

Veröffentlichung

Teile dieser Arbeit wurden unter dem Titel

Cross-dimensional interference between time and distance during spatial navigation is mediated by speed representations in intraparietal sulcus and area hMT+

durch folgende Autor:innen

Martin Riemer^{1,2,3,4,†}, Johannes Achtzehn^{1,2,†}, Esther Kuehn^{2,3}, Thomas Wolbers^{1,3}

[†]geteilte Erstautorenschaft

¹Aging and Cognition Research Group, German Center for Neurodegenerative Diseases (DZNE), 39120 Magdeburg, Germany

²Medical Faculty, Otto-von-Guericke University, 39106 Magdeburg, Germany

³Center for Behavioral Brain Sciences (CBBS), 39106 Magdeburg, Germany

⁴Biological Psychology and Neuroergonomics, Technical University Berlin, 10623 Berlin, Germany

in der Zeitschrift

NeuroImage, Volume 257, 2022, Article 119336

unter der DOI [10.1016/j.neuroimage.2022.119336](https://doi.org/10.1016/j.neuroimage.2022.119336) veröffentlicht.

Lebenslauf

Der Lebenslauf ist in dieser Version aus Datenschutzgründen nicht enthalten.

PART V:
Appendix

A. Methods and Materials

Table A.1: Scanning parameters of anatomical and functional (BOLD) MRI sequences.

Parameter	T1w	T2w	BOLD
In-plane spatial encoding			
Pulse sequence type	MPRAGE	TSE	EPI
PE acceleration factor	2	2	2
PE partial Fourier scheme	224/256	512/512	72/72
Echo spacing	-	-	0.245 ms
Timing Parameters			
TR	2.5 s	6 s	2 s
TE	2.82 ms	73 ms	30 ms
No. of volumes	-	-	500 (per run)
No. of dummy scans	-	-	5 (per run)
Spatial Parameters			
In-plane matrix	256 × 256	512 × 512	72 × 72
No. of slices	192	60	36
FOV	256 mm	256 mm	216 mm
Slice thickness	1 mm	2 mm	3 mm
Inter-slice gap	1 mm	2 mm	3.3 mm
Slice acquisition order	contiguous	contiguous	interleaved
Slice acceleration			
SMS acceleration factor	-	-	1
Sequence name	tfl3d1_16ns	tse2d1_11	epfid2d1_72

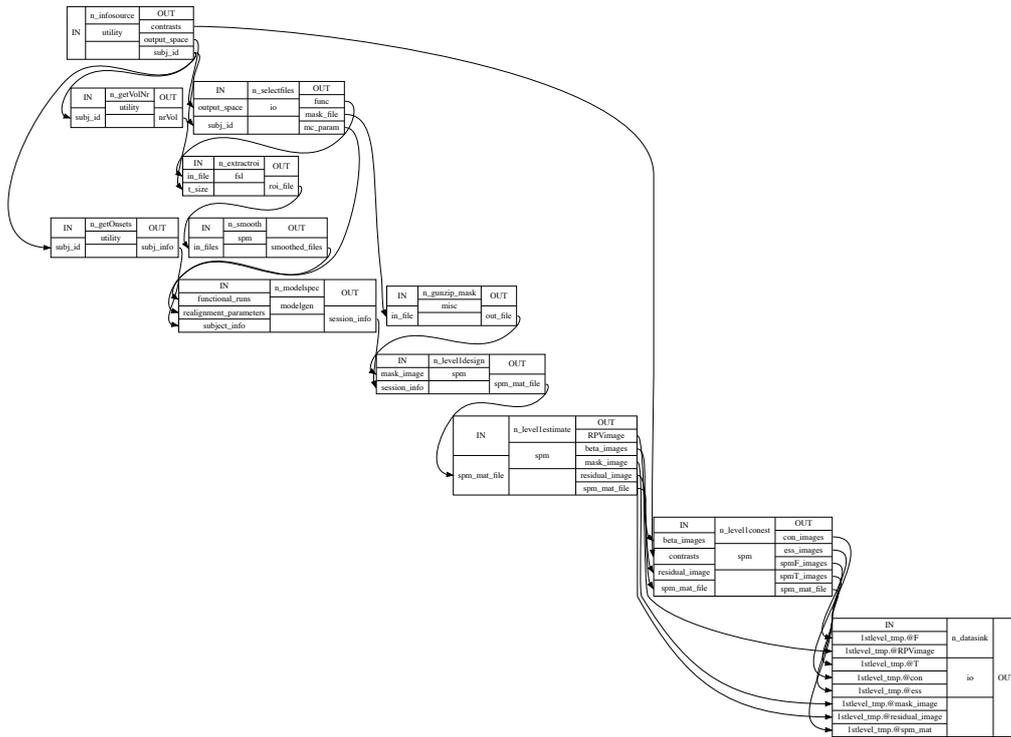


Fig. A.1: Detailed workflow for the univariate analysis.

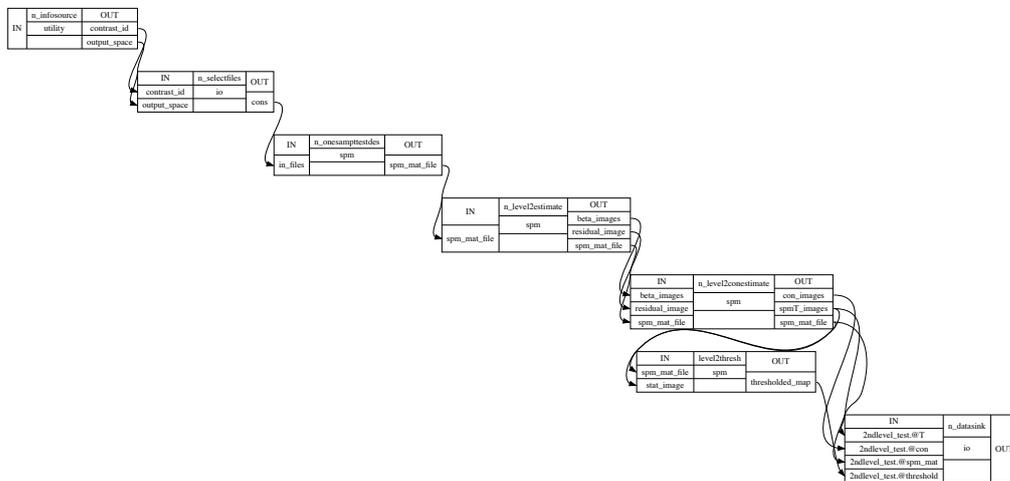


Fig. A.2: Detailed workflow for the second level analysis.

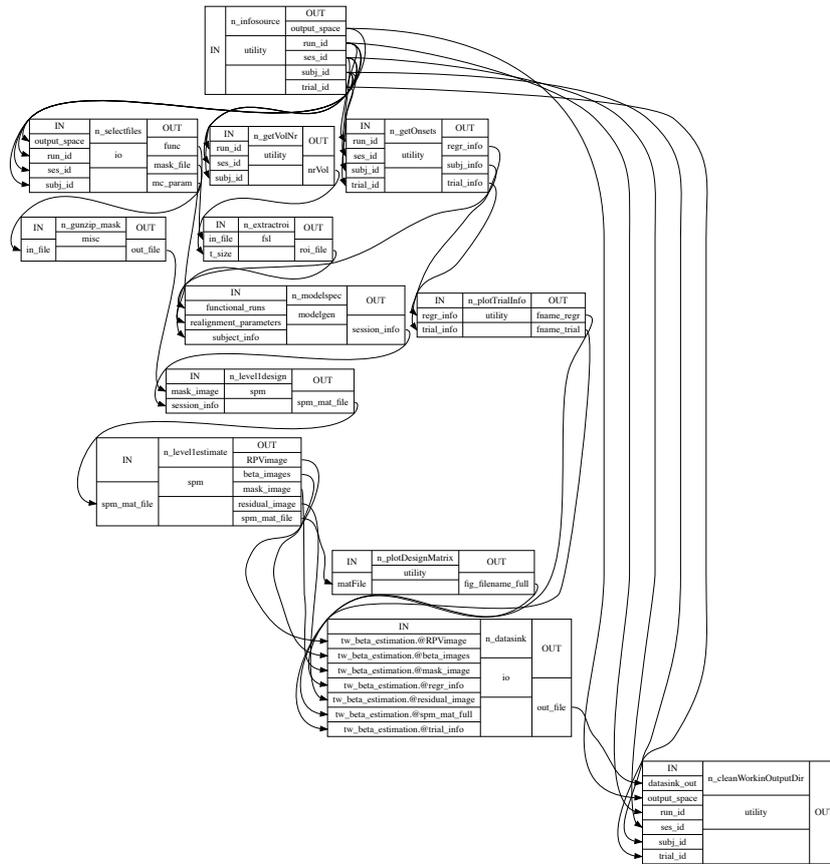


Fig. A.3: Detailed workflow for the trialwise beta map estimation utilised to create decoding samples.

B. Behavioural Results

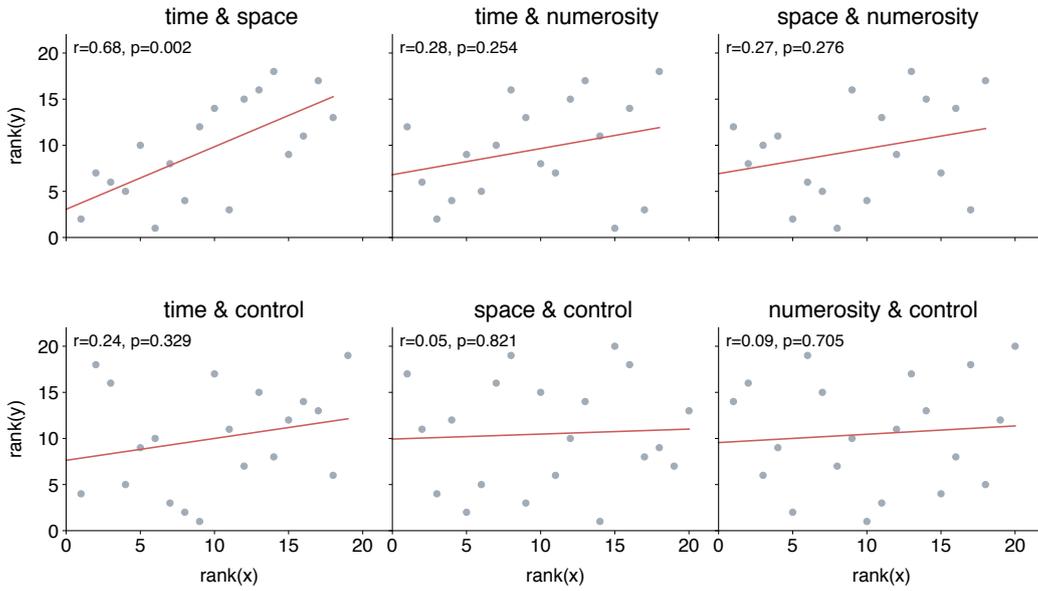


Fig. B.1: Spearman rank plots for all modality pairs comparing the judgement precision.

Table B.1: Cross-dimensional interference on magnitude perception for all possible combinations of dimensions. Numerical values are normalised (to the mean of the current dimension reference values) mean differences in thresholds. Positive numbers indicate higher PSE values if the irrelevant dimension was in high condition and vice versa. Significance was tested with a Wilcoxon signed-rank test and corrected for multiple comparisons (FWE = 0.05).

Relevant dim.	Irrelevant dim.	Difference in mean	W	p-value
Time	Space	0.177	36	0.006
Time	Numerosity	-0.052	106	0.607
Space	Time	-0.139	45	0.013
Space	Numerosity	0.048	89	0.608
Numerosity	Time	0.016	115	0.778
Numerosity	Space	0.015	130	0.5677

C. Neuroimaging Results

C.1 Quality Control

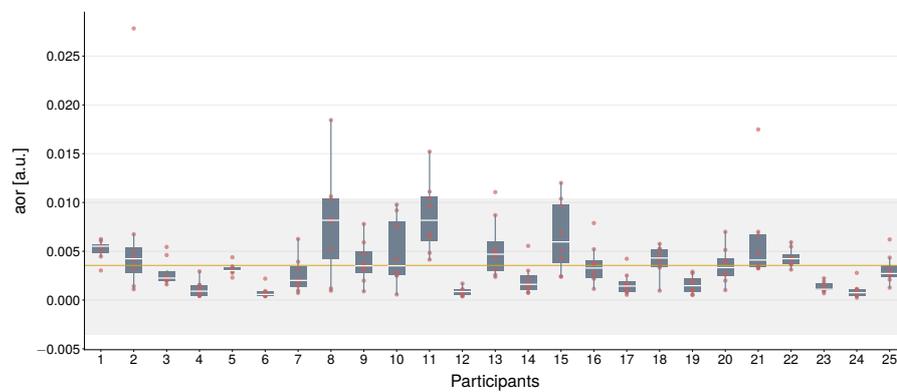


Fig. C.1: Mean fraction of outliers per fMRI volume as given by AFNI's *3dToutcount* (AOR) for all participants.

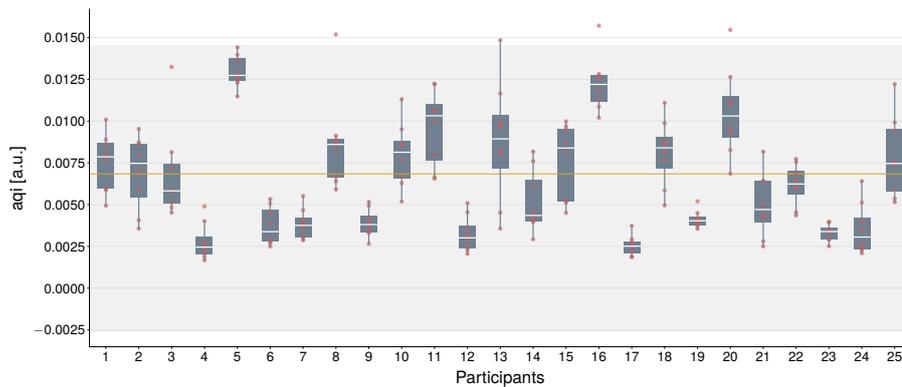


Fig. C.2: Mean quality index as computed by AFNI's *3dTqual*.

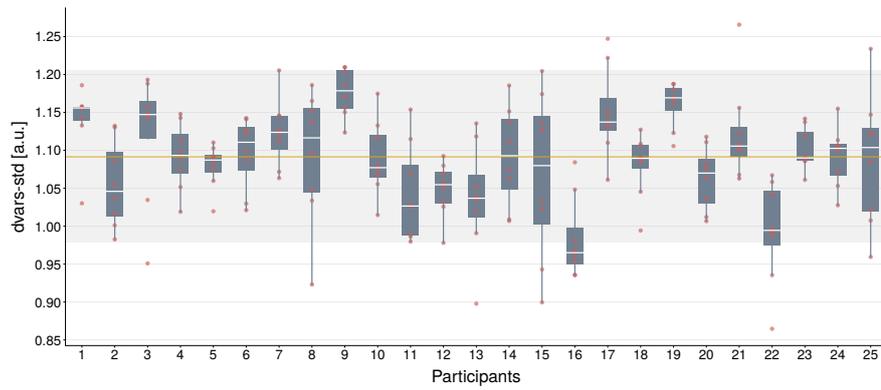


Fig. C.3: DVARS for all participants: D referring to temporal derivative of timecourses, VARS referring to RMS variance over voxels.

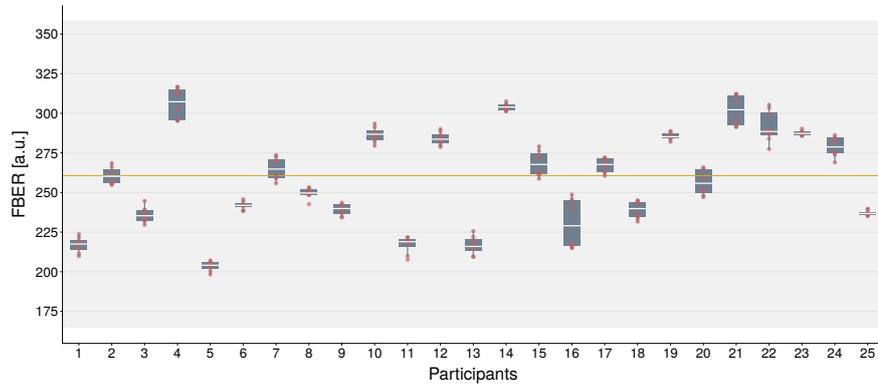


Fig. C.4: Foreground-Background energy ratio (FBER) for all participants.

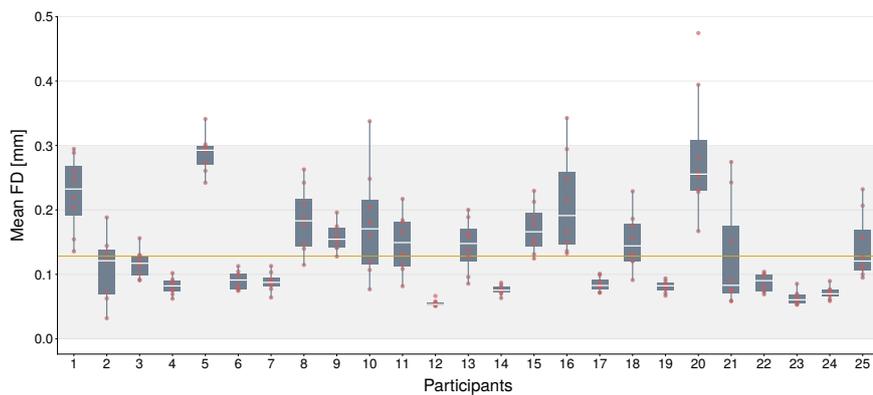


Fig. C.5: Mean fieldwise displacement for all participants.

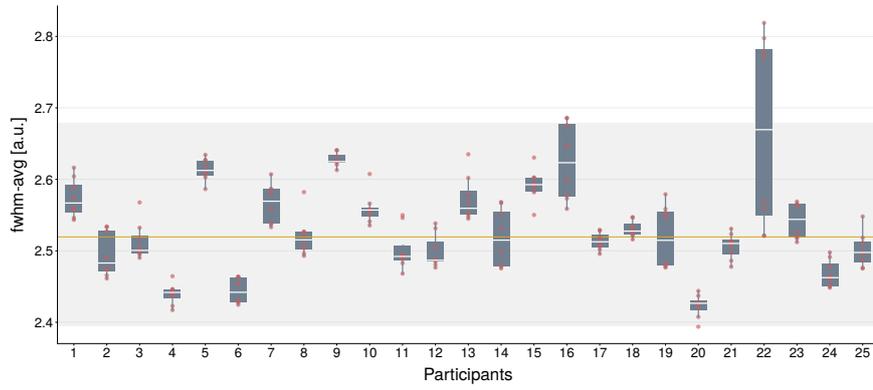


Fig. C.6: Full-width half maximum smoothness for all participants.

C.2 Second Level Analysis

Table C.1: Clusters identified for contrast 1: average activation during accumulation trials. The value of cluster size indicates the number of voxels. Coordinates represent peak level activation. For larger clusters spanning multiple regions, additional maxima have been listed in grey color.

Nr.	Anatomical region	Cluster size	T	p	MNI coordinates
1	R fusiforme gyrus	257	16.75	<0.001	(39, -75, -15)
2	R posterior medial frontal gyrus	409	14.50	<0.001	(3, 15, 47)
	L posterior medial frontal gyrus		13.52	<0.001	(-6, 9, 54)
	R middle cingulate cortex		10.36	<0.001	(9, 24, 41)
3	L inferior frontal gyrus	162	13.81	<0.001	(-30, 27, 1)
4	L intraparietal sulcus	821	13.09	<0.001	(-45, -48, 44)
	L superior parietal lobe		12.85	<0.001	(-30, -66, 47)
5	R inferior frontal gyrus	166	12.87	<0.001	(33, 24, -5)
6	R intraparietal sulcus	255	12.07	<0.001	(30, -69, 51)
7	R cuneus	473	11.72	<0.001	(12, -93, 21)
	L calcarine gyrus		11.50	<0.001	(-3, -93, 8)
	R occipital gyrus		8.20	0.002	(15, -81, -2)
8	L precentral gyrus	160	10.27	<0.001	(-42, 6, 34)
9	L fusiform gyrus	102	10.26	<0.001	(-36, -66, -12)

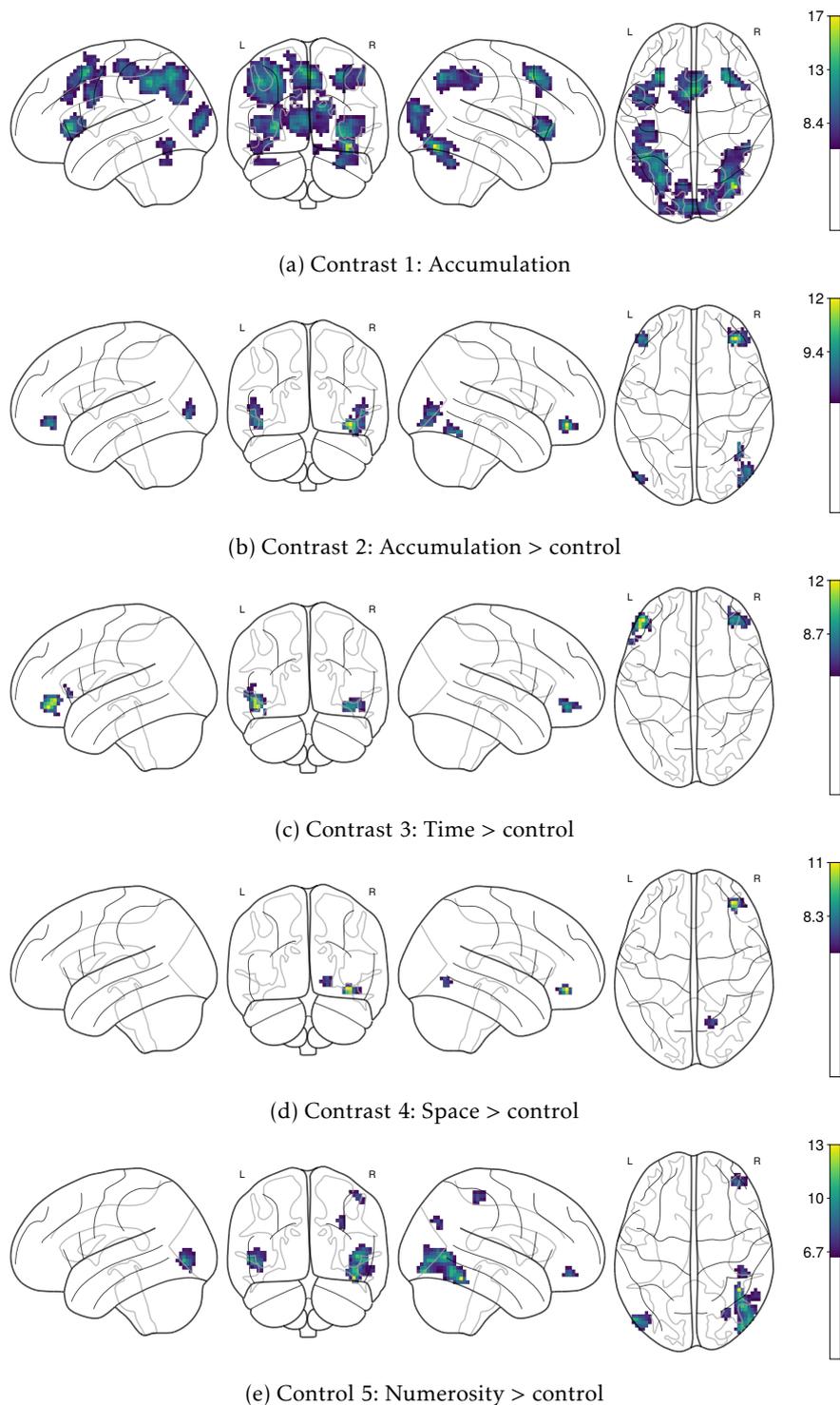


Fig. C.7: Glass brain plots for the second level analysis contrasts 1 - 5, as defined in table 7.1.

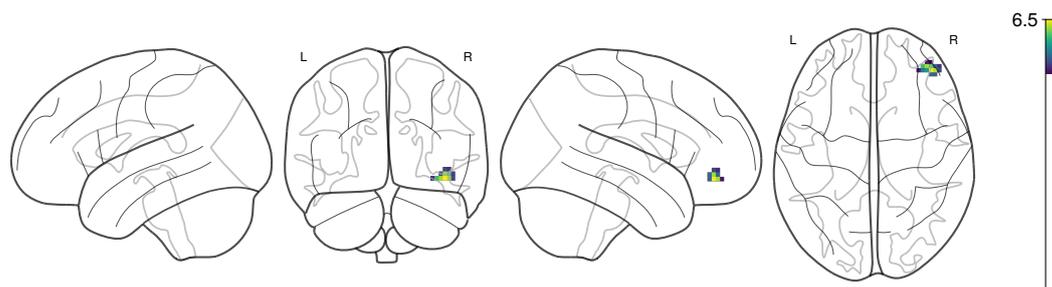
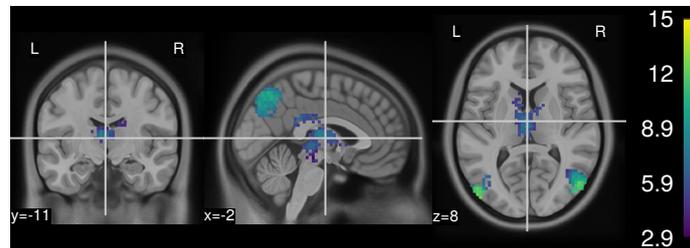
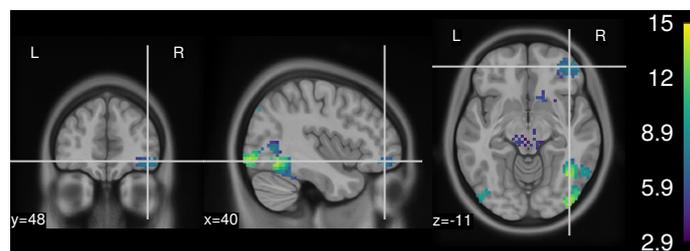


Fig. C.8: Glass brain plots for the conjunction analysis (contrast 6), as defined in table 7.1.

C.3 Multivariate Pattern Analysis



(a) Thalamic cluster



(b) IFG cluster

Fig. C.9: Exemplary plot for two clusters that show significant decoding accuracy across all three modalities when decoded against the control condition. Here, data for the pair space vs. control is shown. Values are percentages above chance level of 50%.

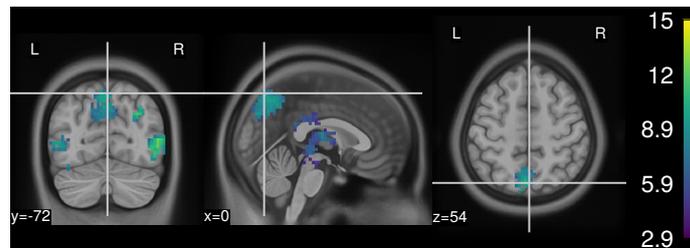


Fig. C.10: Accuracy map for the cluster that exhibits significant decoding accuracy for both space vs. control and numerosity vs. control. Values are percentages above chance level of 50%.