



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2021

INNOVATION AND STANDARDIZATION OF PROCESSING PIPELINES FOR FUNCTIONAL MRI DATA ANALYSIS

Notter Michael P.

Notter Michael P., 2021, INNOVATION AND STANDARDIZATION OF PROCESSING
PIPELINES FOR FUNCTIONAL MRI DATA ANALYSIS

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_62B5C2D0DD7F3

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Service de radiodiagnostic et radiologie interventionnelle

**INNOVATION AND STANDARDIZATION OF PROCESSING
PIPELINES FOR FUNCTIONAL MRI DATA ANALYSIS**

Thèse de doctorat en Neurosciences

présentée à la

Faculté de Biologie et de Médecine
de l'Université de Lausanne

par

MICHAEL P. NOTTER

Maîtrise de sciences en psychologie cognitive et neurosciences
de l'Université de Zurich, Suisse

Jury

Prof. Jean-Philippe Thiran, Président
Prof. Micah M. Murray, Directeur de thèse
Prof. Dr. Michael Hanke, Co-Directeur de thèse
Prof. Dimitri Van De Ville, Expert
Prof. Henning Müller, Expert

Thèse n° 311

Lausanne 2021

*Programme doctoral interuniversitaire en Neurosciences
des Universités de Lausanne et Genève*



UNIL | Université de Lausanne



**UNIVERSITÉ
DE GENÈVE**

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Jean-Philippe	Thiran
Directeur·trice de thèse	Monsieur	Prof.	Micah M.	Murray
Co-Directeur·trice de thèse	Monsieur	Prof.	Michael	Hanke
Expert·e·s	Monsieur	Prof.	Dimitri	Van De Ville
	Monsieur	Prof.	Henning	Müller

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Michael Notter

Master of Science in Psychology,
University of Zürich, Switzerland

intitulée

Innovation and standardization of processing pipelines for functional MRI data analysis

Date de l'examen: 1 juillet 2021

Date d'émission de l'Imprimatur: Lausanne, le 2 août 2021

pour Le Doyen
de la Faculté de Biologie et de Médecine



Prof. Niko GELDNER
Directeur de l'Ecole Doctorale

Acknowledgements

These pages are dedicated to everyone who supported me throughout and helped me to obtain my PhD!

First, I would like to express my sincerest gratitude to Professor Micah M. Murray and thank him for everything he has done for me. Already before supervising my PhD, he was guiding me through the academic world and gave me the great opportunity to work in his dynamic and ever evolving research group. From the beginning, he provided me with many great opportunities in Switzerland and abroad, which led to many important experiences and helped me to grow as a researcher and as a person. From the very start, Prof. Murray made sure that I can work in a fruitful and stimulating environment, and gave me all the freedom and trust I need to work on the things important and interesting to me.

In equal measures, I would like to express my sincerest gratitude and thanks to my thesis co-director Prof. Michael Hanke. Even before my PhD started, Prof. Hanke welcomed me with open hands and mind in his lab as an exchange student and provided everything I needed to learn, feel welcomed and have a stimulating and great time. I cannot thank him enough for the many stimulating and thought-provoking discussions we had through the years. Interacting with Prof. Hanke and learning from his expertise was a great experience and fundamentally changed my opinion of neuroimaging and science in general. Being exposed to his mindset certainly changed mine and started my path as a data scientist. I wish him only the best in his future endeavor and hope that he goes as far as possible.

Besides my thesis directors, I would also like to thank Professor Dimitri Van De Ville, Professor Henning Müller and Professor Jean-Philippe Thiran for accepting to be part of my Jury.

I cannot thank enough Dr. Sandra Da Costa who helped me in all facets of my PhD. Her exhaustive knowledge in neuroscience led to very stimulating discussion, her detailed expertise in neuroimaging was crucial for all of my projects, her continuous presence and support made it possible to pilot and record many different datasets and her friendship made all of it a lot of fun! Without her, this thesis would not have been possible.

I also want to thank Dr. Eveline Geiser, who introduced me to the field of neuroimaging and gave me twice a great opportunity that decisively changed my future. Without her, my life would have evolved much different. I am forever grateful for that it did not.

My academic path that brought me to the end of my PhD was long and winding. I have met a lot of interesting and amazing people and learned from all of them:

A very special thank goes out to Dr. Peer Herholz from the Philipps-University Marburg in Germany and the Montréal Neurological Institute in Canada. With a very similar background and interest for the same things, Peer was my closest work colleague during my thesis and I have spent a lot of amazing moments with him. His friendship, his insights, his discussion about code, topics and everything else were always stimulating, always interesting and always fun. I wish him for his future endeavor only the very best.

A very special thank goes also out to Dr. Kirstie Whitaker from the Alan Turing Institute in London. Her support, motivation and enthusiasm for the field were always very inspiring. Interacting with her was always very inspiring and mind opening. Her mentoring and input were crucial for many important crossroads on my path.

For their friendship, and many great discussions, I want to thank Ayse Ilkay Isik and Dr. Omer Faruk Gulban. I have learned a lot from both of them in respect to science, programming and being a great person in general.

For their great support, friendship and discussion during my time at CHUV, I would like to thank Dr. Rosanna Monteil, Dr. Jean-François Knebel, Dr. Ulrike Toepel, Dr. Camille Crézé, Dr. Sonia Crottaz-Herbette, Dr. Solange Denervaud, Dr. Emeline Mullier, David Zeuglin, Astrid Minier, Paola Suarez Ramirez, Laurence Wahl, Dr. Jacques Anken, Dr. Silvio Ionta, Dr. David Perruchoud, Dr. Céline Richard, Dr. Eleonora Fornari, Dr. Meritxell Bach Cuadra, Dr. Sébastien Tourbier, Dr. Tiffany Grisendi, Dr. Isabel Tissieres, Prof. Aki Kawasaki and Jean-Baptiste Ledoux. My personal gratitude and a special thanks to Prof. Stephanie Clarke and Prof. Reto Meuli who welcomed me in their team or department and made it possible for me to join and stay at CHUV throughout most of my time.

What makes my PhD especially rewarding and interesting were all the amazing people I got to know from around the world. Their presence, support, curiosity, enthusiasm and friendship made my PhD an amazing experience and is something I will always cherish. From Denmark, Dr. Mehdi Ordikhani-Seyedlar and Dr. Karoline Doser. From the Massachusetts Institute of Technology in the United States, Prof. Satrajit Ghosh, Dr. Dorota Jarecka and Prof. John Gabrieli. From Stanford University in the United States, Dr. Krzysztof Gorgolewski, Dr. Oscar Esteban and Chris Markiewicz. From the Inria and NeuroSpin in France, Dr. Gaël Varoquaux. From the University of Cambridge in the United Kingdom, Dr. Olaf Hauk and Dr. Johan Carlin. From the Montreal Neurological Institute in Canada, Ross Markello. From the University of Zürich in Switzerland, Dr. Franz Liem and Dr. Jürgen Hänggi. From Queen's University in Canada, Daniel Gale. From the University of Washington in the United States, Dr. Ariel Rokem. And everybody else from the Neurohackademy 2018 community.

From my non-academic life, I want to especially thank my friend Stephanie Scherrer who supported and encouraged me on my path since my Bachelor. Her faith and trust in me and my skillset brought me to this point and taught me to believe in myself and to never give up. A very special thanks also to my friends Sonja Schneebeli, Carole Wenger and Aude Vogt, Clement Per and Nathanaël Coquoz, for their never-ending support, friendship and motivation.

A very special thanks goes also out to my family and family in-law. For their endless and never tiring support throughout all my studies, for allowing me to realize my dreams and for providing me the environment and tools to do so. Without their help, nothing of this would have been possible.

More important than any other person mentioned so far, I want to thank the most amazing person in the world and my better half, Marie-Laure Notter. Without her constant love, support, motivation and care, this all would have not been possible. Her endless trust and faith in me helped me to become a better person and I'm excited to see in which direction our future paths are leading us.

Last but certainly not least, I want to thank my daughter Isabel Notter, for bringing my life constant sunshine, joy, wonder, excitement, adventure, laughter and meaning!

Abstract (English)

Manual preprocessing and analyzing of functional magnetic resonance imaging (fMRI) datasets can be a cumbersome endeavor. Using readily available processing pipelines can help with this, but such an approach bears many risks.

This thesis first describes the nature of event related fMRI datasets, what it measures and how such data can be preprocessed and analyzed. After listing the main standard neuroimaging toolboxes used to process event-related fMRI datasets, this thesis describes the multiple issues connected with the current processing approaches and the general challenges the field tries to tackle to innovation in technology and methodology.

The four *issues* identified in this thesis are: (1) inaccessibility and stickiness of neuroimaging toolboxes, (2) missing general standards for neuroimaging analyses, (3) a reproducibility and transparency crisis and (4) insufficient data quality control and results reporting. The three *challenges* identified in this thesis are: Due to innovation (1) in the spatial dimension, (2) in the temporal dimension and (3) in signal processing.

The first study (Study A) tackles most of these issues and challenges by introducing a new neuroimaging toolbox, called *fMRIflows*. This toolbox is a consortium of fully automatic processing pipeline capable of performing state-of-the-art data preprocessing, as well as first- and second level univariate and multivariate analysis. Validation of the toolbox is done by analyzing three different fMRI datasets with different temporal resolution (i.e. 2000ms, 1000ms and 600ms) and comparing the output created with fMRIflows to the output created with state-of-the-art neuroimaging software packages fMRIPrep, FSL and SPM. The validation shows that no strong difference between the output of the four toolboxes can be observed. Furthermore, the study shows that an adequate temporal filtering of an fMRI dataset with a sub-second temporal resolution can lead to improved temporal signal-to-noise-ratio (TSNR) after preprocessing and an increased statistical sensitivity in the 1st and 2nd level analysis.

The second study (Study B) tackles many of the beforementioned issues, but focuses on the 4th (i.e. results reporting) in particular, by introducing a new neuroimaging toolbox, called *AtlasReader*. This toolbox can be used to generate coordinate tables, region labels and informative figures from statistical MRI images. Study B successfully introduces a neuroimaging toolbox that allows to create beautiful and informative results reports, independent on the operating system of the user. Furthermore, AtlasReader allows the extraction of association tables from multiple atlases which usually are not accessible to a single operating system.

The third study (Study C), uses these the toolboxes developed in Study A and Study B and show their application in a cognitive neuroscience study in the domain of multisensory integration. Study investigated the brain mechanisms involved during the encoding and subsequent retrieval of semantically congruent multisensory objects. In this study we found that the low-level visual cortex reliably can decode whether an incoming visual stimulus previously had been perceived in a semantically congruent or incongruent context, even if the visual stimuli was only perceived once before. The results from this study further support the notion that the low-level visual cortex has multisensory architecture and that the creation of memories profits from a multisensory semantic congruent stimuli exposure.

As a next step, the thesis assesses the first two studies with respect to the previously mentioned issues and challenges and the third with regards of novelty and new insights gained from this cognitive study. Followed by a critical assessment with regards to the studies limitations. After that, the future directions of fMRI processing analysis routines and of fMRI-based investigation of multisensory integration is discussed. The thesis is finished with a prospect section of what might come and the general conclusion.

Together, the studies comprised in this thesis highlight and address the issues and challenges currently present in the neuroimaging domain and provide a path forward. Furthermore, the findings and outcomes of all three studies contribute to a better understanding of how to correctly preprocess and analyze fMRI datasets, as well as how the mechanism behind multisensory integration takes place.

Résumé (French)

Le pré-traitement et l'analyse manuels des ensembles de données d'imagerie par résonance magnétique fonctionnelle (IRMf) peuvent s'avérer fastidieux. L'utilisation de pipelines de traitement facilement accessibles peut aider, mais une telle approche comporte de nombreux risques.

Cette thèse décrit tout d'abord la nature des ensembles de données évoquées d'IRMf, ce qu'elles mesurent et comment ces données peuvent être pré-traitées et analysées. Après avoir énuméré les principales boîtes à outils de neuro-imagerie utilisées pour traiter les ensembles de données évoquées d'IRMf, cette thèse soulève les multiples problèmes liés aux approches actuelles de traitement et les défis généraux que le domaine tente de relever pour innover en matière de technologie et de méthodologie.

Les quatre *problèmes* identifiés dans cette thèse sont : (1) l'inaccessibilité et l'adhésivité des boîtes à outils de neuro-imagerie, (2) l'absence de normes générales pour les analyses de neuro-imagerie, (3) une crise de reproductibilité et de transparence et (4) l'insuffisance du contrôle de la qualité des données et de la communication des résultats. Les trois *défis* identifiés dans cette thèse sont : l'innovation (1) dans la dimension spatiale, (2) dans la dimension temporelle et (3) dans le traitement du signal.

La première étude (étude A) aborde la plupart de ces questions et défis en introduisant une nouvelle boîte à outils de neuro-imagerie, appelée *fMRIflows*. Cette boîte à outils est un consortium de pipelines de traitement entièrement automatiques, capables d'effectuer un pré-traitement des données de pointe, ainsi qu'une analyse univariée et multivariée de premier et de second niveau. La validation de la boîte à outils est effectuée en analysant trois ensembles de données d'IRMf différents avec une résolution temporelle différente (2000 ms, 1000 ms et 600 ms) puis en comparant les résultats créés avec *fMRIflows* aux résultats créés avec les logiciels de neuro-imagerie de pointe *fMRIPrep*, *FSL* et *SPM*. La validation montre qu'aucune différence importante n'est observée entre les résultats des quatre boîtes à outils. En outre, l'étude montre qu'un filtrage temporel adéquat d'un ensemble de données d'IRMf avec une résolution temporelle inférieure à une seconde peut conduire à une amélioration du rapport signal/ bruit temporel (TSNR) après le pré-traitement et à une sensibilité statistique accrue dans l'analyse de premier et de deuxième niveau.

La deuxième étude (étude B) aborde un grand nombre des questions susmentionnées, en se concentrant particulièrement sur la quatrième (c'est-à-dire la communication des résultats), en introduisant une nouvelle boîte à outils de neuro-imagerie, appelée *AtlasReader*. Cette boîte à outils peut être utilisée pour générer des tableaux de coordonnées, des étiquettes de régions et des figures informatives à partir d'images IRM statistiques. L'étude B introduit avec succès une boîte à outils de neuro-imagerie qui permet de créer des rapports de résultats beaux et informatifs, indépendamment du système d'exploitation de l'utilisateur. De plus, *AtlasReader* permet l'extraction de tables d'association à partir de plusieurs atlas qui ne sont généralement pas accessibles à un seul système d'exploitation.

La troisième étude (étude C), utilise ces boîtes à outils développées dans les études A et B et montre leur application dans une étude de neurosciences cognitives dans le domaine de l'intégration multi-sensorielle. L'étude a examiné les mécanismes cérébraux impliqués dans l'encodage et la récupération ultérieure d'objets multi-sensoriels sémantiquement congruents. Dans cette étude, nous avons découvert que le cortex visuel de bas niveau peut décoder de manière fiable si un stimulus visuel entrant a été perçu précédemment dans un contexte sémantiquement congruent ou incongru, même si le stimulus visuel n'a été perçu qu'une seule fois auparavant. Les résultats de cette étude soutiennent l'idée que le cortex visuel de bas niveau possède une architecture multi-sensorielle et que la création de souvenirs bénéficie d'une exposition multi-sensorielle à des stimuli sémantiquement congruents.

Enfin, la thèse évalue les deux premières études au niveau des problèmes et des défis mentionnés précédemment et la troisième pour la nouveauté et les nouvelles connaissances acquises par cette étude cognitive. Cette évaluation est suivie d'une évaluation critique des limites de l'étude. Ensuite, les directions futures des routines d'analyse du traitement IRMf et de l'investigation de l'intégration multi-sensorielle basée sur l'IRMf sont discutées. La thèse se termine par des perspectives attendues et par une conclusion générale.

Ensemble, les études comprises dans cette thèse soulignent et abordent les questions et les défis actuellement présents dans le domaine de la neuro-imagerie et fournissent une voie à suivre. De plus, les résultats et les conclusions de ces trois études contribuent à une meilleure compréhension de la manière de pré-traiter et d'analyser correctement les ensembles de données IRMf, ainsi que de la manière dont le mécanisme d'intégration multi-sensorielle se produit.

List of Abbreviations

2D-EPI: 2-dimensional echo-planar imaging	fMRIPrep: Robust preprocessing pipeline for fMRI data
AAL: Automated anatomical labelling atlas	GLM: General linear model
aCompCor: Anatomical based CompCor	GM: Gray matter
AFNI: Analysis of Functional NeuroImages	GRAPPA: Generalized autocalibrating partially parallel acquisitions
ANOVA: Analysis of Variance	GUI: Graphical user interface
ANTs: Advanced Normalization Tools	HRF: Hemodynamic response function
BBR: Boundary-Based Registration	Hz: Hertz
BIDS: Brain Imaging Data Structure	ICA: Independent component analysis
BOLD: Blood oxygenation level dependent	ICBM: International Consortium for Brain Mapping
BRAMS: Brain, Music and Sound Research	IFC: Inferior frontal cortex
BSD-3-Clause License: Revised Berkeley Software Distribution License	INU: Intensity non-uniformity
CHUV: Centre hospitalier universitaire vaudois	ITI: Inter-trial interval
CIBM: Center for Biomedical Imaging	JSON: JavaScript Object Notation
COBIDAS: Committee on Best Practices in Data Analysis and Sharing	LIBSVM: Library for linear support vector machine optimization
CSF: Cerebrospinal fluid	LINE: Laboratory for Investigative Neurophysiology
CSV: Comma-separated values	LOC: Lateral occipital cortex
CompCor: Component Based Noise Correction Method	mm: Millimeter
DVARs: Root mean square variance over voxels after temporal differencing	ms: Millisecond
EEG: Electroencephalography	MNI: Montreal Neurological Institute
EPFL: Ecole polytechnique fédérale de Lausanne	MPRAGE: Magnetization prepared rapid gradient echo
FD: Framewise displacement	MRI: Magnetic resonance imaging
FDR: False discovery rate	MRIQC: Quality control toolbox for MRI images
FOV: Field of view	MVPA: Multivariate pattern analysis
FPR: False positive rate	NIFTI: Neuroimaging Informatics Technology Initiative
FSL: FMRIB Software Library	OHBM: Organization for Human Brain Mapping
FWHM: Full width at half maximum	PCA: Principal component analysis
fMRI: Functional magnetic resonance imaging	
fMRIflows: Consortium of fMRI processing pipelines to perform univariate and multivariate analyses	

PyMVPA: Multivariate pattern analysis toolbox in Python	tCompCor: Temporal based CompCor
RAS: Neurological convention for image orientation: right, anterior and superior	TE: Echo Time
ROI: Region of Interest	TPM: Tissue probabilistic maps
SD: Standard deviation	TR: Repetition Time
SMLR: Sparse multinomial logistic regression	TR1000: fMRI dataset with a TR of 1000ms
SMS: Simultaneous multi slice	TR2000: fMRI dataset with a TR of 2000ms
SNR: Signal-to-noise ratio	TR600: fMRI dataset with a TR of 600ms
SPM: Statistical parametric maps	tSNR: Temporal signal-to-noise ratio
SPM12: Statistical Parametric Mapping toolbox – version 12	TSV: Tabular-separated values
STC: Superior temporal cortex	TV: Total brain volume
SVM: Support vector machine	UNIL: University of Lausanne
T1w: T1-weighted MRI image	WM: White matter
	WMA: World Medical Association

Table of contents

1	<u>GENERAL INTRODUCTION</u>	1
1.1	GENERAL APPROACH TO EVENT-RELATED FMRI DATA ANALYSIS	1
1.1.1	WHAT IS EVENT-RELATED FMRI AND WHAT DOES IT MEASURE?	1
1.1.2	PREPROCESSING AND ANALYSIS OF FMRI DATA	3
1.1.3	STANDARD NEUROIMAGING TOOLBOXES USED TO PROCESS FMRI DATA	4
1.2	ISSUES WITH THE CURRENT STANDARDS IN FMRI DATA ANALYSIS	5
1.2.1	INACCESSIBILITY AND STICKINESS OF NEUROIMAGING TOOLBOXES	5
1.2.2	MISSING GENERAL STANDARDS FOR NEUROIMAGING ANALYSES	6
1.2.3	AMIDST A REPRODUCIBILITY AND TRANSPARENCY CRISIS	6
1.2.4	INSUFFICIENT DATA QUALITY CONTROL AND RESULT REPORTING	7
1.3	CHALLENGES DUE TO INNOVATION IN TECHNOLOGY AND METHODOLOGY	8
1.3.1	INNOVATION IN THE SPATIAL DIMENSION	8
1.3.2	INNOVATION IN THE TEMPORAL DIMENSION	9
1.3.3	INNOVATION IN SIGNAL PROCESSING	10
1.4	AIM OF THE THESIS	12
2	<u>STUDY A – FMRI FLOWS: A CONSORTIUM OF FULLY AUTOMATIC UNIVARIATE AND MULTIVARIATE FMRI PROCESSING PIPELINES</u>	13
3	<u>STUDY B – ATLASREADER: A PYTHON PACKAGE TO GENERATE COORDINATE TABLES, REGION LABELS, AND INFORMATIVE FIGURES FROM STATISTICAL MRI IMAGES</u>	15
4	<u>STUDY C – DECODING OF MULTISENSORY SEMANTICS AND MEMORIES IN LOW-LEVEL VISUAL CORTEX</u>	16
5	<u>GENERAL DISCUSSION AND CONCLUSION</u>	17
5.1	ASSESSMENT OF STUDY A	18
5.2	ASSESSMENT OF STUDY B	21
5.3	ASSESSMENT OF STUDY C	23
5.4	LIMITATIONS OF STUDY A, B AND C	23
5.5	FUTURE DIRECTION IN FMRI PROCESSING AND ANALYSIS ROUTINES	25
5.6	FUTURE DIRECTION OF FMRI-BASED INVESTIGATION OF MULTISENSORY INTEGRATIONS	27
5.7	PROSPECT	30
5.8	CONCLUSION	31
6	<u>REFERENCES</u>	32
	<u>ANNEX 1</u>	I
	<u>ANNEX 2</u>	LVII
	<u>ANNEX 3</u>	LXIII

List of figures and tables

Figures

Figure 1: Depiction of hemodynamic response function of the fMRI BOLD response to a period of stimulation.....	2
Figure 2: Depiction of average brain signal and two motion parameters with its corresponding power spectrum from a dataset recorded with a TR of 600ms.....	10
Figure 3: Depiction of average brain signal and two motion parameters with its corresponding power spectrum from a dataset recorded with a TR of 600ms, once with and once without the application of a low-pass filter at 0.2Hz	19
Figure 4: Visualization of a single subject’s statistical map showing the activation difference between audio-visual and auditory only activation, for five different preprocessing pipelines.....	20
Figure 5: Overview figure of a synthetic statistical brain map generated by AtlasReader	22
Figure 6: Number of yearly publications using a specific keyword as listed on PubMed	26
Figure 7: Number of yearly new publications on OpenNeuro.org and number of fMRI images analyzed with MRIQC which were recorded with an multiband acceleration routine or had a TR below 1000ms	27
Figure 8: Depiction of time-resolved HRF signal registered in three distinct brain regions, from a multisensory integration exposure paradigm, with seven distinct stimuli types.....	29

Tables

Table 1: Example output of AtlasReader’s peak table, generated by analyzing the synthetic statistical brain map shown in Figure 5.....	22
Table 2: Example output of AtlasReader’s cluster table, generated by analyzing the synthetic statistical brain map shown in Figure 5.....	22

1 General Introduction

1.1 General approach to event-related fMRI data analysis

1.1.1 What is event-related fMRI and what does it measure?

Functional magnetic resonance imaging (fMRI) is a non-invasive technique for measuring and mapping brain activity using powerful machines, called MRI scanners. Magnetic resonance imaging (MRI) has become a standard tool for neuroscientists to measure diverse properties of the brain, due to its high spatial resolution and good contrast between different tissue types. Depending on the scan sequence used during recording (also known as imaging), an MRI machine is capable of creating a 3-dimensional image of the brain that is sensitive to a particular tissue type. It does so by looking at the proportion of hydrogen nuclei at a given location and measuring their unique energy release profile after being hit with a specific radio frequency pulse. Usually, when the imaging focus is put on the white matter tissue in the brain, we call it a T1-weighted image, and when the focus is put on the gray matter tissue, we call it a T2-weighted image. Because of their particular properties, T1-weighted images are often used to investigate the anatomy of the brain, while T2-weighted images are mostly used to investigate the functional activity of the brain.

Properties of MRI images

Each individual MRI image, also called volume, contains a full picture of the brain. Such a 3-dimensional picture provides an intensity value for each spatial location, also called a voxel. Each voxel has a spatial resolution. If the spatial resolution is identical in the three spatial directions x , y and z , then we talk of an isotropic voxel, otherwise we talk from an anisotropic voxel. Until recently, MRI volumes were usually recorded in a slice-wise manner. Meaning, to record the whole brain volume, a quick succession of 2-dimensional image slices oriented in x and y directions were recorded for each specific z position. The dimension of a slice in x and y directions is specified in a parameter called the field of view (FOV) and the particular extent of the slice in z direction is specified as the slice thickness. Each recording of a 2-dimensional slice takes time, which means that the total number of slices that needs to be recorded directly stands in relation to the total amount of time that is needed to record a full brain volume. This duration to image a single volume is called the repetition time (TR).

What does fMRI measure?

By recording multiple T2-weighted images in a row, we can create a signal time-series for each individual voxel in the volume. This time-series information can then be used to infer the localized functional activity in the brain. However, it is important to stress that this measurement is *indirect* and *relative*. It is *indirect*, because what we are essentially measuring is the ratio between oxygenated blood (containing oxyhemoglobin, which is diamagnetic) and deoxygenated blood (containing deoxygenated hemoglobin, which is paramagnetic) at a particular location, not the activity of the underlying neurons themselves. It is *relative*, because we are not measuring the actual amount of oxygen in the blood, only the

changes in the ratio between oxygenated and deoxygenated blood. This measured signal is also called the blood-oxygen-level dependent (BOLD) signal (Ogawa, Lee, Kay, & Tank, 1990).

Event-related fMRI

When we measure the BOLD signal without any particular external stimulation, we call this a resting state fMRI recording. When we however stimulate the brain with external or internal cues, we call this an event-related fMRI recording. In both cases, whenever a brain region increases its activity, the neurons consume more oxygen than usual, which decreases the BOLD signal at this particular location. To comply with the higher energy demand at this location, our body rushes more oxygenated blood to this location, which increases the BOLD signal and continues to do so as long as the oxygen requirement is present. Once the brain region reduces its activity back to a normal level the BOLD signal decreases as well, often followed by a post-activation undershoot until it goes back to a normal baseline level (Havlicek, Ivanov, Poser, & Uludag, 2017; Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000). This particular evolution of the BOLD signal during activation is also called the hemodynamic response function (HRF) and can especially be observed during event related fMRI, after a brain region is excited by a specific stimulus (see Figure 1). Depending on the intensity and duration of a stimulus, the HRF will be more pronounced in amplitude or duration. Usually, the peak of the HRF is reached around 4-6 seconds post-stimulus onset and returns to a baseline level after 12-16 seconds, given that there is no prolonged stimulation occurring. However, the HRF shape is profoundly nonlinear with respect to the stimuli duration (K. J. Friston et al., 1998; Glover, 1999; Logothetis & Wandell, 2004). The general shape of the HRF is similar across primary sensory areas, but seems to be variable across higher cortical regions and varies considerably across people (Aguirre, Zarahn, & D'Esposito, 1998; Handwerker, Ollinger, & D'Esposito, 2004; Martuzzi et al., 2006, 2007; Miezin et al., 2000; Neumann, Lohmann, Zysset, & von Cramon, 2003). That is why an appropriate preprocessing of the data is crucial.

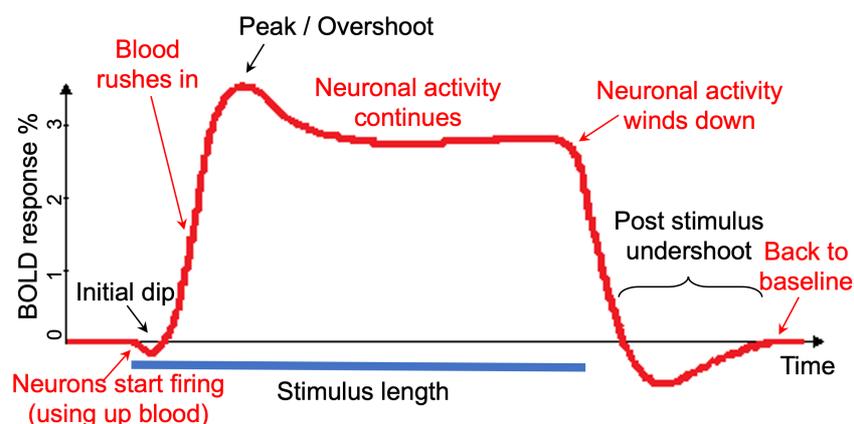


Figure 1: Depiction of hemodynamic response function of the fMRI BOLD response to a period of stimulation. After an initial dip oxygen rich blood is rushed in which creates the characteristic peak of the HRF. With continuous stimulation the HRF will stay at on a certain plateau until the end of stimulation. Without continuous activation, the HRF will directly fall back down after it reached its peak and go back to a baseline level, after going through a post stimulus undershoot. This figure was adapted from Hoge & Pike (2001).

1.1.2 Preprocessing and analysis of fMRI data

To gain any useful insight from fMRI data, appropriate data processing needs to be applied. The steps involved in this regard are commonly separated into a **preprocessing** and an **analysis** stage. Data *preprocessing* is important to make sure that the data are (1) corrected for unwanted effects caused by the measuring technique, (2) corrected for confounding factors due to the biological structure and processes in the brain and (3) to assure that the assumptions required for the analysis are met. Data *analysis* is then used to separate signal from noise and extract meaningful insight from the data by means of statistical inference. Appropriate analysis techniques are crucial to counter the usually very small percentage of signal change observed in the BOLD signal during event-related fMRI.

Data preprocessing

The precise steps involved in an fMRI preprocessing pipeline depend on many different factors, including: (1) scan sequence and acquisition parameters used to record the data, (2) study design, hypothesis and model used for the analysis, (3) neuroimaging toolbox and computer operating system used to conduct the analysis, and (4) experience and background of the person performing the data analysis. Because of all of these factors, the exact preprocessing pipeline of an fMRI study may vary, but in general, all of them include a sequence of the following steps (sequence order might vary): (1) slice-time correction, (2) motion correction, (3) co-registration, (4) spatial normalization, and (5) spatial smoothing. *Slice-time correction* is performed to compensate for the unique time delays between the slices within a volume due to the slice-by-slice wise recording by temporally interpolating slices with adjacent time points. After this correction, the information captured by the image within a volume is considered to represent the same time point. *Motion correction* is performed to correct for participants movement within the scanner during data recording. This correction is necessary to fulfill the assumption that each individual voxel in the image represents the signal time course of only one single location. Unwanted motion between the recording of volumes will violate this assumption and induce noise and uncertainty. Motion correction is usually done by estimating six motion parameters (three rotation and three translation parameters) and correcting for them during preprocessing using a rigid body transformation and accounting for their confounding influence during analysis. *Co-registration* is performed to make sure that each functional image is aligned to the underlying structural image. In other words, to make sure that each voxel of each functional image represents the same location as a corresponding voxel in the anatomical image. This step is required to make sure that all images (functional and anatomical) within a subject are aligned and correspond to same reference space, in this context also called the subject space. *Spatial normalization* is performed to account for subject specific brain characteristics, such as size and shape by transforming the subject specific brain from a unique subject space to a common reference space. This step makes sure that any given voxel in the transformed subject image maps for all individual subjects to the same spatial location in the common reference space. The transformation into this common reference space is usually an affine transformation to a particular template brain, usually one that is close to the MNI152 or the Talairach brain. *Spatial*

smoothing is performed to increase the signal-to-noise ratio (SNR) in an image by applying a spatial convolution to the image. The smoothing has the effect of filtering out the higher spatial frequencies and therefore makes the lower spatial frequencies more apparent. The beneficial effect of smoothing depends on the appropriate amount size of the smoothing kernel, defined by the full width at half maximum (FWHM) parameter.

Data analysis

Once the data are appropriately preprocessed, they are ready for analysis. In event-related fMRI studies, the data analysis is commonly separated into a **1st** and **2nd-level** stage also known as the within- and between-subject or the subject and group analysis. During the *1st-level* analysis, the preprocessed data from one specific subject are analyzed either according to a **univariate** or a **multivariate** approach. A *univariate* analysis investigates the involvement of a particular brain region by looking at the average activation level during an experimental condition. In contrast to this, a *multivariate* analysis investigates the representational content present in a particular brain region by investigating the specific activation pattern present in this region during an experimental condition (Mur, Bandettini, & Kriegeskorte, 2009). A *univariate* approach is usually conducted by using a general linear model (GLM). This is done by creating a model that contains (1) expected neural activation caused by a particular stimulus type during recording, represented by individual stimulus regressors convoluted according to the HRF and (2) confound factors, represented by nuisance regressors or regressors of no-interest in the shape of the known source of noise to account for its unwanted influence. By estimating such a model and computing stimulus-specific contrasts, statistical parametric maps (SPM) can be generated that then can be used for statistical inference. In contrast to the univariate approach, where each voxel is analyzed individually, a *multivariate* approach can analyze multiple voxels at once. By doing so, such a multivariate approach allows the investigation of distributed information patterns. One way to perform multivariate analysis is by using machine learning techniques. Such techniques allow the creation of classifiers capable of distinguishing fine-grained category specific activation pattern at specific brain regions. The outcome of such a multivariate pattern analysis (MVPA) is an accuracy map, indicating the classification accuracy for each location in the brain. From here on out, whenever we mention *multivariate approaches*, we refer to the MVPA approach using machine learning. In the *2nd-level* analysis the subject-specific SPMs or accuracy maps are used to perform statistical inference on the group. Looking at the group instead of one single subject only boosts the overall power and sensitivity of the statistical analysis and allows the generalization of the results to a bigger population by detaching them from the particularity of a specific individual.

1.1.3 Standard neuroimaging toolboxes used to process fMRI data

There exist multiple neuroimaging toolboxes that provide complete processing pipelines to perform preprocessing and analysis of fMRI data, but the most frequently used ones are FSL (Jenkinson,

Beckmann, Behrens, Woolrich, & Smith, 2012), SPM12 (Friston et al., 2006), AFNI (Cox & Hyde, 1997), FreeSurfer (Fischl, 2012) and BrainVoyager (Brain Innovation). While FSL, AFNI and FreeSurfer are freely available toolboxes, BrainVoyager and SPM12 both depend on a monetary license. SPM falls into this category, due to its dependency on MATLAB. BrainVoyager and SPM12 are both available for Windows, macOS and Linux operating systems, whilst FSL, AFNI and FreeSurfer are only available for macOS and Linux operating systems. While all of these toolboxes provide processing pipelines to preprocess and analyze fMRI data, all of them use a slightly different sequence and implementation of the different steps described in the previous section. Additionally, not all of these neuroimaging toolboxes provide all the processing steps required to conduct a state-of-the-art fMRI analysis. For example, not all toolboxes allow natively the possibility to perform a multivariate analysis.

1.2 Issues with the current standards in fMRI data analysis

1.2.1 Inaccessibility and stickiness of neuroimaging toolboxes

Not all neuroimaging toolboxes are available for everyone to use due to the particular license requirement or the specific operating system needed to run them. Additionally, some of them come with an easy to use graphical user interface (GUI), while others require a bit more advanced computer skills. This circumstance, coupled with the software specific expertise researcher gain throughout their career is a likely major reason for why researchers most often stick with their initially learned neuroimaging toolbox, and not because it is necessarily the most suited toolbox for them. This is obviously an issue. By sticking to a particular toolbox, researchers are often only able to profit from the newest developments in the field, when their neuroimaging toolbox is updated to a newer version. Additionally, by sticking to a unique toolbox, researchers become prone to bugs in the code of these toolboxes (Eklund, Nichols, & Knutsson, 2016), potentially overestimate their results by missing results variability between multiple toolboxes (Pauli et al., 2016) or become prone to flawed standards in the field (Vul, Harris, Winkielman, & Pashler, 2009).

To counter the previously mentioned issues, the Nipype toolbox (K. Gorgolewski et al., 2011) was created, of which the thesis author is also part of the core developer team. Nipype serves as an interface to most neuroimaging toolboxes (e.g. SPM12, FSL, FreeSurfer, AFNI) and allows the creation of flexible and complex processing pipelines across all of these neuroimaging toolboxes. Thanks to new technologies such as Docker (<https://www.docker.com>) and Singularity (<https://www.sylabs.io>), Nipype provides access to most of the commonly used neuroimaging toolboxes on any operating system (e.g. Windows, macOS, Linux). Nipype itself is written in the easy to learn programming language Python (<https://www.python.org>) and comes with many useful tutorials. While this is already a good attempt to counter the inaccessibility of some toolboxes and helps researchers to profit from uncharted toolboxes, it does not solve other issues prevalent in the field.

1.2.2 Missing general standards for neuroimaging analyses

Until recently, most of the neuroimaging toolboxes mentioned in the previous section used their own and unique data format to store MRI images. Because of this, exchange of data between toolboxes or researchers with different toolbox proficiencies was difficult. To solve this issue the Neuroimaging Informatics Technology Initiative (NIfTI) was created and introduced a new common standard data type for MRI images, called the NIfTI standard (<https://nifti.nimh.nih.gov/>). While the NIfTI standard was an important step in the right direction, missing standards for datasets and processing pipelines in general were still missing.

Still now, each researcher used their own particular structure to store the dataset and output of an MRI study. Such a setup makes sharing of datasets difficult and the processing and analysis of fMRI data prone to human errors caused by unwanted overwriting of files, chaotic folder structures or unclear file naming conventions. Additionally, unique dataset structures make it difficult to share processing and analysis scripts between researchers. To counter this issue, the so-called Brain Imaging Data Structure (BIDS) (Gorgolewski et al., 2016) was introduced. Using an agreed upon structure to store MRI datasets and a clear naming convention to save raw and processed files is the second important step to a common general standard in neuroimaging analysis.

The third and last important step to a complete general standard in neuroimaging analysis is the introduction of clear pipeline standards. Being able to directly compare different processing and analysis pipelines between each other will help to better understand the unique properties of each pipeline and therefore improve the quality and trustworthiness of the pipeline.

1.2.3 Amidst a reproducibility and transparency crisis

One of the key parts of the scientific method is the topic of reproducibility. Scientific results which cannot be replicated are unsound and their claim to new insights would fall short. It would be impossible to say if the results' occurrence is due to chance or due to the actual scientific work. Many scientific fields are currently struggling through something called a reproducibility crisis as many results fall short of being reproduced (Baker, 2016). The reason for the lack of reproducibility is rarely due to malicious intent, but most often due to missing standards or transparency (Gorgolewski & Poldrack, 2016). In-house closed source software make it almost impossible to understand how the results were achieved. But even in open source software packages, bugs and faulty processes can appear (Eklund et al., 2016). Missing standards as described in the previous section mean that results are described in a subjective manner and are reported with a varying number of toolboxes. Missing transparency of code, data and method make it almost impossible to replicate studies and further hinder the advancement of the field. The best way to tackle all of these deficits is by being completely open (Halchenko & Hanke, 2015).

To counter the issue of subjective description of data analysis and results reporting, the Organization for Human Brain Mapping (OHBM) has created the Committee on Best Practices in Data Analysis and

Sharing (COBIDAS), which has created a clear set of guidelines how studies and results should be reported (Nichols et al., 2017). To counter the lack of transparency, researchers are encouraged to share their code via online services such as GitHub (<https://github.com>) and to provide more detailed information about the computational environment and exact software used during processing of the data. Additionally to that, researchers are encouraged by many journals to publicly share their data through services such as OpenNeuro.org (Gorgolewski, Esteban, Schaefer, Wandell, & Poldrack, 2017) and to share their results, such as the statistical maps of their MRI analysis via online platforms, such as NeuroVault.org (K. J. Gorgolewski et al., 2015).

1.2.4 Insufficient data quality control and result reporting

The quality of data and its processing throughout the study are crucial factors for any scientific study. Unfortunately, quality control is often done subjectively and therefore is prone to human biases. Additionally, looking at the raw data is unfortunately too often neglected or insufficiently done. Connected to the previous points, this is mostly due to a lack of standards and toolboxes available. To counter this, the MRI Quality Control toolbox (MRIQC) was created (Esteban et al., 2017). MRIQC is a fully automated quality control toolbox, providing many important statistics and figures about MRI raw data, i.e. before it was preprocessed. MRIQC is a great example of tackling the other aforementioned issues. It uses the BIDS data format to conform to current standards, its code is open-source and therefore fully transparent, it runs on any system and can quickly adapt to newest advancement in the field thanks to the fact that its processing pipelines are implemented with Nipype. MRIQC is a great example of showing how supportive statistics and figures can help assessing the quality of a dataset. Building on these insights, the authors of MRIQC created a robust preprocessing pipeline for fMRI data, called fMRIPrep (Esteban, Markiewicz, et al., 2019), and added informative figures to the different fMRI data preprocessing steps. Together, both of these toolboxes showed that a state-of-the-art neuroimaging toolbox needs to provide informative figures, which force their users to investigate and acknowledge the quality of their dataset and processing pipelines.

What is still missing is an informative toolbox for the communication of results. This is problematic, as very often, the result reporting in fMRI studies is insufficient to fully communicate the location and extent of significant brain regions. Most commonly, results are only reported through the location of their primary peak and the size of the cluster. Such reporting neglects the actual shape and extent of a cluster and misses the opportunity to report secondary peaks. Additionally, depending on the researchers' background and computational environment used, the reference atlas to identify the region of a peak, the information about secondary peaks and the toolbox to gather all this information may vary and might again be subjectively tinted. Therefore, better standards and automated and transparent reporting toolboxes are necessary. An alternative solution to this issue is to share the results directly as statistical maps via platforms like NeuroVault.org (K. J. Gorgolewski et al., 2015). While such an approach is a great solution

to explore the results yourself and provides a great opportunity to meta-analysis, it does not solve the problem of sufficient reporting within a paper.

1.3 Challenges due to innovation in technology and methodology

Additional to the previous mentioned issues, new innovations in measuring techniques and analysis methods constantly cause new challenges to existing neuroimaging toolboxes and their processing pipelines and force researchers who use them to be more flexible and adaptive in their approaches to analyze fMRI data. The following section highlights important innovations in the way we record or process fMRI data and separate them by the spatial, temporal and signal dimension.

1.3.1 Innovation in the spatial dimension

Thanks to gradual improvement in measuring techniques, fMRI data can now be recorded with a high enough spatial resolution to detect localized distinct activation patterns in images. However, as mentioned above, univariate fMRI analyses are only sensitive to overall activation differences between experimental conditions. Therefore, even if two conditions are very distinct in their fine-grained activation pattern, as long as their spatial-average leads to the same activation, standard univariate analyses will not be able to distinguish them. In contrast to this, multivariate analysis, or more specifically MVPA, are capable of investigating how a specific object is represented in a region and if this representation is fundamentally different from another object representation, even if their average activation is identical (Mur et al., 2009).

Even though the advantage and fruitfulness of multivariate approaches is clear (Haxby, 2012), most neuroimaging toolboxes lack analysis pipelines capable of conducting such analyses. Currently, the main toolboxes to perform state-of-the-art MVPA on neuroimaging data are PyMVPA (Hanke et al., 2009), CoSMoMVPA (Oosterhof, Connolly, & Haxby, 2016) and scikit-learn (Abraham et al., 2014). All of them are written in Python or MATLAB and need some degree of programming skill to use, which is why they are not equally accessible to all researchers in the field. Additionally, multivariate analyses profit from slightly adapted data preprocessing pipelines. Some of them, such as moving the normalization to template step after the MVPA analysis or applying a weaker spatial smoothing to preserve the underlying pattern as much as possible, can easily be implemented in current processing pipelines. However, others, such as spatial band-pass filtering to improve signal decoding accuracy (Sengupta, Pollmann, & Hanke, 2018) can only be implemented with custom code and are not yet possible in standard preprocessing pipelines.

1.3.2 Innovation in the temporal dimension

The temporal resolution in which fMRI images can be recorded is an important, but also restrictive factor in neuroimaging analysis. Currently, fMRI studies are usually recorded with a TR above one second. This is not a big problem when we are looking at signal changes that happen over multiple seconds. In these cases, the comparably slow HRF curve serves us well enough to model the underlying brain activation. However, using things like event onset jitter to better map the full extent of the HRF can only bring us so far. The less sampling points we have to model a curve, the fewer revealing characteristics we can observe about our model. In the case of the HRF, a slow TR might only reveal the overall amplitude of the HRF but will not provide information to better understand the exact shape or onset delay of the curve. For this, we would need to have a much higher TR that allows us to sufficiently sample the HRF. Additionally, having a low temporal resolution will prevent us from observing any fast-changing signal changes, may they be localized general fluctuations or pattern variations.

Fortunately, new advancements in MRI measuring techniques allow the recording of fMRI images with a much higher temporal resolution than previously thought possible. The trick to this was the introduction of new scan sequences which use acceleration techniques such as GRAPPA (Griswold et al., 2002), simultaneous multi-slice/multiband acquisitions (Feinberg et al., 2010; Feinberg & Setsompop, 2013; Moeller et al., 2010) or even one shot 3D volume recordings such as 3D-EPI-CAIPI (Narsude, Gallichan, van der Zwaag, Gruetter, & Marques, 2016). Depending on the sequences used and on the brain volume that needs to be covered, a TR reduction of 2 to 8-fold is possible while keeping comparable data quality as previous methods.

However, the introduction of sub-second temporal resolution creates new challenges for the processing of functional images. By recording brain activity with a high enough temporal resolution, we will inevitably also be able to sample the respiration (0.2-0.6Hz) and cardiac (1-1.7Hz) signal (Viessmann, Möller, & Jezzard, 2018). Typically, things like low-pass filtering are not performed during preprocessing as it is assumed that high frequency signals like cardiac or respiratory cycles would be severely aliased by such filters. However, given their oscillation frequency and considering the Nyquist–Shannon sampling theorem, a sufficient sampling of the average respiration and cardiac signal will become possible if the TR is below 1250ms and 400ms, respectively. Thanks to the aforementioned acceleration techniques, a complete sampling of the respiratory signal becomes possible (see Figure 2). One consequence is that appropriate new filtering techniques need to be introduced during the preprocessing pipeline. While high-pass filtering is a standard procedure in most processing pipelines, an appropriate temporal low-pass or band-pass filter is often not possible. And with these toolboxes where flexible temporal filtering is possible, none so far is capable of providing an appropriate preprocessing pipeline that keeps different preprocessing filters orthogonal to each other, to prevent the reintroduction of previously removed noise (Lindquist, Geuter, Wager, & Caffo, 2019).

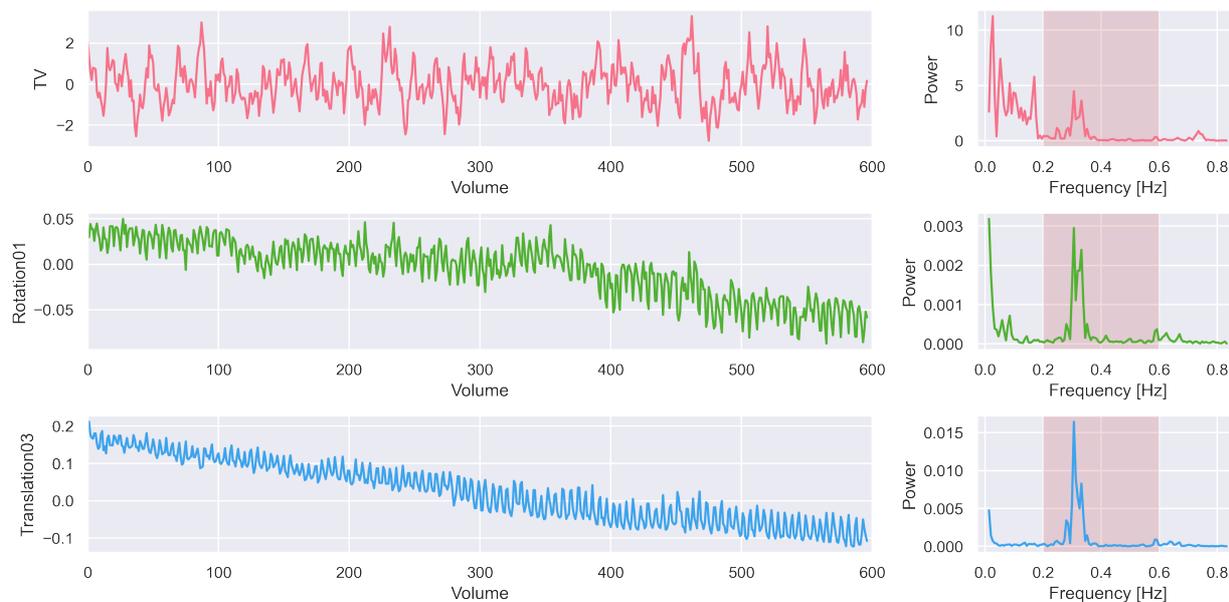


Figure 2: Depiction of average brain signal and two motion parameters with its corresponding power spectrum from a dataset recorded with a TR of 600ms. The three panels show from top to bottom, the average signal in the total brain volume (TV), the estimated rotation along the x-axis (Rotation01) and the estimated translation along the z-axis (Translation03). The panels on the left show the before mentioned data over time, represented in fMRI volume, and the panels on the right show the corresponding power spectrum, with the frequency depicted on the x-axis. The sampling window to record the respiration signal (0.2-0.6Hz) is highlighted in red.

1.3.3 Innovation in signal processing

Deducing brain activation by looking at changes in the BOLD signal is by no means straightforward. Due to the measuring technique itself and because of the properties of the biophysical process we want to observe, proper data preprocessing is crucial. Being an aggregation of multiple biological, physical and thermal factors, the BOLD signal is prone to many confounding factors. Amongst others, these include: (1) scanner field inhomogeneity during recording (Vovk, Pernus, & Likar, 2007), (2) respiratory and cardiovascular fluctuations (Viessmann et al., 2018), (3) in-scanner motion during recording (K. J. Friston, Williams, Howard, Frackowiak, & Turner, 1996; Hajnal et al., 1994), (4) improper data processing (Lindquist et al., 2019), and (5) signal fluctuations that is not due to neural activity (Liu, 2016; Thomas, Harshman, & Menon, 2002; Triantafyllou et al., 2005).

Fortunately, many new methods were developed to improve the data analysis with respect to these factors. While some of them require the recording of additional data, such as the distortion field map for inhomogeneity correction (Hutton et al., 2002) or the external measurement of respiratory and cardiac signals (Glover, Li, & Ress, 2000; Kasper et al., 2017), others work directly on the data at hand and do not require additional data recordings. For the purpose of universal applicable methods, I will focus here only on the latter category of methods. For each of the aforementioned confounding factors, a multitude of techniques exist to tackle them. The following is therefore only a brief listing of solutions: *Scanner field inhomogeneity during recording* can be corrected with the N4ITK algorithms (Tustison et al., 2010),

available in the toolbox ANTs (Avants et al., 2011). *Respiratory and cardiovascular fluctuations* can be handled through appropriate temporal filtering (Viessmann et al., 2018) or component analysis, such as independent component analysis (ICA) or dictionary learning (Mensch, Varoquaux, & Thirion, 2016). Potential toolboxes to do this are Nilearn (Abraham et al., 2014) or the Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) toolbox (Beckmann & Smith, 2004), as implemented in the neuroimaging toolbox FSL. The correction for *in-scanner motion* is implemented as such in almost any fMRI processing pipeline. However, as mentioned in the previous section, current techniques were developed for fMRI images with temporal resolution above 1000ms and might therefore not be adapted to properly handle things such respiratory and cardiac fluctuations, sampled by scan sequences with higher temporal resolution. *Improper data processing* can reintroduce previously removed noise back into the data. This can be avoided by applying orthogonal filtering during preprocessing (Lindquist et al., 2019). However, so far, none of the standard neuroimaging preprocessing pipelines is capable to do this correctly. To remove *signal fluctuations which are not due to neural activity*, i.e. noise, approaches such as CompCor (Behzadi, Restom, Liao, & Liu, 2007) or ICA could be used. While CompCor uses a principal component analysis (PCA) to identify noise components in predefined regions of no interest, ICA performs a blind source separation with the goal to separate noisy components from relevant signal components. In either case, these components can then either directly be removed from the processed signal or be included as nuisance regressors in the GLM during data analysis.

1.4 Aim of the thesis

The thesis at hand aimed at addressing the previously mentioned issues with current standards in fMRI data analysis and the new challenges due to innovation in technology and methodology by introducing two new neuroimaging software packages. An additional cognitive fMRI study was conducted to validate these toolboxes under normal conditions. Therefore, the three studies reported in this thesis, which address these points, can be separated as follows:

The *first study* built on previous advancements in the field of fMRI processing pipelines and introduces a new neuroimaging toolbox called *fMRIflows*, capable of performing a complete fMRI analysis, from data preprocessing to subject and group-level analysis and result reporting. This toolbox is able to do this fully automatically for univariate and multivariate analysis. As such, fMRIflows addresses all of the issues and challenges mentioned before, except for the topic of results reporting. This study is further referred to as *Study A* and led to the submission of “fMRIflows: a consortium of fully automatic univariate and multivariate fMRI processing pipelines” in *NeuroImage*. Please refer to chapter 2 for a summary and to Annex 1 for the submitted article.

The *second study* introduces a new neuroimaging toolbox called *AtlasReader*. This toolbox addresses the remaining issue with result reporting by providing an improved and automated way to create publication ready outputs (figures and tables) for fMRI studies. This study is further referred to as *Study B* and led to the publication of “AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI images” in the *Journal of Open Source Software* (Notter et al., 2019). Please refer to chapter 3 for a summary and to Annex 2 for the published article.

The *third study* investigates if the low-level visual cortex is capable of decoding multisensory semantics and memories through the usage of MVPA. As such, this study also validates the two newly introduced neuroimaging toolboxes from Study A and Study B. This study is further referred to as *Study C* and led to the submission of “Decoding of multisensory semantics and memories in low-level visual cortex” in *NeuroImage*. Please refer to chapter 4 for a summary and to Annex 4 for the submitted article.

2 Study A – fMRIflows: a consortium of fully automatic univariate and multivariate fMRI processing pipelines

Authors: Michael P. Notter^{1*}, Peer Herholz^{2,3*}, Sandra Da Costa⁴, Omer F. Gulban^{5,6}, Ayse I. Isik⁷ and Micah M. Murray^{1, 8-10}

1. The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, University Hospital Center and University of Lausanne, Switzerland
2. International Laboratory for Brain, Music and Sound Research, Université de Montréal & McGill University, Montréal, Canada
3. McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, Canada
4. Animal Imaging and Technology Section, Center for Biomedical Imaging (CIBM), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
5. Department of Cognitive Neuroscience, Maastricht University, The Netherlands
6. Brain Innovation B.V., Maastricht, The Netherlands
7. Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany
8. Sensory, Perceptual and Cognitive Neuroscience Section, Center for Biomedical Imaging (CIBM), Lausanne, Switzerland
9. Department of Ophthalmology, Fondation Asile des aveugles and University of Lausanne, Lausanne, Switzerland
10. Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

(* shared first co-authorship)

Submitted to *NeuroImage*, 2021 Mar 23. To be found in Annex 1.

Contribution: The candidate created the project, wrote and developed the toolbox, performed the data acquisition, conducted the validation analyses, prepared the manuscript for submission in a peer-reviewed journal and handled the reviewing process.

This study examined the strength and weaknesses of current neuroimaging processing pipelines and identified important new technologies and standards in the field. Combining all of these insights, this study introduced a new neuroimaging toolbox called fMRIflows. fMRIflows, is a fully automatic processing pipeline capable of performing state-of-the-art data preprocessing, as well as first- and second level univariate and multivariate analysis. fMRIflows is built upon the advancements of fMRIPrep and MRIQC, enforces the newly introduced BIDS standards for datasets and allows for flexible temporal and spatial filtering, while considering orthogonality between filters. All of these are important new properties and features. The study explains all different processing steps included in the pipeline in much detail and highlights the advantages of a community developed and open-source toolbox. The validation of this software package was done on three distinct fMRI datasets, each with different characteristics. Most notably is the difference in the temporal resolution, as the three datasets have a sampling rate of 2000ms, 1000ms and 600ms. Especially the last dataset was crucial to showcase one of fMRIflows most pronounced feature, the appropriate handling of fMRI datasets recorded with high temporal resolution, i.e. datasets

with a sampling rate below 1000ms. Because of the lack of publicly available fMRI dataset with such high sampling rates, the author of this thesis recorded a complete with a temporal resolution of 600ms, including a total of 17 subjects. To allow the direct comparison of fMRIflows to other processing pipelines, the beforementioned three datasets were all preprocessed with comparable neuroimaging software packages FSL, SPM and fMRIPrep. To highlight the relevance of temporal filtering, the datasets were preprocessed with fMRIflows, once with and once without a low-pass filter at 0.2Hz. The results show that no clear differences can be observed between the preprocessing of fMRIflows and the other toolboxes fMRIPrep, FSL and SPM on all three datasets. However, fMRIflows' flexible temporal filtering approach led to improved temporal signal-to-noise-ratio (TSNR) after preprocessing and increased statistical sensitivity, in particular for datasets at or below a sampling rate of 1000ms.

3 Study B – AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI images

Authors: **Michael P. Notter**¹, Dan Gale², Peer Herholz³⁻⁵, Ross Markello⁵, Marie-Laure Notter-Bielser⁶, and Kirstie Whitaker⁷

1. The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, University Hospital Center and University of Lausanne, Switzerland
2. Centre for Neuroscience Studies, Queen's University, Kingston, Canada
3. Laboratory for Multimodal Neuroimaging, Philipps-University Marburg, Hesse, Germany
4. International Laboratory for Brain, Music and Sound Research, Université de Montréal & McGill University, Montréal, Canada
5. McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, Canada
6. Centre Leenaards de la Mémoire, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland
7. Alan Turing Institute, London, UK; Department of Psychiatry, University of Cambridge, Cambridge, UK

Published in *Journal of Open Source Software*, 2019 Feb 24. To be found in Annex 2.

Contribution: The candidate created the project, wrote and developed the toolbox, conducted the validation analyses, prepared the manuscript for submission in a peer-reviewed journal and handled the reviewing process.

This study details the need of a toolbox that can create publication-ready output for fMRI studies in a flexible way. It provides a solution in the form of a neuroimaging toolbox called AtlasReader. Many comparable software packages are often restricted, either to the atlas they use to lookup locations or the operating system or software framework they use. To counter this thematic, AtlasReader was written in Python but is also available as a command line tool, is freely available, community developed and allows the usage of nine different lookup atlases. To avoid tedious manual interventions, the toolbox tries to perform as much as possible in an automatic way. It is able to do so by taking advantages of many other neuroimaging toolboxes, notably Nilearn. For each provided MRI image, AtlasReader will provide an overview figure containing all relevant datapoints in the volume, a more detailed informative figure for each individual data cluster, an overview table with information about location and atlas affiliation of each peak, as well as an overview table with information about the location and atlas affiliation of each cluster and its extent. In particular the last point is something that so far no other neuroimaging toolbox provides.

4 Study C – Decoding of multisensory semantics and memories in low-level visual cortex

Authors: **Michael P. Notter**¹, Sandra Da Costa², Anna Gaglianese¹, and Micah M. Murray^{1,3-5}

1. The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, University Hospital Center and University of Lausanne, Switzerland
2. Center for Biomedical Imaging (CIBM), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
3. Sensory, Perceptual and Cognitive Neuroscience Section, Center for Biomedical Imaging (CIBM), Lausanne, Switzerland
4. Department of Ophthalmology, Fondation Asile des aveugles and University of Lausanne, Lausanne, Switzerland
5. Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

Submitted to *NeuroImage*, revision in preparation. To be found in Annex 3.

Contribution: The candidate performed the data analysis, and prepared the manuscript for a submission in a peer-reviewed journal.

This study investigated the brain mechanisms involved during the encoding and subsequent retrieval of semantically congruent multisensory objects. Twelve healthy participants partook in the study and were presented with multiple visual stimuli. Each visual stimulus was presented twice, first during an encoding phase and the second time during a so-called decoding phase. The first presentation of a visual stimuli was sometimes paired with a congruent or incongruent auditory stimulus. During the second presentation of the visual stimuli, participants were asked if they saw the object before. During this continuous recognition task fMRI images with a temporal resolution of 2000ms were recorded. The data was preprocessed with fMRIflores and a univariate and multivariate analysis was conducted to gain new insights. The univariate analysis was used to identify regions involved in the information processing of semantic context dependent multisensory memory traces. In contrast to that, the multivariate analysis was used to locate where in the brain the representational content of these traces is encoded and where they are later on again retrieved. The results show that the low-level visual cortex reliably can decode whether an incoming visual stimulus previously had been perceived in a semantically congruent or incongruent context, even if the visual stimuli was only perceived once before. The results from this study further support the notion that the low-level visual cortex has multisensory architecture and that the creation of memories profits from a multisensory semantic congruent stimuli exposure.

5 General Discussion and Conclusion

The thesis at hand introduces new techniques and standards to process fMRI data in the form of the two toolboxes *fMRIflows* and *AtlasReader*. Both software packages were created to tackle many presiding issues and challenges with current neuroimaging practices (see Study A and Study B). To showcase the relevance of such innovations, the thesis also includes a cognitive neuroscience study in the domain of multisensory integration (see Study C).

Before assessing the outcome and limitations of the two toolboxes reported in Study A and B, and the functional neuroimaging Study C, let me briefly summarize again the current issues and challenges in the neuroimaging domain, which we identified in the Introductions:

Issue 1: Neuroimaging toolboxes are inaccessible and package routines are difficult to interchange. The solution should provide flexibility and accessibility.

Issue 2: While standards for neuroimaging data (i.e. NIfTI) and datasets (i.e. BIDS) exist, standards for neuroimaging analysis are still missing. The solution should introduce validated, tested and community developed analysis pipeline standards.

Issue 3: Closed-off research work and analysis practices created a reproducibility and transparency crisis. The solution is sharing of computational environments, data, code and procedures.

Issue 4: Insufficient quality control and result reporting. The solution should provide detailed quality control for processing pipelines and allow for detailed and publication ready results report.

Challenge 1: Innovation in the spatial dimension of fMRI data recording allows for more detailed acquisition of signal patterns. The solution should allow flexible preprocessing of such data and provide suitable analysis methods.

Challenge 2: Innovation in the temporal dimension of fMRI data recording allows for faster acquisition schemes which lead to new signal confounds. The solution should be adapted to this new kind of data and allow flexible data preprocessing.

Challenge 3: Innovation in signal processing of fMRI data allows for better preprocessing and confound estimations. The solution should be capable of using these new techniques and extract these new confounds.

5.1 Assessment of Study A

The neuroimaging toolbox *fMRIflows*, which was developed in *Study A*, introduces a fully automatic processing pipeline capable of performing state-of-the-art data preprocessing, as well as first- and second level univariate and multivariate analysis. To validate the software package *Study A* processed and analyzed three distinct fMRI datasets. The most notable difference between the three datasets was the temporal resolution (TR) with which they were recorded, which was 2000ms, 1000ms and 600ms. The dataset with a TR of 600ms was specifically recorded for Study A and was crucial to showcase one of *fMRIflows* most pronounced features, the appropriate handling of fMRI datasets recorded with sub-second temporal resolution. To compare *fMRIflows* performance to other state-of-the-art neuroimaging processing pipelines the three datasets were as well preprocessed with the comparable neuroimaging pipelines from FSL, SPM and fMRIPrep. The outcome of this validation showed that *fMRIflows* performs comparably or better than these other software packages. However, when an additional temporal filtering, using a low-pass filter at 0.2Hz, was applied - something that is currently only available with *fMRIflows* - temporal signal-to-noise-ratio (TSNR) after preprocessing was significantly improved and the statistical sensitivity was significantly increased, in particular for datasets at or below a sampling rate of 1000ms. The additional temporal filtering did not impact the dataset with a TR of 2000ms.

With respect to the initially mentioned *issues* in the neuroimaging domain, *fMRIflows* addresses all four of them: The toolbox is developed as a Python based open-source software package, publicly available via GitHub, and the relevant environment to execute the toolbox is freely available via DockerHub. Additionally, the source code of *fMRIflows* is provided in accessible and easy to adapt Jupyter Notebooks. Furthermore, thanks to the Docker environment in which *fMRIflows* is distributed, software routines usually only available in UNIX or MATLAB environments can now be profited from, even if a user does not have access to these tools. All of these points address the *first* and *third issue* mentioned during the introduction and make the toolbox *flexible* and *accessible*. Due to the fact that *fMRIflows* requires the input data to be in NIfTI format and the dataset according to the BIDS standard, the toolbox reinforces these standards and therefore addresses issue 2. Additionally, with each new code snippet added to the *fMRIflows* routine, multiple tests are undertaken to guarantee the correct functioning of the toolbox. *Issue 4* is addressed through the many intermediate visual outputs *fMRIflows* generates, which help with the quality control of the preprocessing and analysis pipeline.

With respect to the initially mentioned *challenges* in the neuroimaging domain, *fMRIflows* addresses all three of them. Addressing challenge 1, the toolbox allows spatial low-, high- and band-pass filtering of the fMRI dataset. To our knowledge, no other available neuroimaging processing pipeline allows such customized spatial filtering of the data. The flexible spatial filtering is especially important for multivariate analysis, where such a spatial band-pass filtering approach potentially can improve the signal decoding accuracy (Sengupta et al., 2018). Similarly, *fMRIflows* also addresses challenge 2 with flexible temporal

low-, high- and band-pass filtering. However, to prevent the reintroduction of artifacts into the data (Lindquist et al., 2019), this filtering step is performed orthogonally to other preprocessing steps. The flexible temporal filtering is especially important for fMRI datasets with sub-second temporal resolution, where the sampling of confounds like respiration or cardiac signal can become an issue (Viessmann et al., 2018). To address challenge 3, *fMRIflows* estimates multiple state-of-the-art confounds and makes them available in accessible TSV files. Furthermore, the toolbox can readily integrate this information during the univariate and multivariate analysis of the data.

The **most important innovation** introduced by *fMRIflows* is its capability of adequate temporal filtering of fMRI datasets with a sub-second temporal resolution. To better illustrate the importance of this feature, we will look at the estimated confound signal (see Figure 3) and the first-level statistical map (see Figure 4) of an individual subject, once where the data was preprocessed with and once without a low-pass filter at 0.2 Hz.

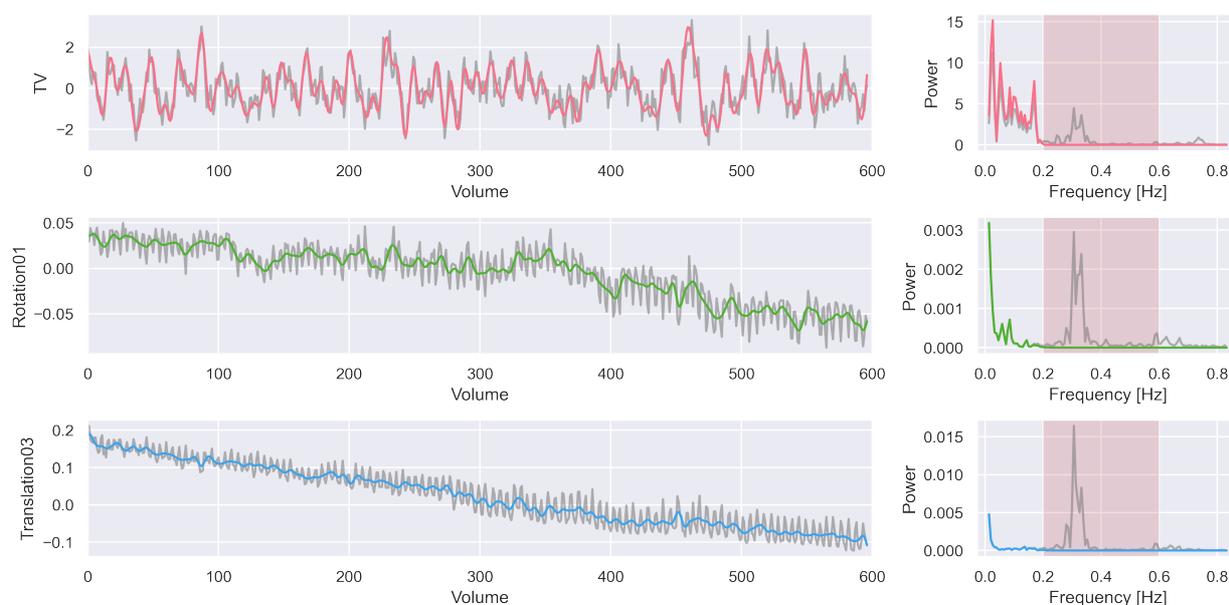


Figure 3: Depiction of average brain signal and two motion parameters with its corresponding power spectrum from a dataset recorded with a TR of 600ms, once with and once without the application of a low-pass filter at 0.2Hz. This figure shows the same information as was provided in Figure 2. However, this time the original depicted signal, i.e. the one preprocessed with a low-pass filter is depicted in light gray, while the signal and power spectrum of the data preprocessed with a low-pass filter at 0.2 Hz is depicted in color.

When looking at the average signal in the total brain volume (TV) and the estimated motion parameters of a dataset with a TR of 600ms, that was not preprocessed with a temporal low-pass filter (as depicted in gray in Figure 3), a clear high frequency oscillation and a potential confound can be observed. This signal oscillation is assumed to be due to the breathing of the participant during data acquisition, as it fits perfectly in the frequency band between 0.2-0.6 (Viessmann et al., 2018). Using the unfiltered brain

signal to estimate brain activation might already be problematic in itself. However, using the shown motion parameters to correct for the in-scanner motion during preprocessing has a high chance to do more harm than not, given that the absolute volume-to-volume motion in certain cases is almost as high as the general motion drift over 100 volumes, which corresponds to one minute. Therefore, an appropriate temporal low-pass filtering is crucial to correctly preprocess such data. This point is further supported by the visualization shown in Figure 4.

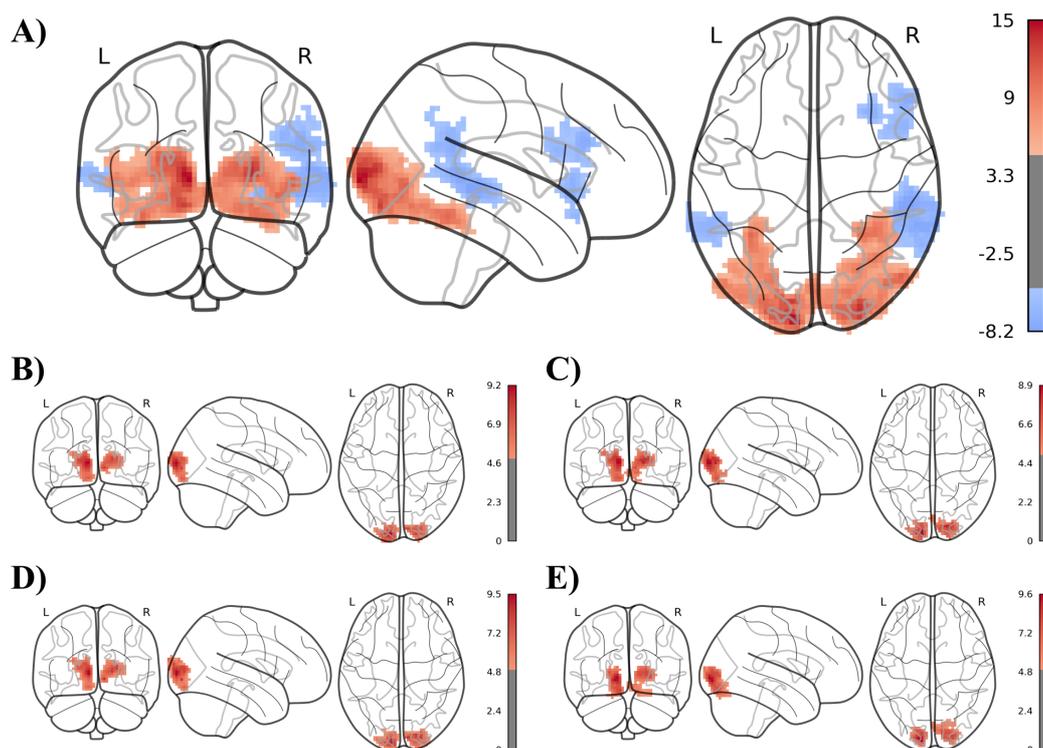


Figure 4: Visualization of a single subject's statistical map showing the activation difference between audio-visual and auditory only activation, for five different preprocessing pipelines. The 'glass-brain' visualizations show the statistical maps (z-scores) for the same individual as shown in Figure 3, for five different preprocessing approaches: A) fMRIflows with a temporal low-pass filter at 0.2Hz; B) fMRIflows without a temporal low-pass filter; C) fMRIPrep, D) FSL and E) SPM. Each image was identically thresholded at the same arbitrary z-value of 5. Furthermore, only clusters of a size bigger than 1350 mm^3 were kept, which corresponds to a cluster size of at least 50 voxels.

Figure 4 shows the statistical maps of a single subject's brain activation difference when perceiving an audio-visual versus an auditory only stimulus. The data used for this visualization is from the third dataset in *Study A*, was recorded with a TR of 600ms, and depicts the same participant as shown in Figure 3. Each of the five panels show a different preprocessing approach and was thresholded at the same height value, and cluster extent. The outcomes of the preprocessing with fMRIflows (without a low-pass filter), fMRIPrep, FSL and SPM, as shown in panel B to E, all look comparable and led to similar maximum z-values between 8.9 and 9.6. This supports the assessment that fMRIflows' preprocessing is comparable to other neuroimaging pipelines in the field. However, the clear advantage of using fMRIflows to preprocess a dataset recorded with a sub-second temporal resolution becomes visible when looking at panel A. This

panel shows the outcome when the preprocessing was done with fMRIflows using a temporal low-pass filter at 0.2 Hz and shows much bigger activation clusters and z-score values, even already at the individual level. This supports the assessment that an adequate preprocessing can lead to an improvement of temporal signal-to-noise ratio and an overall increased statistical sensitivity.

5.2 Assessment of Study B

The neuroimaging toolbox *AtlasReader*, which was developed in *Study B*, introduces an easy-to-use and new way to quickly generate informative and publication-ready outputs. For each provided MRI image, the toolbox will create (1) an overview figure in the shape of a glass-brain plot, highlighting at once all relevant clusters in the volume, (2) more detailed informative figures showing the exact peak of each individual cluster and its respective extent in the sagittal, transversal, and coronal planes, and (3) an overview table with information about the location and atlas affiliation of each cluster peak and its extent.

With respect to the initially mentioned issues in the neuroimaging domain, *AtlasReader* addresses all four of them. Addressing issue 1, the toolbox can be considered *flexible* and *accessible* for multiple reasons. First, *AtlasReader* exploits up to nine different atlases for the lookup of location properties. This consortium of different atlases can rarely be encountered on the same operating system at once, because these atlases depend on different neuroimaging toolboxes such as SPM, FSL or FreeSurfer. Some of these toolboxes are tied either to pricey MATLAB licenses or UNIX capable operating systems. Second, the toolbox and its source code are freely available via GitHub and can be run on any major operating system. It provides a straightforward command line interface or can be integrated more smoothly into any processing script using the easy to learn programming language Python. As an example, embedding this toolbox in a Nipype framework will streamline and optimize the creation of informative output figures and tables. Addressing issue 2, *AtlasReader* supports the standardization initiative in the neuroimaging domain with a well-tested and community developed open-source toolbox. Its numerous available execution parameters allow enough space for customizability in outcome presentation, while its code base keeps the main capability restricted to its standards in results reporting. These factors are simultaneously also the ways how the toolbox addresses issue 3. But most importantly, *AtlasReader* is tackling issue 4 by providing detailed and publication ready results tables and figures. Additionally, these tables and figures can of course also be used as quality control mechanisms of the processing pipelines, ensuring that created output are leading to expected outputs.

The **most important innovation** introduced by *AtlasReader* is its capability of providing atlas affiliation information with respect to the extent of a cluster, not just its peaks. To the knowledge of the author, no other freely available toolbox currently provides this information. To better illustrate the importance of this feature, a synthetic statistical brain map was generated, containing two elongated clusters (see Figure 5).

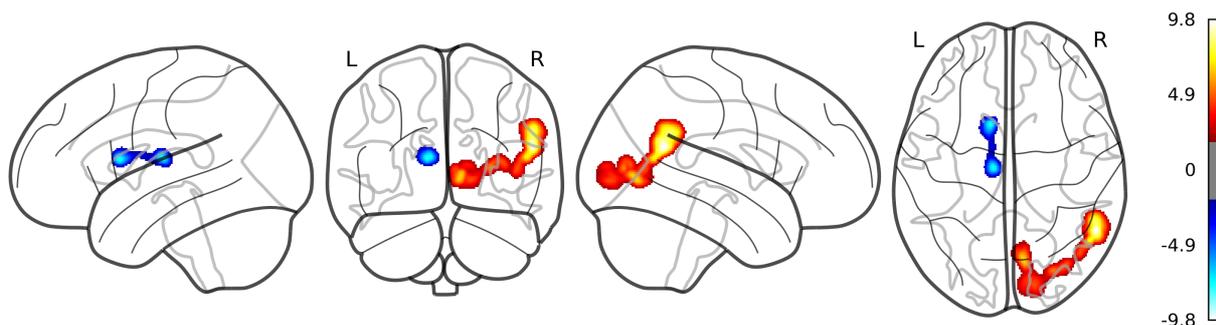


Figure 5: Overview figure of a synthetic statistical brain map generated by AtlasReader. The ‘glassbrain’ visualization generated by AtlasReader, shows that the synthetic statistical brain map contains two activation clusters, both being unusual elongated. The positive activation, shown in red to yellow, covers a region from the right angular gyrus to calcarine. The negative activation, shown in blue to turquoise, covers a region from the left caudate to thalamus

Table 1: Example output of AtlasReader’s peak table, generated by analyzing the synthetic statistical brain map shown in Figure 5.

Cluster ID	Peak Location [x, y, z]	Peak Values [t-value]	Cluster size [mm ³]	Atlas AAL	Atlas Harvard Oxford
1	54, -54, 20	9.80	12163	Temporal Mid R	65.0% Right Angular Gyrus
1	9, -71, 0	6.65	12163	Lingual R	54.0% Right Lingual Gyrus
1	45, -67, 7	4.74	12163	Temporal Mid R	50.0% Right Lateral Occipital Cortex inferior division
2	-11, 12, 13	-7.62	2349	Caudate L	88.0% Left Caudate; 11.0% Left Lateral Ventricular
2	-9, -15, 13	-7.41	2349	Thalamus L	100.0% Left Thalamus

Of those analysis tools which are able to provide atlas affiliation information based on statistical brain maps, many can only provide the information for the main peak in the cluster (as highlighted in green in Table 1). While others are able to provide this information for multiple peaks within a cluster (see other peaks in Table 1), none of them is able to provide this information for multiple atlases at once (see Table 1). Furthermore, *AtlasReader* additionally provides the atlas affiliation for all voxels in a cluster, and is therefore able to more accurately report the extent of an activation cluster (see Table 2). This information can provide relevant insights about otherwise overlooked parts of the cluster.

Table 2: Example output of AtlasReader’s cluster table, generated by analyzing the synthetic statistical brain map shown in Figure 5.

Cluster ID	Peak Values [t-value]	Cluster size [mm ³]	Atlas AAL	Atlas Harvard Oxford
1	3.79	12163	21.00% Angular R 20.06% Calcarine R 17.73% Temporal Mid R 15.65% Lingual R	33.90% Right Angular Gyrus 16.32% Right Lateral Occipital Cortex inferior division 14.50% Right Lingual Gyrus 11.23% Right Occipital Pole 11.21% Right Intracalcarine Cortex
2	-3.69	2349	49.26% Caudate L 43.00% Thalamus L	46.79% Left Thalamus 42.06% Left Caudate 11.15% Left Lateral Ventricular

5.3 Assessment of Study C

The cognitive neuroscience study in *Study C* uses the two toolboxes *fMRIflows* and *AtlasReader*, developed in *Study A* and *Study B*, respectively, and showed how they can be applied to an actual neuroimaging study to gain valuable new insights. The study's focus was in the domain of multisensory integration and addressed the question: What are the mechanisms that allow encoding and retrieval of multisensory information. Using the multivariate approach MVPA, this study investigated if the low-level visual cortex is capable of decoding multisensory semantics and memories of audio-visual stimuli.

Participants in this study performed a continuous recognition task, during which they were presented with either visual stimuli or audiovisual stimuli that were either semantically congruent or incongruent. During the second presentation of a stimulus, participants were only presented with the visual information of the previously audio-visual, or visual-only stimuli. This allowed us to study the retrieval of information that was multisensory during encoding without the interference of presenting multisensory information during retrieval.

The univariate analysis revealed that during the second presentation of a previously audio-visual congruent stimuli, the activation in the lateral occipital cortex (LOC) was higher compared to a previously visual only or audio-visual incongruent stimuli. For the multivariate analysis, an MVPA approach was able to analyze the activation pattern present during encoding of the multivariate stimuli and identify its re-occurrence during the memory retrieval present during the reactivation during the second visual-only presentation of the stimuli. This multivariate analysis revealed several clusters where semantic congruence can be decoded during initial and repeated presentations, including a cluster in the occipital pole. These findings suggest the involvement of low-level visual regions in multisensory and more abstract processes, and further support a multisensory role for supposedly unisensory cortices and that the creation of memories profits from a multisensory semantic congruent stimulus exposure.

With respect to the initially mentioned issues in the neuroimaging domain. Using *fMRIflows* developed in *Study A*, *Study C* showed that it is possible to acquire reliable and informative univariate and multivariate results using accessible, transparent and automatic processing pipelines, which follow a clear state-of-the-art standard. The usage of *AtlasReader* developed in *Study B* allows the creation of output figures, as well as more detailed information tables than otherwise possible.

5.4 Limitations of Study A, B and C

The studies comprised in this thesis have multiple strengths, but their potential to acquire new insights are also tempered by their limitations. In this section, we will critically assess the limitations of each study and provide potential solutions.

Study A introduced the new neuroimaging toolbox *fMRIflows*, capable of performing data preprocessing and analysis, with a univariate and multivariate approach. While this can be a strength of the toolbox, the vast extent of neuroimaging analysis routines that this toolbox tries to cover might make it prone to bugs in the code. For example, software dependency issue in the future will require frequent adaptation to new insights in the field. The solution to this limitation is a good test suite for the software, a transparent development via GitHub as well as a helpful forum where users can address problems and find solutions. The hope of the author is that such a forum is found in the shape of Neurostars.org. However, for this solution to work, enough users need to use the toolbox and participate in its development and support of other users, as the support of the toolbox's main authors can only cover a certain amount of the workload. Another limitation of Study A is due to a general time constraint in software exploration. While the study introduces flexible temporal and spatial filtering routines for univariate and multivariate analyses, the paper only reports the effect of the flexible temporal filtering, but does not further explore the origin of the improved outcome. Additionally, the effect of the spatial filtering for the multivariate analysis, as well as the multivariate analysis in general was not explored in Study A. While these two topics did not have to be revisited in the study, a replication of previous results could have provided relevant insights.

Study B introduced the new results reporting toolbox *AtlasReader*, capable of providing useful figures and reports, as well as new information in the form of cluster tables. Given that this toolbox focuses only on a very small part of the neuroimaging domain and that the code base is rather compact, it does not suffer from the same limitations as mentioned in the previous paragraph. Nonetheless, the toolbox could be further extended to be even more inclusive by allowing other fMRI data formats than NIfTI and data in other spaces than the MNI-space. For example, a transformation of atlas association from MNI to Talairach space could potentially broaden the user base. Furthermore, the toolbox could be improved by providing a web-based interface or API based accessibility to allow the usage of AtlasReader without the need of installing it first. Another limitation of Study B is its lack of exploration of the new 'cluster table' feature. While the main author has already seen in multiple studies of his own that the usage of cluster tables provides additional and often more important information about which brain regions are associated with a certain statistical map, the usefulness and potential of these cluster tables should be explored further. One potential path to better understand the implication of these cluster tables could be the analysis of statistical maps from other studies, freely available via Neurovault.org, and see in which ones reported significant results contain clusters with unusual cluster associations. For example, clusters with more than 50% association to ventricles, white matter or other brain regions usually not associated with brain activation. The source of such clusters with high ratio of non gray matter activation can at this point only be speculated. It could be due to unfavorable spatial smoothing, missing application of appropriate brain masks, incorrect image normalization which could lead to heterogenous outcomes in the mapping into template space, or many other problematic data properties or erroneous pre-processing or analysis steps which could bias the final outcome. Another potentially limiting factor to consider is the number of atlases

which can be used to extract location information. Providing information tables for multiple atlases might allow users to pick and choose the ones they prefer, leading to subjective biases and counteracting the intended standardization purpose of this toolbox, with the goal of providing objective results reporting. However, even though AtlasReader currently provides information from ten different atlases, some of them only provide information for cortical or subcortical regions. Similarly, some atlases were created by analyzing histological architectonic, while others are based on functional connectivity maps.

Study C investigated the brain mechanisms involved during the encoding and subsequent retrieval of semantically congruent multisensory objects using univariate and multivariate approaches. The limitations of this study are due to the paradigm employed in its design and the characteristics of the dataset that was recorded. While the paradigm itself was not a limitation, increasing the complexity of the stimulus categories would have allowed further investigations than what was already previously done in comparable EEG and fMRI studies. Due to the fact that the dataset used in this study was acquired in 2006, the scan sequence to record the data, as well as the resulting data characteristics were sub-optimal for a multivariate approach. Additionally, the number of trial repetitions and participants was (with respect to today's standards) rather low and further complicated the multivariate analysis. Future attempts should acquire the data with a more recent scan sequence, collect more stimuli samples per subject and more subjects overall, and should record the signal with a higher temporal resolution to allow for a time-resolved analysis of the BOLD signal. This last point is further explored in the section "5.6 Future direction of fMRI-based investigation of multisensory integrations", given that the third dataset in Study A addresses many of these issues.

5.5 Future direction in fMRI processing and analysis routines

Every year new studies come out establishing new standards to conduct fMRI studies (Nichols et al., 2017), to store fMRI datasets (Gorgolewski et al., 2016), introducing new quality control tools (Esteban et al., 2017) and automated fMRI preprocessing pipelines (Esteban, Markiewicz, et al., 2019), highlighting the importance of correctly applying a spatial (Sengupta et al., 2018) or temporal filter (Lindquist et al., 2019), or pointing to problematic analysis routines (Eklund et al., 2016) or variability in the way different research teams analyze the exact same fMRI dataset (Botvinik-Nezer et al., 2020). In short, the technology and methodology around neuroimaging processing and analysis is ever evolving.

This is great news for those who are capable to adapt to these changes and who can adapt their processing and analysis pipelines to these new insights. Unfortunately, many of the standard neuroimaging toolboxes take a long time to adapt to these changes and to release a new version. For example, SPM8 and SPM12 were released in 2009 and 2014, respectively. FSL version 5.0 and 6.0 were released in 2012 and 2019, respectively. FreeSurfer's versions 5, 6 and 7 were released in 2010, 2017 and 2020, respectively.

Even though the latest releases of some of these toolboxes are only a few years old, many researchers still perform their analysis on previous versions and are consequently using outdated processing and analysis methods. However, as highlighted in the introduction of this thesis and in Study A, the introduction of Nipype, BIDS, MRIQC, fMRIPrep, Nilearn, OpenNeuro.org and Neurovault.org has created new opportunities, new standards and new benchmarks against neuroimaging toolboxes need to measure themselves.

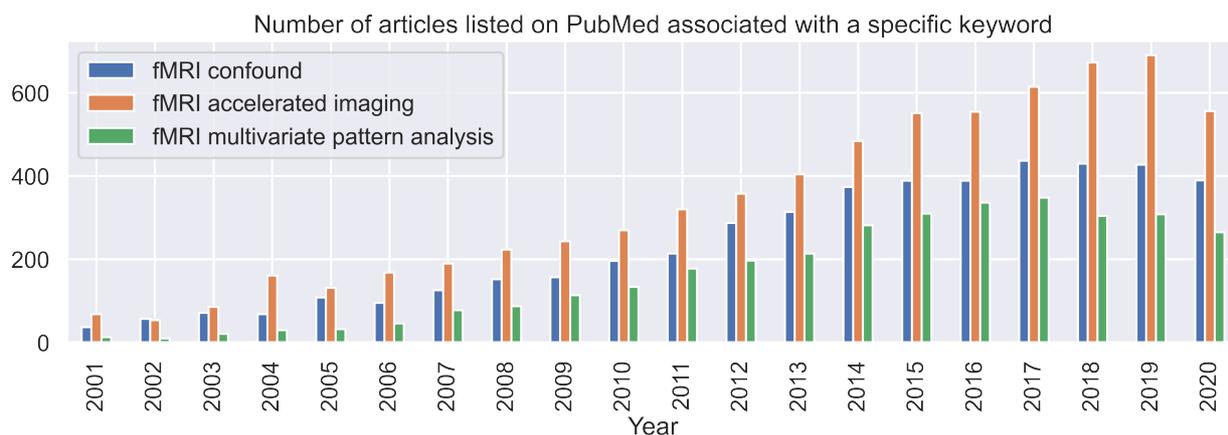


Figure 6: Number of yearly publications using a specific keyword as listed on PubMed. This histogram shows the yearly number of articles listed on PubMed which use the keyword “fMRI confound” (in blue), “fMRI accelerated imaging (in orange) or “fMRI multivariate pattern analysis” (in green) for the years 2001 to 2020.

An assessment of the total numbers of articles listed on PubMed each year (see Figure 6) reveals that multiple topics are gaining importance. The topic “fMRI confounds” addresses issues of fMRI preprocessing and how this can be improved. The topic “fMRI accelerated imaging” highlights an increased usage of accelerated acquisition sequences, and consequently also an increase of datasets recorded with higher temporal resolution. The topic “fMRI multivariate pattern analysis” shows an increased interest in MVPA. In a similar fashion, analyzing the number of publicly available datasets published on OpenNeuro.org or investigating the number of fMRI images analyzed with MRIQC which had used a multiband acceleration factor or which had a temporal resolution below 1000ms is increasing yearly (see Figure 7).

Together, Figures 6 and 7 show that the neuroimaging field is increasing the total number of freely available data, that more and more datasets are recorded with higher temporal resolutions, and that topics like confounds and multivariate approaches gain in traction every year. For these reasons, the field needs to make sure that the four issues and three challenges as addressed by this thesis are met and sufficiently addressed.

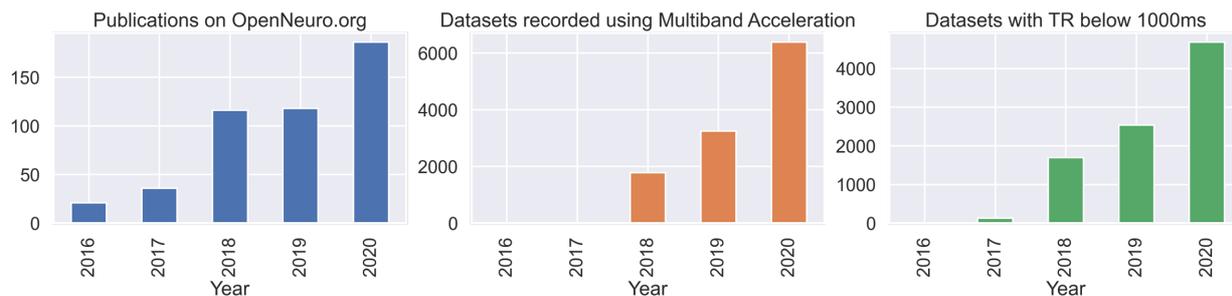


Figure 7: Number of yearly new publications on OpenNeuro.org and number of fMRI images analyzed with MRIQC which were recorded with an multiband acceleration routine or had a TR below 1000ms. The panel on the left shows in blue the number of yearly new available datasets on OpenNeuro.org. The data was acquired directly via the homepage using web scraping. The middle and right panel show the yearly total number of fMRI images which were analyzed with MRIQC, which used a multiband acceleration routine (in orange) or had a TR below 1000ms (in green). The data for the middle and right panel were acquired using MRIQC’s feely available Web API.

5.6 Future direction of fMRI-based investigation of multisensory integrations

Study C led to important new results on how univariate and multivariate analyses can lead to new insights for the domain of multisensory integration, in particular in the context of semantic congruency. Future directions of fMRI-based investigations should build on these insights and combine them with the knowledge from Study A. In particular, the research could be extended in two important directions: *First*, all multivariate and univariate stimuli should also have an equivalent of pure visual and/or auditory noise. Such a paradigm will allow the investigation of multisensory integration not just with respect to vision, but also audition. Additionally, it allows to investigate if it is important for multisensory integration if the stimuli show naturalistic stimuli, or if pure noise is also enough to elicit a comparable brain response. *Second*, increasing the temporal resolution with which an fMRI dataset is recorded, will allow to move away from the idea that univariate fMRI analysis can only analyze amplitude-based activation differences and start analyzing the data with a time-resolved analysis approach. Looking at peak onset or signal rising time is usually only reserved for electroencephalography (EEG) studies. In the context of multisensory integration, such a time-resolved approach using EEG data has already let to important new insights (Matusz et al., 2015). Other researchers have already shown, that peak latencies and slopes differences in BOLD signal activation due to unisensory and multisensory activation can be observed (Martuzzi et al., 2007). However, in that case the hemodynamic response function (HRF) had to be estimated as an intermediate tool to estimate the time sensitive information. Since the publication of that study, the scan sequences to register brain activation in the scanner have improved, and led to the before mentioned sub-second temporal resolution capability. Other researchers have shown, that increasing the temporal resolution high enough, allows us to sufficiently sample the HRF, and therefore circumventing the error prone estimation (Narsude et al., 2016).

To account for these potential two future directions of investigation, the third dataset in Study A, was recorded with a high enough temporal resolution (i.e. 600ms) and a complex enough multisensory integration paradigm. Preliminary results are already very promising (see Figure 8). They indicate that a time-resolved analysis of fMRI data is possible and that stimuli specific time curves can be observed. The preliminary results shown in Figure 8 reveal that in both primary areas (A1 and V1) the ‘noise’ conditions elicit always less brain activation as their ‘object’ counterpart, except for the non-corresponding stimuli modality (i.e. visual stimuli in A1 and auditory stimuli in V1). Additionally, multisensory congruent stimuli seem to elicit slightly more brain activation as their incongruent counterpart, but not necessarily as the corresponding unisensory stimuli. Furthermore, as shown in previous studies and in Study C, LOC seems to be particularly sensitive for stimuli with realistic multisensory stimuli (i.e. with ‘objects’) and less reactive to pure noise multisensory stimuli.

While the preliminary results depicted in Figure 8 are promising, it is important to highlight the full potential of this dataset. Already in the current analysis, differences in HRF between stimuli and brain regions can be observed. Further research will need to focus on more restrictive and smaller regions of interest and investigate the time course with respect to the aforementioned differences in peak latencies and slopes differences. Additionally, more in-depth multivariate analysis can be performed to investigate the activation patterns between these seven conditions, not just with respect to their pattern in amplitude, but also through time and other HRF specific characteristics.

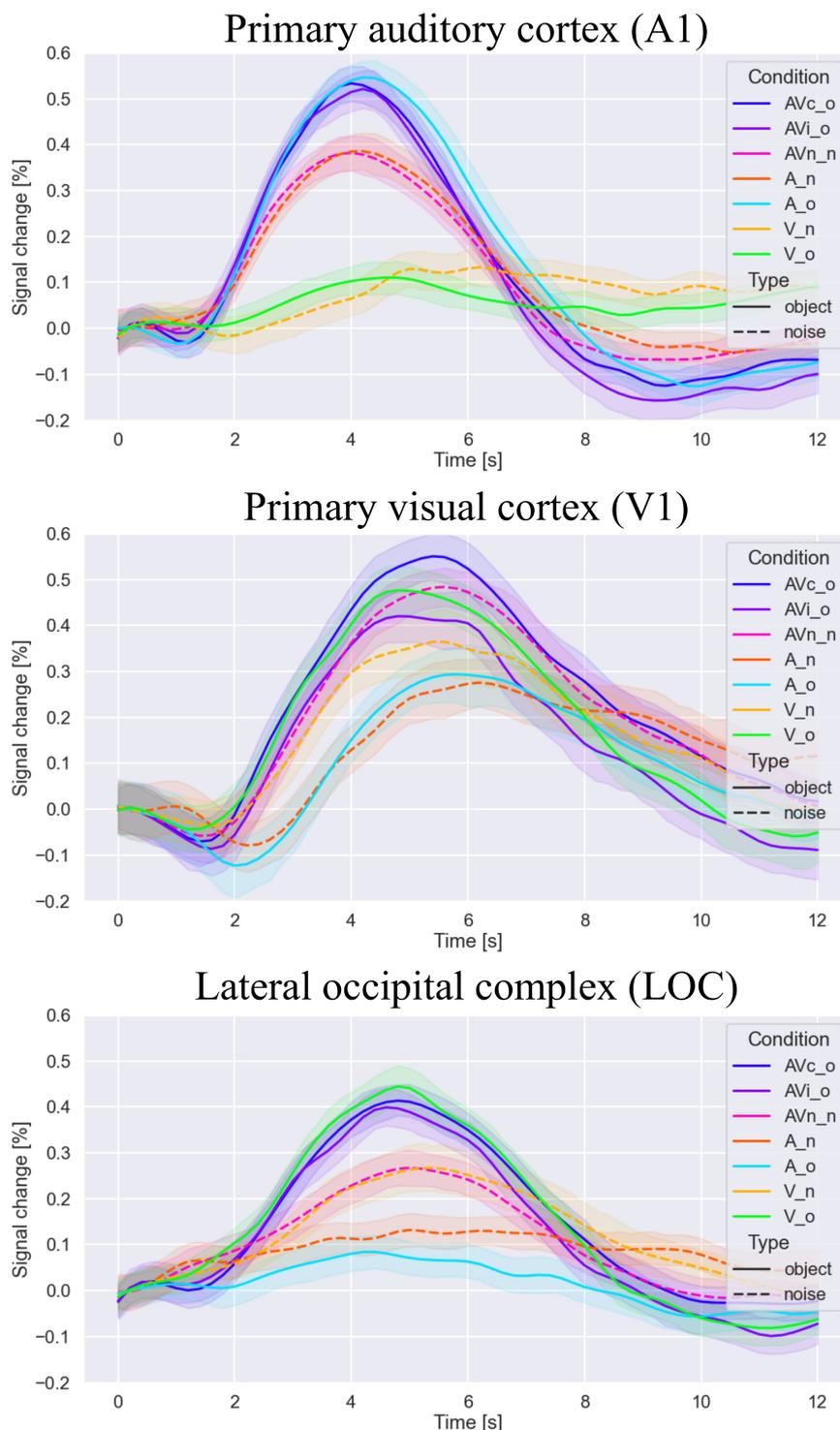


Figure 8: Depiction of time-resolved HRF signal registered in three distinct brain regions, from a multisensory integration exposure paradigm, with seven distinct stimuli types. Each panel shows the percentage of signal change (y-axis) over time in seconds (x-axis) for the seven distinct stimuli: AVc_o = Audio-visual congruent stimuli depicting an animal; AVi_o = Audio-visual incongruent stimuli depicting an animal; AVn_n = Audio-visual stimuli of pure noise; A_n = Auditory only stimuli of pure noise; A_o = Auditory only stimuli depicting an animal; V_n = Visual only stimuli of pure noise; V_o = Visual only stimuli depicting an animal. The region of interests used for this depiction were taken from Neurosynth.org and contain the top 0.1% of voxels associated with the regions (top) primary auditory cortex or A1, (middle) primary visual cortex or V1 and (bottom) lateral occipital complex (LOC). The volume covered by the three rois was respectively (A1) 322 voxels or 8694 mm³, (V1) 240 voxels or 6480 mm³ and (LOC) 281 voxels or 7587 mm³ and included bilateral regions.

5.7 Prospect

Due to the invention of many new neuroimaging toolboxes and an ever growing and supportive open-source community, the processing and analysis of fMRI data becomes more accessible, easier to do and leads to more robust results. The introduction of tools like Nipype, BIDS, fMRIPrep and OpenNeuro.org have helped the field to quickly evolve and will be responsible for many new innovations. For example, the main developers of Nipype, which initially was created to provide an interface to common neuroimaging pipelines, are currently working on a more advanced toolbox called Pydra (Jarecka et al., 2020). Pydra provides the same functionality as Nipype, but can be considered a general-purpose dataflow engine, allowing the integration of more advanced machine learning routines, as well as the potential analysis of other neuroimaging modalities such as EEG. The introduction of BIDS (Gorgolewski et al., 2016), initially only intended for MRI datasets, led to the development of BIDS standards for other modalities, such as EEG (Pernet et al., 2019), iEEG (intracranial EEG) (Holdgraf et al., 2019), MEG (magnetoencephalography) (Niso et al., 2018), PET (positron emission tomography), microelectrode recordings (MER), genetic information and eye tracking data, amongst others. Furthermore, more and more neuroimaging toolboxes package their functionality into Docker containers and provide their service in standalone and operating system independent BIDS Apps (see <https://bids-apps.neuroimaging.io/apps/>). The introduction of fMRIPrep, intended for the preprocessing of functional MRI data, has led to the development of its equivalent for structural MRI data (sMRIPrep, see <https://github.com/nipreps/smriprep>) and diffusion MRI data (dMRIPrep, see <https://github.com/nipreps/dmriprep>). The introduction of OpenfMRI.org, initially only intended for the sharing of functional MRI images, has led to the creation of OpenNeuro.org and allows the publication of any kind of MRI datasets, as well as EEG, MEG and PET datasets. The explosion of so many different innovative toolboxes, standards and supportive infrastructures is in large part due to an ever growing and welcoming open-source community in the neuroimaging field. To ensure that the strong community persists and can evolve further into the future, some of the main contributors behind these innovations created the NeuroImaging PREProcessing tools (NiPreps) framework (Esteban, Wright, et al., 2019).

5.8 Conclusion

The field of neuroimaging is ever evolving. New measuring technics, data processing and analysis approaches are developed constantly. It therefore is crucial to integrate these new innovations to current standards and adapt them to new insights from publications pointing to confounds, issues and opportunities. Consequently, good standards and processing toolboxes are needed. This thesis introduces two such new toolboxes in the form of *fMRIflows* (Study A) and *AtlasReader* (Study B) and showed how they can be used to conduct a cognitive fMRI study, using amongst others a multivariate analysis approach (Study C). As described later in this thesis, these studies created insights which can help us to guide our future direction in how we process and analyze fMRI datasets and how we can further investigate the mechanisms behind multisensory integration using fMRI. These three studies contribute to the innovation observed in the neuroimaging domain, strengthen the open-source community, allow the investigation of new neuroscience questions and support the establishment of clearer standards for processing pipelines for fMRI data analysis.

6 References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8(February), 14. <https://doi.org/10.3389/fninf.2014.00014>
- Aguirre, G. K., Zarahn, E., & D'esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *NeuroImage*, 8(4), 360–369. <https://doi.org/10.1006/nimg.1998.0369>
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, 54(3), 2033–2044. <https://doi.org/10.1016/j.neuroimage.2010.09.025>
- Baker, M. (2016). Reproducibility crisis? *Nature*, 533(26), 353–366.
- Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 23(2), 137–152. <https://doi.org/10.1109/TMI.2003.822821>
- Behzadi, Y., Restom, K., Liu, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., ... Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, 582(7810), 84–88. <https://doi.org/10.1038/s41586-020-2314-9>
- Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, 10(4–5), 171–178. [https://doi.org/10.1002/\(SICI\)1099-1492\(199706/08\)10:4/5<171::AID-NBM453>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L)
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 113(28), 7900–7905. <https://doi.org/10.1073/pnas.1602413113>
- Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PloS One*, 12(9), e0184661. <https://doi.org/10.1371/journal.pone.0184661>
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., ... Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1), 111–116. <https://doi.org/10.1038/s41592-018-0235-4>
- Esteban, O., Wright, J., Markiewicz, C. J., Thompson, W. H., Goncalves, M., Ciric, R., ... Poldrack, R. (2019). *NiPreps: enabling the division of labor in neuroimaging beyond fMRIPrep*. 7–9. <https://doi.org/10.31219/osf.io/ujxp6>
- Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Gunther, M., ... Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PloS One*, 5(12), e15710. <https://doi.org/10.1371/journal.pone.0015710>

- Feinberg, D. A., & Setsompop, K. (2013). Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *Journal of Magnetic Resonance (San Diego, Calif.: 1997)*, 229, 90–100. <https://doi.org/10.1016/j.jmr.2013.02.002>
- Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-Related fMRI: Characterizing Differential Responses. *NeuroImage*, 7(1), 30–40. <https://doi.org/10.1006/nimg.1997.0306>
- Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996). Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3), 346–355. <https://doi.org/10.1002/mrm.1910350312>
- Friston, K., Penny, W., Ashburner, J., Kiebel, S., & Nichols, T. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (Vol. 8). <https://doi.org/10.1016/B978-012372560-8/50002-4>
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage*, 9(4), 416–429. <https://doi.org/10.1006/nimg.1998.0419>
- Glover, G. H., Li, T., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine*, 44(1), 162–167. [https://doi.org/10.1002/1522-2594\(200007\)44:1<162::AID-MRM23>3.3.CO;2-5](https://doi.org/10.1002/1522-2594(200007)44:1<162::AID-MRM23>3.3.CO;2-5)
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5(August), 13. <https://doi.org/10.3389/fninf.2011.00013>
- Gorgolewski, K., Esteban, O., Schaefer, G., Wandell, B., & Poldrack, R. (2017). OpenNeuro - a free online platform for sharing and analysis of neuroimaging data. *Organization for Human Brain Mapping. Vancouver, Canada*, 1677. <https://doi.org/10.1038/sdata.2016.44.3>
- Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., ... Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3, 160044. <https://doi.org/10.1038/sdata.2016.44>
- Gorgolewski, K. J., & Poldrack, R. a. (2016). A Practical Guide for Improving Transparency and Reproducibility in Neuroimaging Research. *PLoS Biology*, 14(7), e1002506. <https://doi.org/10.1371/journal.pbio.1002506>
- Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., ... Margulies, D. S. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in Neuroinformatics*, 9(April), 8. <https://doi.org/10.3389/fninf.2015.00008>
- Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., ... Haase, A. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in*

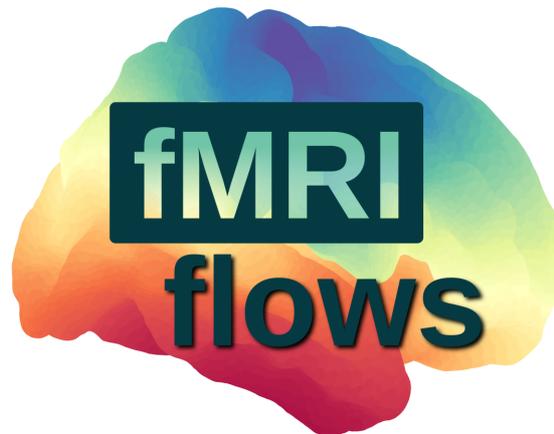
- Medicine*, 47(6), 1202–1210. <https://doi.org/10.1002/mrm.10171>
- Hajnal, J. V., Myers, R., Oatridge, A., Schwieso, J. E., Young, I. R., & Bydder, G. M. (1994). Artifacts due to stimulus correlated motion in functional imaging of the brain. *Magnetic Resonance in Medicine*, 31(3), 283–291. <https://doi.org/10.1002/mrm.1910310307>
- Halchenko, Y. O., & Hanke, M. (2015). Four aspects to make science open “by design” and not as an after-thought. *GigaScience*, 4(1), 31. <https://doi.org/10.1186/s13742-015-0072-7>
- Handwerker, D. A., Ollinger, J. M., & D’Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, 21(4), 1639–1651. <https://doi.org/10.1016/j.neuroimage.2003.11.029>
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, 7(1), 37–53. <https://doi.org/10.1007/s12021-008-9041-y>
- Havlicek, M., Ivanov, D., Poser, B. A., & Uludag, K. (2017). Echo-time dependence of the BOLD response transients – A window into brain functional physiology. *NeuroImage*, 159(July), 355–370. <https://doi.org/10.1016/j.neuroimage.2017.07.034>
- Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: the early beginnings. *NeuroImage*, 62(2), 852–855. <https://doi.org/10.1016/j.neuroimage.2012.03.016>
- Hoge, R. D., & Pike, G. B. (2001). Quantitative measurement using fMRI. *Functional Magnetic Resonance Imaging: An Introduction to Methods*, 159–174.
- Holdgraf, C., Appelhoff, S., Bickel, S., Bouchard, K., D’Ambrosio, S., David, O., ... Hermes, D. (2019). iEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology. *Scientific Data*, 6(1), 102. <https://doi.org/10.1038/s41597-019-0105-7>
- Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image Distortion Correction in fMRI: A Quantitative Evaluation. *NeuroImage*, 16(1), 217–240. <https://doi.org/10.1006/nimg.2001.1054>
- Jarecka, D., Goncalves, M., Markiewicz, C., Esteban, O., Lo, N., Kaczmarzyk, J., & Ghosh, S. (2020). Pydra - a flexible and lightweight dataflow engine for scientific analyses. *Proceedings of the 19th Python in Science Conference*, (November), 132–139. <https://doi.org/10.25080/majora-342d178e-012>
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage*, 62(2), 782–790.
- Kasper, L., Bollmann, S., Diaconescu, A. O., Hutton, C., Heinzle, J., Iglesias, S., ... Stephan, K. E. (2017). The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. *Journal of Neuroscience Methods*, 276, 56–72. <https://doi.org/10.1016/j.jneumeth.2016.10.019>
- Lindquist, M. A., Geuter, S., Wager, T. D., & Caffo, B. S. (2019). Modular preprocessing pipelines can reintroduce artifacts into fMRI data. *Human Brain Mapping*, (January), 407676. <https://doi.org/10.1002/hbm.24528>
- Liu, T. T. (2016). Noise contributions to the fMRI signal: An overview. *NeuroImage*, 143, 141–151.

<https://doi.org/10.1016/j.neuroimage.2016.09.008>

- Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology*, *66*(1), 735–769. <https://doi.org/10.1146/annurev.physiol.66.082602.092845>
- Martuzzi, R., Murray, M. M., Maeder, P. P., Fornari, E., Thiran, J. P., Clarke, S., ... Meuli, R. A. (2006). Visuo-motor pathways in humans revealed by event-related fMRI. *Experimental Brain Research*, *170*(4), 472–487. <https://doi.org/10.1007/s00221-005-0232-6>
- Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J.-P., Maeder, P. P., Clarke, S., & Meuli, R. A. (2007). Multisensory Interactions within Human Primary Cortices Revealed by BOLD Dynamics. *Cerebral Cortex*, *17*(7), 1672–1679. <https://doi.org/10.1093/cercor/bhl077>
- Matusz, P. J., Thelen, A., Amrein, S., Geiser, E., Anken, J., & Murray, M. M. (2015). The role of auditory cortices in the retrieval of single-trial auditory-visual object memories. *European Journal of Neuroscience*, *41*(5), 699–708. <https://doi.org/10.1111/ejn.12804>
- Mensch, A., Varoquaux, G., & Thirion, B. (2016). Compressed online dictionary learning for fast resting-state fMRI decomposition. *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016-June*(1), 1282–1285. <https://doi.org/10.1109/ISBI.2016.7493501>
- Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage*, *11*(6 D), 735–759. <https://doi.org/10.1006/nimg.2000.0568>
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Ugurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic Resonance in Medicine*, *63*(5), 1144–1153. <https://doi.org/10.1002/mrm.22361>
- Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI—an introductory guide. *Social Cognitive and Affective Neuroscience*, *4*(1), 101–109. <https://doi.org/10.1093/scan/nsn044>
- Narsude, M., Gallichan, D., van der Zwaag, W., Gruetter, R., & Marques, J. P. (2016). Three-dimensional echo planar imaging with controlled aliasing: A sequence for high temporal resolution functional MRI. *Magnetic Resonance in Medicine*, *75*(6), 2350–2361. <https://doi.org/10.1002/mrm.25835>
- Neumann, J., Lohmann, G., Zysset, S., & von Cramon, D. Y. (2003). Within-subject variability of BOLD response dynamics. *NeuroImage*, *19*(3), 784–796. [https://doi.org/10.1016/S1053-8119\(03\)00177-0](https://doi.org/10.1016/S1053-8119(03)00177-0)
- Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., ... Yeo, B. T. T. (2017). Best practices in data analysis and sharing in neuroimaging using MRI. *Nature Neuroscience*, *20*(3), 299–303. <https://doi.org/10.1038/nn.4500>
- Niso, G., Gorgolewski, K. J., Bock, E., Brooks, T. L., Flandin, G., Gramfort, A., ... Baillet, S. (2018). MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Scientific Data*, *5*, 1–5. <https://doi.org/10.1038/sdata.2018.110>

- Notter, M., Gale, D., Herholz, P., Markello, R., Notter-Bielser, M.-L., & Whitaker, K. (2019). AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI images. *Journal of Open Source Software*, 4(34), 1257. <https://doi.org/10.21105/joss.01257>
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(24), 9868–9872. <https://doi.org/10.1073/pnas.87.24.9868>
- Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, 10(JUL), 27. <https://doi.org/10.3389/fninf.2016.00027>
- Pauli, R., Bowring, A., Reynolds, R., Chen, G., Nichols, T. E., & Maumet, C. (2016). Exploring fMRI Results Space: 31 Variants of an fMRI Analysis in AFNI, FSL, and SPM. *Frontiers in Neuroinformatics*, 10(July), 24. <https://doi.org/10.3389/fninf.2016.00024>
- Pernet, C. R., Appelhoff, S., Gorgolewski, K. J., Flandin, G., Phillips, C., Delorme, A., & Oostenveld, R. (2019). EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Scientific Data*, 6(1), 103. <https://doi.org/10.1038/s41597-019-0104-8>
- Sengupta, A., Pollmann, S., & Hanke, M. (2018). Spatial band-pass filtering aids decoding musical genres from auditory cortex 7T fMRI. *F1000Research*, 7(0), 142. <https://doi.org/10.12688/f1000research.13689.2>
- Thomas, C. G., Harshman, R. A., & Menon, R. S. (2002). Noise reduction in BOLD-based fMRI using component analysis. *NeuroImage*, 17(3), 1521–1537. <https://doi.org/10.1006/nimg.2002.1200>
- Triantafyllou, C., Hoge, R. D., Krueger, G., Wiggins, C. J., Potthast, A., Wiggins, G. C., & Wald, L. L. (2005). Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. *NeuroImage*, 26(1), 243–250. <https://doi.org/10.1016/j.neuroimage.2005.01.007>
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6), 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>
- Viessmann, O., Möller, H. E., & Jezzard, P. (2018). Dual regression physiological modeling of resting-state EPI power spectra: Effects of healthy aging. *NeuroImage*, 187(December 2017), 68–76. <https://doi.org/10.1016/j.neuroimage.2018.01.011>
- Vovk, U., Pernus, F., & Likar, B. (2007). A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Transactions on Medical Imaging*, 26(3), 405–421. <https://doi.org/10.1109/TMI.2006.891486>
- Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 4(3), 274–290. <https://doi.org/10.1111/j.1745-6924.2009.01125.x>

Annex 1



Study A – fMRIflows: a consortium of fully automatic univariate and multivariate fMRI processing pipelines

Michael P. Notter, Peer Herholz, Sandra Da Costa,
Omer F. Gulban, Ayse I. Isik and Micah M. Murray

Submitted to *Neuroimage*, 2021 Mar 23.

fMRIflows: a consortium of fully automatic univariate and multivariate fMRI processing pipelines

Authors: **Michael P. Notter**^{1*}, Peer Herholz^{2,3*}, Sandra Da Costa⁴, Omer F. Gulban^{5,6}, Ayse Ilkay Isik⁷ and Micah M. Murray^{1,4,8-9}

1. The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, University Hospital Center and University of Lausanne, Switzerland
2. International Laboratory for Brain, Music and Sound Research, Université de Montréal & McGill University, Montréal, Canada
3. McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, Canada
4. CIBM Center for Biomedical Imaging, Lausanne, Switzerland
5. Department of Cognitive Neuroscience, Maastricht University, The Netherlands
6. Brain Innovation B.V., Maastricht, The Netherlands
7. Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany
8. Department of Ophthalmology, Fondation Asile des aveugles and University of Lausanne, Lausanne, Switzerland
9. Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

** shared first co-authorship*

Address correspondence to:

Michael P. Notter
Chemin de la Milliere 1
1040 Villars-le-Terroir
Switzerland
Michael.Notter@epfl.ch

Abstract

How functional MRI (fMRI) data are analyzed depends on the researcher and the toolbox used. It is not uncommon that the processing pipeline is rewritten for each new dataset. Consequently, code transparency, quality control and objective analysis pipelines are important for improving reproducibility in neuroimaging studies. Toolboxes, such as Nipype and fMRIPrep, have documented the need for and interest in automated analysis pipelines. Here, we introduce fMRIflows: a consortium of fully automatic neuroimaging pipelines for fMRI analysis, which performs standard preprocessing, as well as 1st- and 2nd-level univariate and multivariate analysis. In addition to the standardized processing pipelines, fMRIflows also provides flexible temporal and spatial filtering to account for datasets with increasingly high temporal resolution and to help appropriately prepare data for multivariate analysis and improve signal decoding accuracy. This paper *first* describes fMRIflows' structure and functionality, *then* explains its infrastructure and access, and *lastly* validates the toolbox by comparing it to other neuroimaging processing pipelines such as fMRIPrep, FSL and SPM. This validation was performed on three datasets with varying temporal resolution to ensure flexibility and robustness, as well as to showcase the improved capability of fMRIflows. The outcome of the validation analysis shows that fMRIflows preprocessing pipeline performs comparably to the ones obtained from other toolboxes. Importantly, fMRIflows' flexible temporal filtering approach leads to improved signal-to-noise-ratio after preprocessing and increased statistical sensitivity, particularly in datasets with high temporal resolution. fMRIflows is a fully automatic fMRI processing pipeline which uniquely offers univariate and multivariate single-subject and group analyses as well as preprocessing. It offers flexible spatial and temporal filtering and thereby leads to more pronounced results for datasets with temporal resolutions at or below 1000ms.

Keywords: Python, neuroimaging, data processing, pipeline, reproducible research

1 Introduction

Functional magnetic resonance imaging (fMRI) is a well-established neuroimaging method used to analyze activation patterns in order to understand brain function. A full fMRI analysis includes preprocessing of the data, followed by statistical analysis and inference on the results, usually separated into 1st-level analysis (the statistical analysis within subjects) and 2nd-level analysis (the group analysis between subjects). The goal of preprocessing is to identify and remove nuisance sources, measure confounds, apply temporal and spatial filters and to spatially realign and normalize images to make them spatially conform (Caballero-Gaudes & Reynolds, 2017). A good preprocessing pipeline should improve the signal-to-noise ratio (SNR) of the data, ensure validity of inference and interpretability of results (Ashburner, 2009), reduce false positive and false negative errors in the statistical analysis and therefore improve the statistical power.

Even though the consequences of inappropriate preprocessing and statistical inference are well documented (Power, Plitt, Laumann, & Martin, 2017; Strother, 2006), most fMRI analysis pipelines are still established ad-hoc, subjectively customized by researchers to each new dataset (Carp, 2012). This usage can be explained by the circumstance that most researchers, by habit or lack of time, stick with the neuroimaging software at-hand or reuse and modify scripts and code snippets from colleagues and previous projects, and do not always adapt their processing pipelines to the newest standard in neuroimaging processing. Rehashing processing pipelines is associated with problems like persisting bugs in the code and delays in updating individual analysis steps to the most recent standards. This can lead to far reaching consequences. Of course, the constant updating of pipelines to newest standards and softwares also bears the risk of introducing new bugs and might lead to the pitfall of blindly trusting new untested procedures.

One solution to tackle this issue will require code transparency, good quality control and a collective development of well-tested objective analysis pipelines (Gorgolewski et al., 2016). Recent years have brought some important reformations to the neuroimaging community that go in this direction.

First, the introduction of Nipype (Gorgolewski et al., 2011) made it easier for researchers to switch between different neuroimaging toolboxes, such as AFNI (Cox & Hyde, 1997), ANTs (Avants et al., 2011), FreeSurfer (Fischl, 2012), FSL (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012), and SPM (Friston, Penny, Ashburner, Kiebel, & Nichols, 2006). Nipype together with other software packages such as Nibabel (Brett et al., 2018) and Nilearn (Abraham et al., 2014) opened up the whole Python ecosystem to the neuroimaging community. Code can be shared between researchers via online services such as GitHub (<https://github.com>), and the whole neuroimaging software ecosystem can be run on any machine or server through the use of container software such as Docker (<https://www.docker.com>) or Singularity (<https://www.sylabs.io>). Combined with a continuous integration service such as CircleCI (<https://circleci.com>) or TravisCI (<https://travis-ci.org>), this allows the creation of easy-to-read, transparent, shareable and continuously tested open-source neuroimaging processing pipelines.

Second, the next major advancement in the neuroimaging field was the introduction of a common dataset standard, such as the NIfTI standard (<https://nifti.nimh.nih.gov/>). This was important for the formatting of neuroimaging data. The neuroimaging community gathered together in a consortium to define a standard format for the storage of neuroimaging datasets, the so-called Brain Imaging Data Structure (BIDS) (Gorgolewski et al., 2016). A common data structure format facilitates the sharing of datasets and makes it possible to create universal neuroimaging toolboxes that work out-of-the-box on any BIDS conforming dataset. Additionally, through services like OpenNeuro (Gorgolewski, Esteban, Schaefer, Wandell, & Poldrack, 2017), a free online platform for sharing neuroimaging data, one can test the robustness and flexibility of a new neuroimaging toolbox on hundreds of different datasets.

Software toolboxes like MRIQC (Esteban et al., 2017) and fMRIPrep (Esteban et al., 2019) have shown how fruitful this new neuroimaging ecosystem can be and have highlighted the importance and need of good quality control and high-quality preprocessing workflows with consistent results from diverse datasets. Thus, the present software package, called fMRIflows, will be based on this new ecosystem and toolboxes, and will expand it to fully automated pipelines for univariate and multivariate individual and group analyses.

Moreover, fMRIflows provides flexible temporal and spatial filtering, to account for two recent findings in the field. *First*, flexible spatial filtering can become of importance when performing multivariate analysis, as it has been shown that the correct spatial band-pass filtering can improve signal decoding accuracy (Sengupta, Pollmann, & Hanke, 2018). *Second*, correct temporal filtering during preprocessing is important and can lead to improved signal-to-noise ratio (SNR), especially for fMRI datasets with a temporal resolution rate below one second (Viessmann, Möller, & Jezzard, 2018), but only if the filter is applied orthogonally to the other filters during preprocessing to ensure that previously removed noise is not again reintroduced into the data (Lindquist, Geuter, Wager, & Caffo, 2019). Due to technical improvements in imaging recording through acceleration techniques such as GRAPPA (Griswold et al., 2002) and simultaneous multi-slice/multiband acquisitions (Feinberg et al., 2010; Feinberg & Setsompop, 2013; Moeller et al., 2010), faster sampling rates became possible, to the point that respiratory and cardiac signals can be sufficiently sampled in the BOLD signal. This creates new challenges for the preprocessing of functional images, especially when the external recording of those physiological sources cannot be achieved. fMRIflows tackles this challenge in an innovative way by leveraging notebook format.

fMRIflows presents a consortium of fully automatic neuroimaging pipelines for fMRI analysis, performing standardized preprocessing, as well as 1st- and 2nd-level analyses for univariate and multivariate analysis, with the additional creation of informative quality control figures. fMRIflows is predicated on the insights and code base of MRIQC (Esteban et al., 2017) and fMRIPrep (Esteban et al., 2019), extending their functionality with regard to the following aspects: (a) flexible temporal and spatial filtering of fMRI data, i.e. low- or band-pass filtering allowing for the exclusion of high-frequency oscillations introduced through physiological noise (Viessmann et al., 2018); (b) accessible and modifiable

code base; (c) automatic computation of 1st-level contrasts for univariate and multivariate analysis; and (d) automatic computation of 2nd-level contrasts for univariate and multivariate analysis.

In this paper, we (1) describe the different pipelines included in fMRIflows and illustrate the different processing steps involved, (2) explain the software structure and setup, and (3) validate fMRIflows' performance by comparing it to other widely used neuroimaging toolboxes, such as fMRIPrep (Esteban et al., 2019), FSL (Jenkinson et al., 2012) and SPM (Friston et al., 2006).

2 Materials and Methods

2.1.1 fMRIflows' processing pipelines

The complete code base of fMRIflows is open access and stored conveniently in six different Jupyter notebooks on <https://github.com/miykael/fmriflows>. The first notebook does not contain any processing pipeline, but rather serves as a user input document that helps to create JSON files, which will contain the execution specific parameters for the five processing pipelines contained in fMRIflows: (1) anatomical preprocessing, (2) functional preprocessing, (3) 1st-level analysis, (4) 2nd-level univariate analysis and (5) 2nd-level multivariate analysis. Each of these five pipelines stores its results in a sub hierarchical folder, specified as an output folder by the user. In the following section, we explain the content of those six Jupyter notebooks.

Specification preparation

Each fMRIflows processing pipeline needs specific input parameters to run. Those parameters range from subject ID and, number of functional runs per subject, to requested voxel resolution after image normalization, etc. Each notebook will read the relevant specification parameters from a predefined JSON file that starts with the prefix "fmriflows_spec". There is one specification file for the anatomical and functional preprocessing, one for the 1st and 2nd level univariate analysis, and one for the 2nd-level multivariate analysis. For an example of these three JSON files, see Supplementary Note 1. The first notebook contained in fMRIflows, called `01_spec_preparation.ipynb`, can be used to create those JSON files, based on the provided dataset and some standard default parameters. It does so by using Nibabel v2.3.0 (Brett et al., 2018), PyBIDS v0.8 (Yarkoni et al., 2019) and other standard Python libraries. It is up to the user to change any potential processing parameter should they be different from the used default values.

Anatomical preprocessing

The anatomical preprocessing pipeline is contained within the notebook `02_preproc_anat.ipynb` and uses the JSON file `fmriflows_spec_preproc.json` for parameter specification such as voxel resolution. If a specific value is not set, `fmriflows` normalizes to an isometric voxel resolution of 1 mm^3 by default. However, the user can also choose an anisometric voxel resolution that varies in all three dimensions. Additionally, the user can decide to have a fast or precise normalization. The precise normalization can take up to eight times as long as the fast approach but guarantees a normalization of high precision. For an example of the JSON file content, see Supplementary Note 1.

The anatomical preprocessing pipeline only depends on the subject specific T1-weighted (T1w) anatomical images as input files. The individual processing steps are visualized in Figure 1 and consist of: (1) image reorientation, (2) cropping of field of view (FOV), (3) correction of intensity non-uniformity (INU), (4) image segmentation, (5) brain extraction and (6) image normalization. For a more detailed description of the steps involved in this processing pipeline, see Supplementary Note 2.

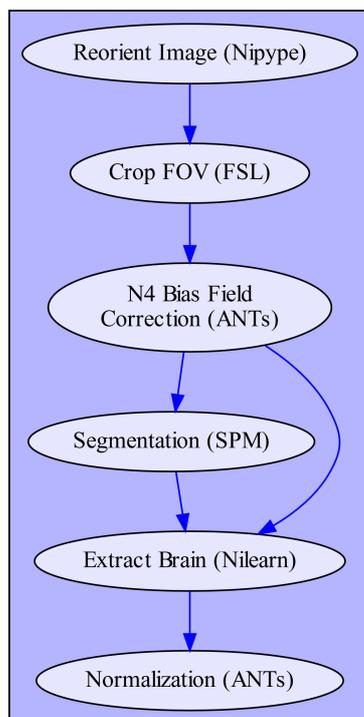


Figure 1: Depiction of `fmriflows`' anatomical preprocessing pipeline. Arrows indicate dependency between the different processing steps and data flow. Name of each node describes functionality, with the corresponding software dependency mentioned in brackets.

Functional preprocessing

The functional preprocessing pipeline is contained within the notebook `03_preproc_func.ipynb` and uses the JSON file `fmriflows_spec_preproc.json` for parameter specification. As specification parameters, users can indicate if slice-time correction should be applied or not, and if so which reference timepoint should be used. The user can also indicate to which isometric or anisometric voxel resolution functional images should be sampled to, and if the sampling is into subject or template space. For the template space, the ICBM 2009c nonlinear asymmetric brain template is used (Fonov et al., 2011). Furthermore, users can specify possible values for low-, high- or band-pass filters in the temporal or spatial domain. Additionally, to investigate nuisance regressors, users can specify the number of CompCor (Behzadi, Restom, Liao, & Liu, 2007) or independent component analysis (ICA) components they want to extract and which threshold values they want to use to detect outlier volumes. The implications of those parameters will be explained in more details in the following sections. For an example of the JSON file content, see Supplementary Note 1.

The functional preprocessing pipeline depends as inputs on the output files from the anatomical preprocessing pipeline, as well as the subject-specific functional images and accompanying descriptive JSON file that contains information about the temporal resolution (TR) and slice order of the functional image recording. This JSON file is part of the BIDS standard and therefore should be available in the BIDS conform dataset. The individual processing steps are schematized in Figure 2 and consist of: (1) image reorientation, (2) non-steady-state detection, (3) creation of functional brain mask, (4) slice time correction, (5) estimation of motion parameters, (6) two-step estimation of coregistration parameters between functional and anatomical image, (7) finalization of motion parameters, (8) single-shot spatial interpolation applying motion correction, coregistration and if specified normalizing images to the template image, (9) construction and application of brain masks, (10) temporal filtering and (11) spatial filtering. It is important to mention that the functional preprocessing is done for each functional run separately to prevent inter-run contaminations. For a more detailed description of the steps involved in this processing pipeline, see Supplementary Note 3.

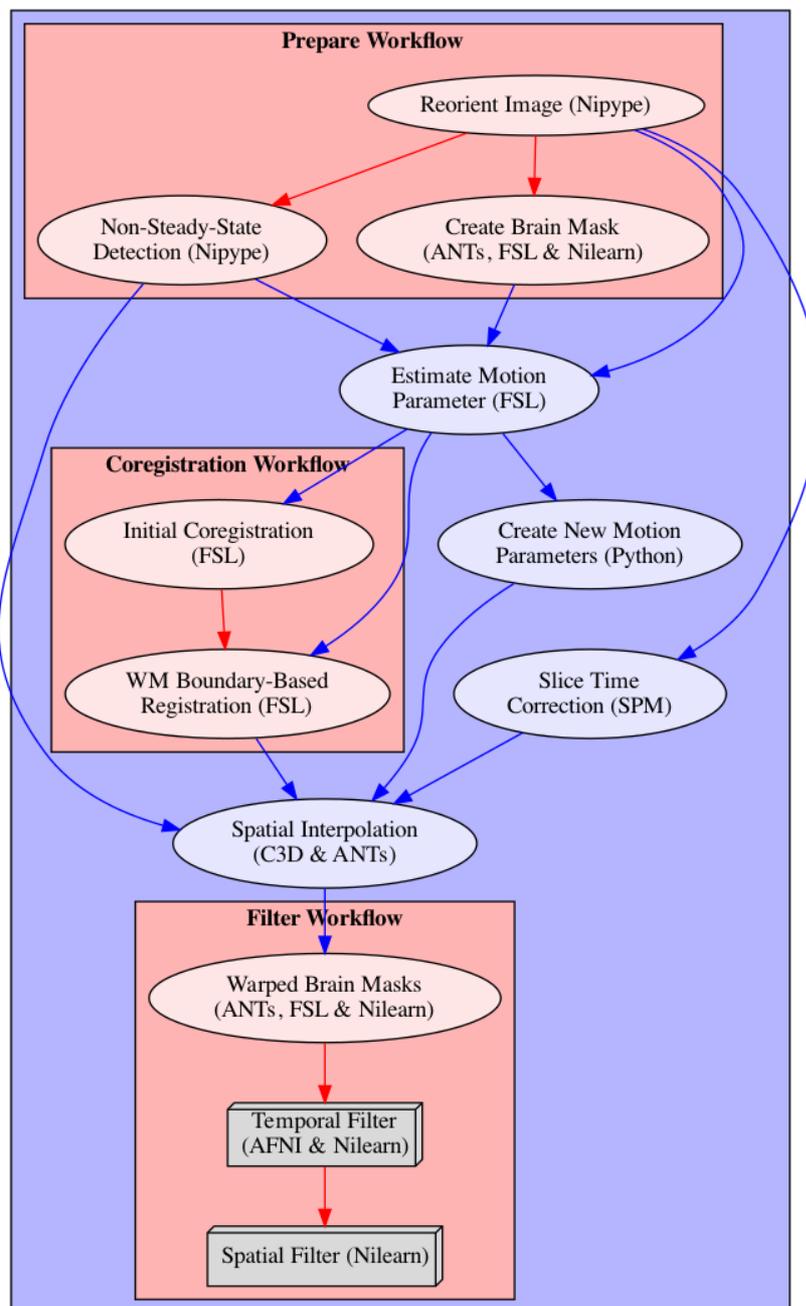


Figure 2: Depiction of fMRIfloWS' functional preprocessing pipeline. Arrows indicate dependency between the different processing steps and data flow. Name of each node describes functionality, with the corresponding software dependency mentioned in brackets. Steps that can be grouped into specific sections are contained within a red box to facilitate understanding of the pipeline. Color of arrows indicated if connection stays within a section (red) or not (blue). Nodes depicted as gray boxes indicate that they can be run multiple times with iterating input values, i.e. performing a spatial smoothing with an FWHM value of 4 and 8mm.

1st-level analysis

The first level analysis pipeline is contained within the notebook `04_analysis_1st-level.ipynb` and uses the JSON file `fmriflows_spec_analysis.json` for parameter specification. As specification parameters, users can indicate which nuisance regressors to include in the GLM, if outliers should be considered, and if the data is already in template space or if this normalization should be done after the estimation of the contrasts. Users can also specify other GLM model parameters, such as the high-pass filter value and the type of basis function that should be used to model the haemodynamic response function (HRF). Additionally, the users will also specify a list of contrasts they want to be estimated, or if they want to create specific contrasts for each stimulus column in the design matrix, and/or for each session separately, which then later might also be used for multivariate analysis. For an example of the JSON file content, see Supplementary Note 1.

The 1st-level analysis pipeline depends on a number of outputs from the previous anatomical and functional preprocessing pipelines, i.e. the TSV (tab separated value) file containing motion parameters and confound regressors, a text file indicating the number of non-steady-state volumes removed from the functional image, and a text file containing a list of indexes identifying outlier volumes. Additionally, the 1st-level analysis pipeline also requires BIDS conform events files containing information on the applied experimental design, including types of conditions and their respective onsets and durations. The individual processing steps included in the 1st-level analysis consist of: (1) collecting preprocessed files and model relevant information, (2) model specification and estimation, (3) univariate contrast estimation, (4) optional preparation for multivariate analysis, (5) optional spatial normalization of contrasts (Figure 3). All of the relevant steps, that is model creation, estimation and contrast computation are performed with SPM version 12. For a more detailed description of the steps involved in this processing pipeline, see Supplementary Note 4.

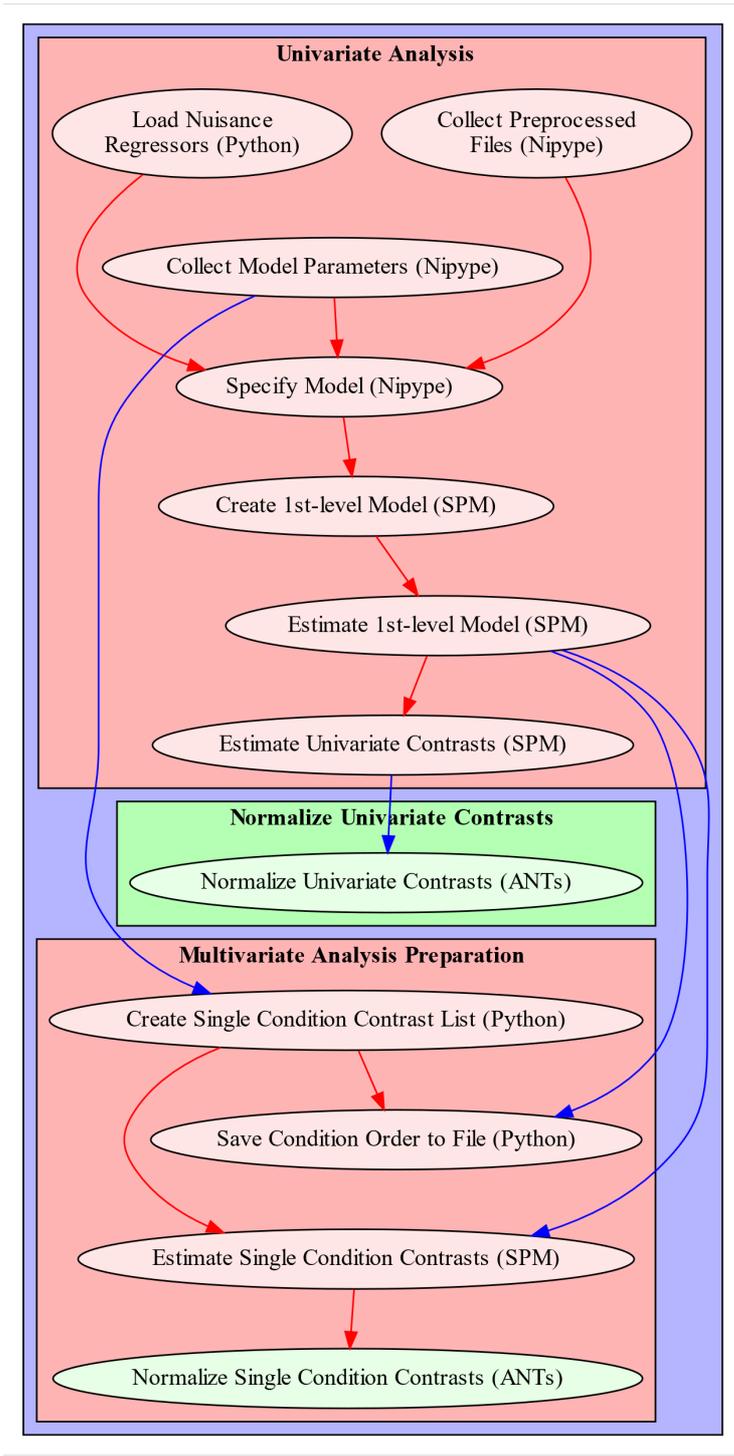


Figure 3: Depiction of fMRIflores’ 1st-level analysis pipeline. Arrows indicate dependency between the different processing steps and data flow. Name of each node describes functionality, with the corresponding software dependency mentioned in brackets. Sections that can be grouped into specific sections are contained within a red box to facilitate understanding of the pipeline. Color of arrows indicated if connection stays within a section (red) or not (blue). Nodes depicted in green are optional and can be run if spatial normalization was not yet performed during preprocessing.

2nd-level univariate analysis

The second level univariate analysis pipeline is contained within the notebook `05_analysis_2nd-level.ipynb` and uses the JSON file `fmriflows_spec_analysis.json` for parameter specification. Users can specify the probability value used as a cutoff for the threshold of the GM probability tissue map in template space that is later used during the model estimation. Additionally, users can specify voxel- and cluster-threshold topological thresholding of the statistical contrast, as well as relevant AtlasReader (Notter et al., 2019) parameters for the creation of the output tables and figures.

The 2nd-level univariate analysis pipeline depends only on the estimated contrasts from the 1st-level univariate analysis. No further contrast specification is required as `fMRIfflows` currently only implements a simple one-sample T-test. The individual processing steps included in the 2nd-level univariate analysis consist of: (1) gathering of the 1st-level contrasts, (2) creation and estimation of 2nd-level model, (3) estimation of contrast estimation, (4) topological thresholding of contrasts, (5) results creation with AtlasReader. As for the 1st-level analysis, all of the relevant model creation, estimation and contrast computation are performed with SPM12. For a more detailed description of the steps involved in this processing pipeline, see Supplementary Note 5.

2nd-level multivariate analysis

The second level multivariate analysis pipeline is contained within the notebook `06_analysis_multivariate.ipynb` and uses the JSON file `fmriflows_spec_multivariate.json` for parameter specification. Users can define a list of classifiers to use for the multivariate analysis, the sphere radius and step size of the searchlight approach. To perform a 2nd-level analysis of searchlight results users can decide between a classical GLM approach testing against chance level and a more recommended permutation based method as described in Stelzer, Chen, & Turner (2013) with the option of determining the number of permutations. Additionally, users can specify voxel- and cluster-threshold topological thresholding of the statistical contrast, as well as relevant AtlasReader parameters for the creation of the output tables and figures.

The 2nd-level multivariate analysis pipeline depends on the estimated contrasts from the 1st-level multivariate analysis, the associated CSV file containing a list of the corresponding contrast labels and a list of binary classification identifiers. In contrast to the other notebooks, this notebook uses Python 2.7 to accommodate the requirements of PyMVPA v2.6.5 (Hanke et al., 2009). The individual processing steps included in the 2nd-level multivariate analysis consist of: (1) data preparation for the analysis with PyMVPA, (2) searchlight classification, (3) computation of group analysis using a T-test, (4) computation of group analysis according to Stelzer et al. (2013), and (5) results creation with AtlasReader. For a more detailed description of the steps involved in this processing pipeline, see Supplementary Note 6.

2.1.2 Infrastructure and access to fMRIfflows

The source code of fMRIfflows is available at GitHub (<https://github.com/miykael/fmrifflows>) and is licensed under the BSD 3-Clause “New” or “Revised” License. The code is written in Python v3.7.2 (<https://www.python.org>), stored in Jupyter Notebooks v4.4.0 (Kluyver et al., 2016) and distributed via Docker v18.09.2 (<https://docker.com>) containers that are publicly available via Docker Hub (<https://hub.docker.com>). The usage of Docker allows the user to run fMRIfflows on any major operating system, with the following command:

```
docker run -it -p 9999:8888 -v /home/user/ds001:/data miykael/fmrifflows
```

The first flag `-it` indicates that the docker container should be run in interactive mode, while the second flag `-p 9999:8888` defines the port (here 9999) that we want to use to access the Jupyter Notebooks via the web-browser. The third flag, `-v /home/user/ds001:/data` tells fMRIfflows the location of the BIDS conform dataset that should be mounted in the docker container, here located at `/home/user/ds001`. Once the docker container is launched, the interactive Jupyter Notebooks can be accessed through the web-browser.

fMRIfflows uses many different software packages for the individual processing steps. The neuroimaging software that are used are: Nipype v1.1.9 (Gorgolewski et al., 2011), FSL v5.0.9 (Smith et al., 2004), ANTs v2.2.0 (Avants et al., 2011), SPM12 v7219 (Penny, Friston, Ashburner, Kiebel, & Nichols, 2011), AFNI v18.0.5 (Cox & Hyde, 1997), Nilearn v0.5 (Abraham et al., 2014), Nibabel v2.3.0 (Brett et al., 2018), PyMVPA v2.6.5 (Hanke et al., 2009), Convert3D v1.1 (<https://sourceforge.net/p/c3d>), AtlasReader v0.1 (Notter et al., 2019) and PyBIDS v0.8 (Yarkoni et al., 2019). In addition to some standard Python libraries, fMRIfflows also uses Numpy (Oliphant, 2007), Scipy (Jones, Oliphant, Peterson, & others, 2001), Matplotlib (Hunter, 2007), Pandas (McKinney & others, 2010) and Seaborn (<http://seaborn.pydata.org>).

With every new pull request pushed to the GitHub repository of fMRIfflows, a test instance on CircleCI (<https://circleci.com>) is deployed to test the complete code base for execution errors. This framework allows the continuous integration of new code to fMRIfflows, and guarantees the general functionality of the software package. Outputs are not controlled for their correctness.

2.1.3 Validation of fMRIfflows

fMRIfflows was validated in two phases. In *Phase 1*, we validated the proficiency of the toolbox by applying it on different kinds of fMRI datasets conforming to the BIDS standard (Gorgolewski et al., 2016) available via OpenNeuro.org (Gorgolewski et al., 2017). Insights during this phase allowed us to improve the code base and make fMRIfflows robust to a diverse set of datasets. In *Phase 2*, we compared the performance of the toolbox to similar neuroimaging preprocessing pipelines such as fMRIPrep, FSL, and

SPM. To better understand where fMRIfloWS overlaps or diverges from comparable processing pipelines, we investigated the preprocessing, subject-level and group-level outcomes for all four toolboxes, run on three different datasets.

Phase 1: Proficiency validation

To investigate the capabilities and flaws of the initial implementation of the toolbox, fMRIfloWS was run on different datasets, either available publicly via OpenNeuro.org or available privately to the authors. Such an approach allowed the exploration of datasets with different temporal and spatial resolutions, SNRs, FOVs, numbers of slices, scanner characteristics, and other sequence parameters, such as acceleration factors and flip angles.

Phase 2: Performance validation

To validate the performance of fMRIfloWS, we used three different task-based fMRI datasets and compared its preprocessing to the three neuroimaging processing pipelines fMRIPrep, FSL and SPM. Comparison was done on preprocessing, subject-level and group-level outputs. Because of differences in how FSL and SPM perform subject- and group-level analyses and due to the lack of such routines in fMRIPrep, all subject- and group-level analyses for the performance validation were performed using identical Nistats (Abraham et al., 2014) routines.

The three datasets (see Table 1) were all acquired on scanners with a magnetic field strength of 3 Tesla and differ in many sequence parameters, most notably in the temporal resolution with which they were recorded. This is especially important as we aim to highlight that the right handling of temporal filtering is crucial for datasets with a temporal resolution below 1000ms.

Table 1: Overview of the datasets used to validate fMRIflows.

Dataset	TR2000	TR1000	TR600
Temporal resolution	2000ms	1000ms	600ms
Spatial resolution	3.5 x 3.5 x 3.3	2.0 x 2.0 x 2.4	3.0 x 3.0 x 3.0
Number of slices	36	64	24
Slice Order	Descending	Unknown	Interleaved
Coverage	Whole brain	Whole brain	Slab
Volumes per run	275	453	600
Number of runs	4	4	6
Acceleration Factor	None	4	3
Magnetic strength	3 Tesla	3 Tesla	3 Tesla
Number of subjects	12	20	17
Sequence type	2D-EPI	Multi-Band	SMS
Task	Audio-visual memory task	Mixed gamble task	Audio-visual observation task
Data Availability	OpenNeuro.org (ds001345, v 1.0.1)	OpenNeuro.org (ds001734, v.1.0.4)	OpenNeuro.org (will be made available after publication of experimental work)

Dataset TR2000 has a comparably low temporal and spatial resolution. It serves as a standard dataset, recorded with a standard EPI scan sequence. The dataset and paradigm are described in more details in Notter et al. (under review). In short, participants performed a continuous recognition task and indicated for each image whether it is old or new. When the image was presented for the first time (new) it was either presented with no sound (unisensory visual context) or together with a sound (multisensory context).

Dataset TR1000 has a rather high temporal and spatial resolution and serves as an advanced dataset, recorded with a scan sequence using a multiband acceleration technique. The dataset and paradigm are described in more detail in Botvinik-Nezer et al. (2019). In short, participants performed a mixed gambling task in which they were asked to either accept or reject a possible monetary gain or loss.

Dataset TR600 has a very high temporal resolution with a moderate spatial resolution and serves as an extreme dataset, recorded with scan sequences using a simultaneous multi-slice (SMS) acceleration technique (Feinberg et al., 2010). This dataset was collected for another project. In short, participants were shown auditory, visual or audiovisual stimuli containing either an animal (as an image or sound), pure noise or both together. Participants performed a discrimination task in which they had to indicate if they perceived a stimuli with an animal in it or not, independent of the stimuli modality. The stimuli were either presented in a unisensory or multisensory context.

All participants of the Datasets TR2000 and TR600 have been included in the performance validation, while only the first 20 out of the 120 total participants of the Dataset TR1000 was used in order

to reduce computation time and make this dataset comparable to the other two. Datasets TR2000 and TR1000 are already publicly available through the OpenNeuro platform. Dataset TR600 is in preparation to be published on OpenNeuro as well. Until then, this dataset is available upon request.

The **preprocessing pipelines** with fMRIfloWS, fMRIPrep, FSL and SPM were based on the default parameters and only differed in the following points from their standard implementations: (1) Functional images were resampled to an isometric voxel resolution according to the dominant resolution dimension within a dataset; (2) Spatial smoothing of the functional images is applied after preprocessing of the images, using a Nilearn routine and a smoothing kernel with a Full Width at Half Maximum (FWHM) of 6mm, in order to keep the approaches comparable, as fMRIPrep does not allow spatial smoothing of functional images; (3) Anatomical images in the FSL pipeline were first cropped to a standard FOV, followed by brain extraction using FSL's BET before FSL's FEAT was launched; (4) In the case of FSL, the normalization from structural to standard space was done using a non-linear warping approach with 12 degrees of freedom and a spline interpolation model; (5) In the case of SPM, the template brain for the normalization was its standard tissue probability brain TPM, while for fMRIfloWS, fMRIPrep and FSL, the ICBM 2009c nonlinear asymmetric brain template was used.

The statistical inference was not performed on any of the investigated toolboxes to prevent the introduction of a software specific bias. The 1st- and 2nd-level analysis was performed using Nistats, Nilearn and other Python toolboxes and only differed between the toolboxes in the following ways: (1) the estimated motion parameters added to the design matrix during the 1st-level analysis differed for each toolbox as they were based on the software-specific preprocessing routine; (2) the number of volumes per run used during the 1st-level analysis of fMRIfloWS might differ slightly from the other approaches, as the fMRIfloWS routine removes non-steady state volumes during the preprocessing; (3) SPM used its own tissue probability map to create a binary mask restricted to gray matter voxels during the group analysis, while the other three toolboxes used the ICBM 2009c gray matter probability map instead.

To compare the unthresholded group statistic maps between the toolboxes, we created for each pairwise combination of preprocessing approach a Bland-Altman 2D histogram plot, as described by Bowring, Maumet, & Nichols (2018). These plots show the difference between the statistic value (y-axis), against the mean statistic value (x-axis) for all voxels within the intersection of the respective brain mask. In other words, it summarized in a 2D histogram plot, for each voxel how much higher the statistical value in toolbox B is (y-axis), in comparison to toolbox A's statistical value (x-axis).

The complete lists of parameters, the scripts to perform preprocessing, 1st- and 2nd-level analysis and the scripts to create individual figures can be found on fMRIfloWS GitHub page (<https://github.com/miykael/fmriflows/tree/master/paper>). Derivatives generated for the validation in phase 2 can be inspected and downloaded on NeuroVault (Gorgolewski et al., 2015) under the following links: (1) Standard deviation maps of temporal averages after preprocessing (<https://identifiers.org/neurovault.collection:5645>), (2) temporal SNR maps after preprocessing (<https://identifiers.org/neurovault.collection:5713>), (3) binarized 1st-level activation count maps

(<https://identifiers.org/neurovault.collection:5647>), (4) 2nd-level activation maps
 (<https://identifiers.org/neurovault.collection:5646>).

2.2 Results

2.2.1 Summary of outputs obtained by fMRIfloWS' processing pipelines

Output generated after executing the anatomical preprocessing pipeline

After the execution of the anatomical preprocessing pipeline, the following files are generated for each subject: (1) image of the inhomogeneity-corrected full head image, (2) image of the extracted brain, (3) binary mask used for the brain extraction, (4) individual tissue probability maps for gray matter (GM), white matter (WM), cerebrospinal fluid (CSF), skull and head, (5) normalized anatomical image in template space (6) reverse-normalized template image in subject space, (7) plus the corresponding transformation matrices used for output 5 and 6. Each anatomical preprocessing output folder also contains (8) the ICBM 2009c brain template used for the normalization, sampled to the requested voxel resolution.

In addition to these files, the following three informative figures are generated: (1) tissue segmentation, (2) brain extraction and (3) spatial normalization of the anatomical image. A shortened version of those three figures, as well as their explanation are shown in Figure 4.

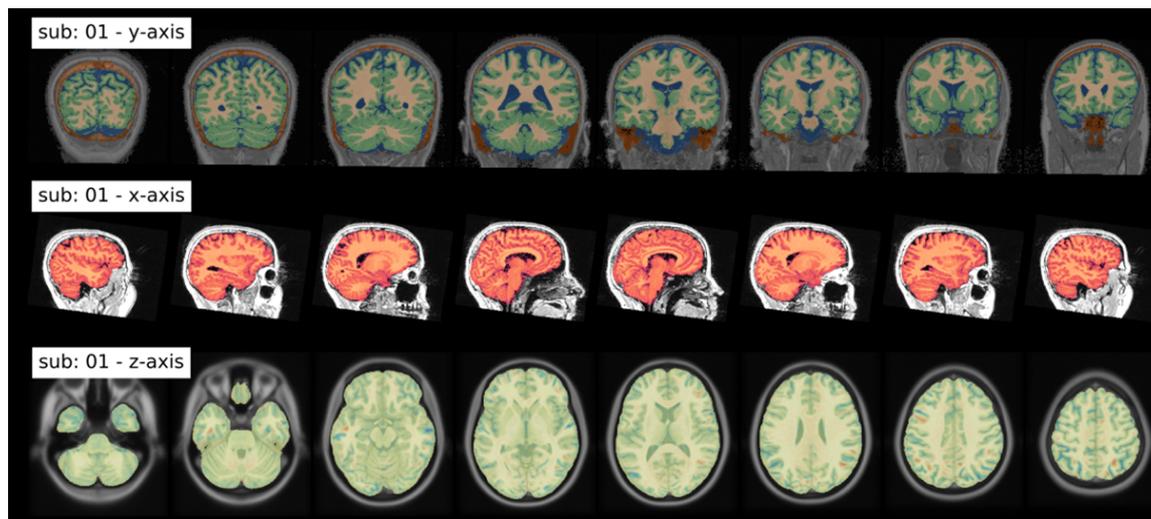


Figure 4: Summary of output figures generated by fMRIfloWS after executing the anatomical preprocessing pipeline. (Top) Coronal view of the image segmentation output, showing gray matter tissue in green, white matter tissue in beige, cerebrospinal fluid in blue. **(Middle)** Sagittal view of the brain extraction output, showing the extracted brain image in red, and the original anatomical image in gray. **(Bottom)** Axial view of the spatial normalization output, showing the normalized brain image highlighted in yellow, overlaid over the ICBM 2009c brain template in gray. Regions in red and blue show negative and positive deformation discrepancy between the normalized subject image and the template.

Output generated after executing the functional preprocessing pipeline

After the execution of the functional preprocessing pipeline, the following files are generated separately for each subject, each functional run and each temporal filtering: (1) text file indicating which volumes were detected as outliers, (2) tabular separated (TSV) file containing all extracted confound regressors, (3) text file containing the six motion parameter regressors according to FSL's output scheme, (4) binary masks for the brain, (5) masks for anatomical and functional component based noise correction, (6) functional mean image, and (7) completely preprocessed functional images, separated by spatial smoothing approaches. Each subject folder also contains (8) one text file per functional run indicating the number of non-steady-state volumes at the beginning of run.

The following is a more detailed description of the multiple confounds fMRIflows estimates during functional preprocessing:

Confounds based on motion parameters: In addition to the head motion parameters created during preprocessing, fMRIflows also computes (1) 24-parameter Volterra expansion of the motion parameters (Friston, Williams, Howard, Frackowiak, & Turner, 1996) using custom scripts and (2) Framewise Displacement (FD) component (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012) using Nipype.

Confounds based on global signal: Functional images before spatial smoothing were used to compute confound regressors, such as (1) DVARS, which represents the spatial standard deviation of the signal after temporal differencing, to identify motion-affected frames (Power et al., 2012), using Nipype and (2) four global signal curves representing the average signal in the total brain volume (TV), GM, WM and CSF, using Nilearn.

Detection of outlier volumes: The user can specify which of the six signal curves for FD, DVARS and average signal in TV, GM, WM and CSF to use to identify outlier volumes (see Figure 5A). Those are volumes that have larger fluctuations in the signal values in a given volume, compared to the z-scored standard deviation throughout the time course. The exact threshold for each curve can be adapted by the user, but its default value is set to a z-value of 3.27, representing 99%, for the FD, DVARS and TV signal. The identification number of each outlier volume is stored in a text file that might be used in the 1st-level pipeline during the GLM model estimation to remove the effect of those volumes from the overall analysis, also known as censoring (Caballero-Gaudes & Reynolds, 2016).

Confounds based on signal components: Using the temporal filtered functional images, two different kinds of approaches are performed to extract components that could be used for denoising or dimensionality reduction of the data. The first approach is called CompCor (Behzadi et al., 2007) and uses principal component analysis (PCA) to estimate the main sources of noise within specific confound regions. Regions are either defined by their temporal or anatomical characteristics. The temporal CompCor approach (tCompCor) considers the 2% most variable voxels within the confound brain mask as sources of confounds. The anatomical CompCor approach (aCompCor), considers voxels within twice eroded WM and CSF brain masks as sources of confounds. The user can specify how many aCompCor and tCompCor components should be computed, but the default value is set to five each. The second approach uses

independent component analysis (ICA) to perform source separation in the signal. Using Nilearn's CanICA routine, fMRIflows computes by default the top ten independent components throughout the confound masks. The number of confounds to extract can be adjusted by the user.

Storage of confound information: All of the confound curves computed after functional preprocessing are stored in a TSV file to allow for easy access.

Diverse set of overview figures: To allow for visual inspection of the numerous outputs generated after the execution of the functional preprocessing pipeline, fMRIflows creates many informative overview figures. These overviews cover the motion parameters used for head motion correction, the anatomical and temporal CompCor components, FD, DVARS, average signal in TV, GM, WM and CSF, and the ICA components. fMRIflows also creates a brain overview figure showing the extent of the different masks applied during functional preprocessing, a spatial correlation map between the ICA components and the individual voxel signal, and a carpet plot according to Power (2017) and Esteban et al. (2019). To better visualize underlying structures in the carpet plot the time series traces are sorted by their correlation coefficients to the average signal within a given region, allowing for a positive or negative time lag of 2 volumes. A shortened version of all these figures, as well as their explanations are shown in Figures 5-7.

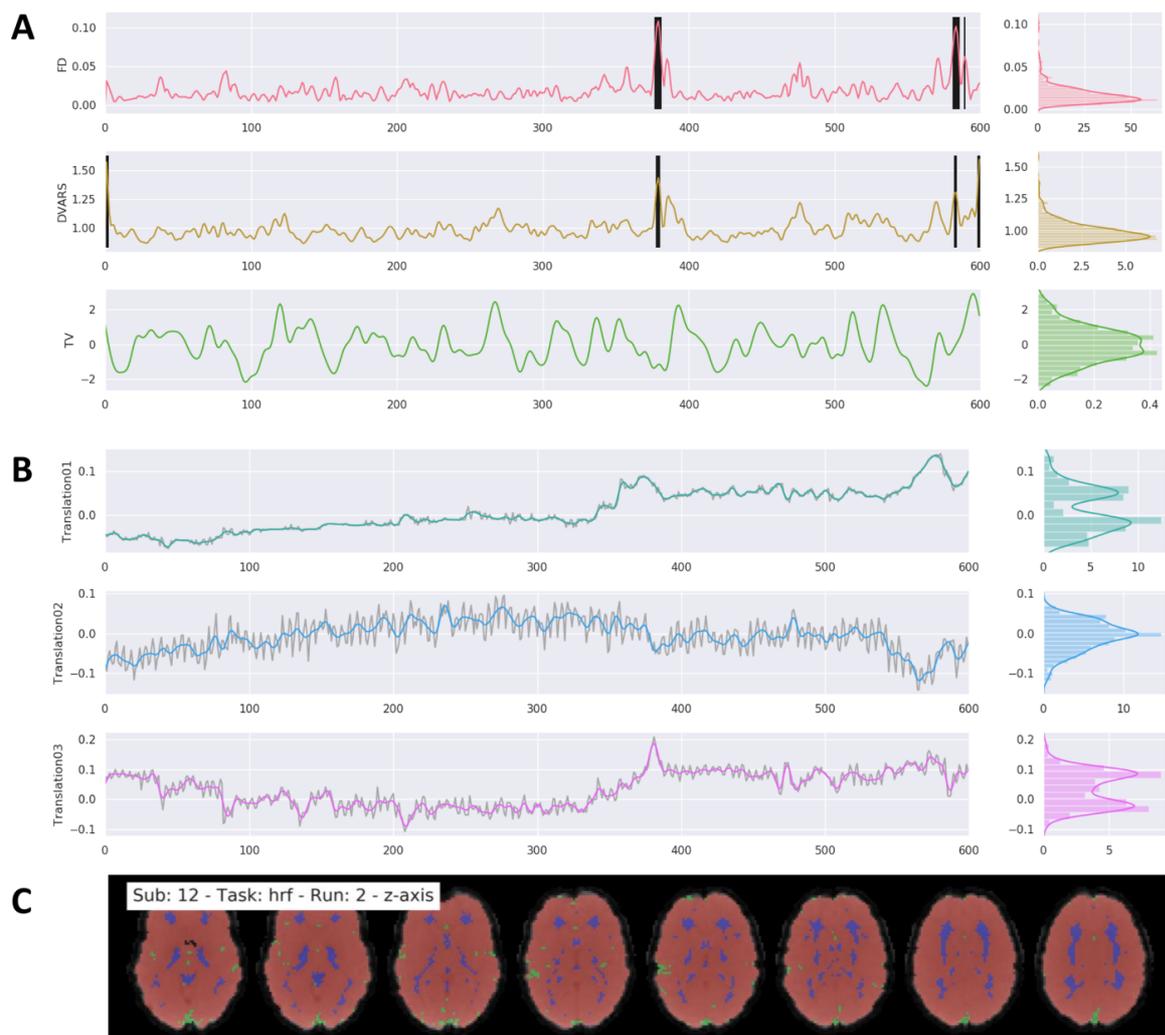


Figure 5: Example of general output figures generated by fMRIflows after executing the functional preprocessing pipeline. The dataset used to generate these figures was recorded with a TR of 600ms and had a total of 600 volumes per run. Preprocessing included a low-pass filter at 0.2 Hz. Distribution plots on the right side of the figures in part A and B represent value frequency in y-direction. **A)** Depiction of the nuisance confounds FD, DVARS and TV. Detected outlier volumes are highlighted with vertical black bars. **B)** Estimation of translation head motion after application of low-pass filtering at 0.2 Hz in color, and before temporal filtering in light gray. **C)** Depiction of brain masks used to compute DVARS (red), and temporal (green) and anatomical (blue) CompCor confounds, overlaid on the mean functional image (grey).

Output generated after executing the 1st-level analysis pipeline

After the execution of the 1st-level analysis, the following files are generated for the univariate analysis: (1) contrasts and statistical map of the specified contrasts, (2) SPM.mat file containing the information relevant for the model, (3) visualization of the design matrix used in the 1st-level model depicting the regressor for the stimuli, motion and confounds, and (4) glass brain plot for each estimated

contrast thresholded at the top 2% of positive and negative values created with AtlasReader (Notter et al., 2019) to provide a general overview of the quality of contrasts. The multivariate analysis part of this notebook creates: (1) one contrast image per condition and session which later can be used as samples for the multivariate analysis, and (2) a label file identifying the condition of each contrast.

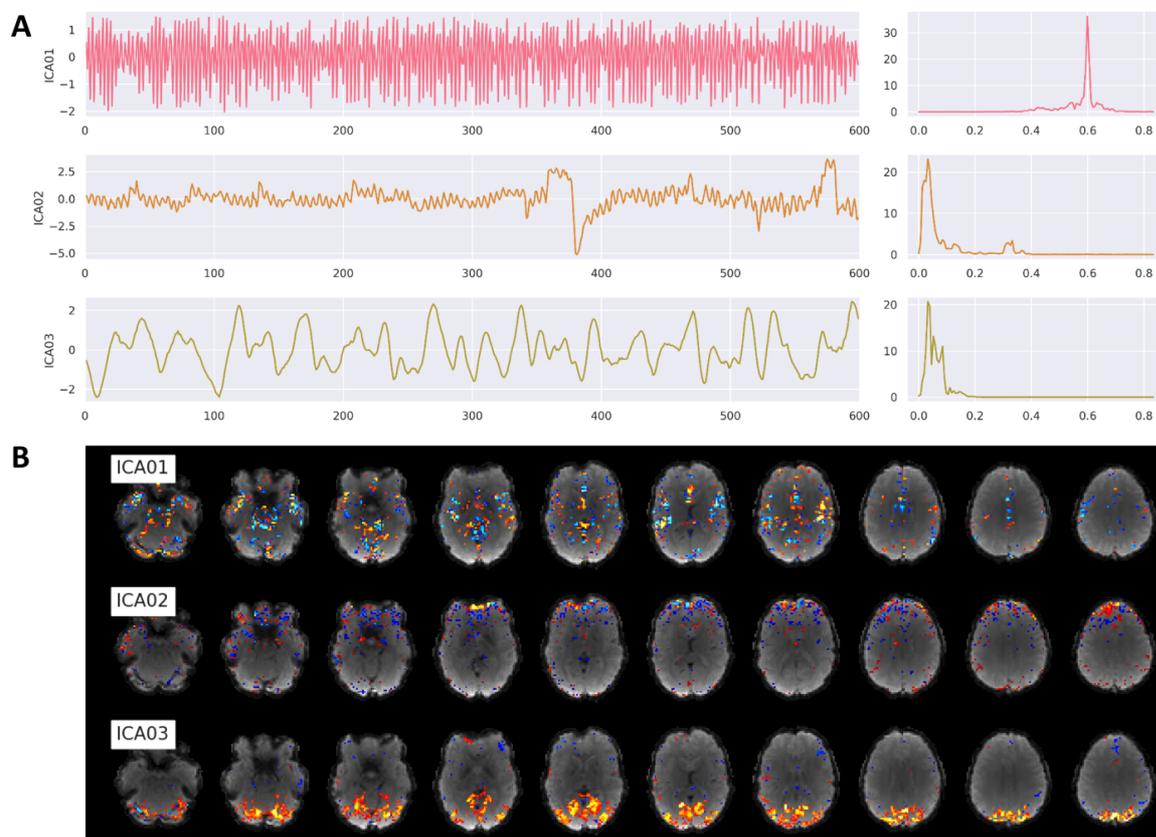


Figure 6: Example of ICA output figures generated by fMRIflores after executing the functional preprocessing pipeline. The dataset used to generate these figures was recorded with a TR of 600ms and had a total of 600 volumes per run. **A)** Correlation between the first three ICA components and the functional image over time (left) and the corresponding power density spectrum with frequency on the x-axis (right). First component most likely depicts respiration at 0.6 Hz, while third component is most likely visual activation induced by the visual stimulation task during data acquisition. **B)** Correlation strength between a given ICA component and the location in the brain volume for the first three ICA components.

Output generated after executing the 2nd-level analysis pipeline

After the execution of the 2nd-level univariate analysis, the following files are generated, individually for each contrast and spatial and temporal filter that was applied: (1) contrasts and statistical map of one-sample t -test contrast, (2) SPM.mat file containing the information relevant for the model, (3) thresholded statistical maps with corresponding AtlasReader outputs (i.e. glass brain plot to provide a

result overview, cross section plot showing each significant cluster individually, informative tables concerning the peak and cluster extent of each cluster).

After the execution of the 2nd-level multivariate analysis, the following files are generated, for each specified comparison individually: (1) subject-specific permutation files needed for correction according to Stelzer et al. (2013), (2) group-average prediction accuracy maps as well as corresponding feature-wise maps representing chance level acquired via bootstrapping approach (Stelzer et al., 2013), (3) group-average prediction accuracy maps after correction for multiple comparisons and (4) thresholded statistical result maps with corresponding AtlasReader outputs (i.e. glass brain plot to provide a result overview, cross section plot showing each significant cluster individually, informative tables concerning the peak and cluster extent of each cluster).

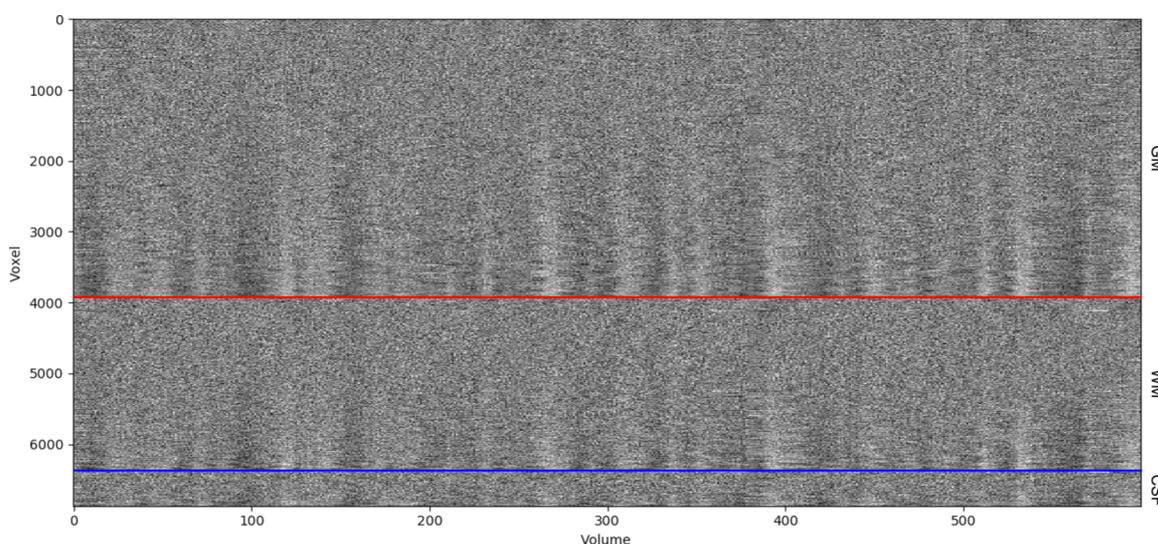


Figure 7: Example of a carpet plot figures generated by fMRIflows after executing the functional preprocessing pipeline. The dataset used to generate these figures was recorded with a TR of 600ms and had a total of 600 volumes per run. This panel shows the signal after preprocessing for every other voxel (y-axis), over time in volumes (x-axis). The panel shows voxels in the gray matter (top part), white matter (between blue and red line) and CSF (bottom section). The data was standardized to the average signal, and ordered within a given region according to the correlation coefficient between a voxel and to the average signal of this region.

2.2.2 Results of phase 1: Proficiency validation

Due to differences in scanner hardware, scan protocols, research requirements and expertise of the person who records the images, fMRI datasets can come in many different shapes and forms. We ran fMRIflows on several datasets to make sure that it is capable of dealing with differences inherent to each of them. In this section, we summarize the main issues we encountered during this process and describe how we tackled each of them.

Image orientation

fMRIflows reorients all anatomical and functional images at the beginning of the preprocessing pipeline to the neurological convention RAS (right, anterior, superior) to prevent failures of coregistration between anatomical and functional images due to orientation mismatches within subjects.

Image extent

Some datasets have unusually large image coverage along the inferior-superior axis, which means that their anatomical images also often contain part of the participant's neck. This can lead to unwanted outcomes in certain neuroimaging routines, as they were not tested for such additional tissue coverage. This is most pronounced in the case of FSL's BET routine, which has difficulty finding the center and extent of the brain, or SPM's segmentation routine that depends on the distribution of the voxel intensities within the whole volume. To prevent these and other unforeseen behaviors, fMRIflows uses FSL's robustfov routine to restrict all anatomical images to the same spatial extent.

Image inhomogeneity

Depending on the scan sequence protocol or the scanner hardware itself, some datasets can contain strong image intensity inhomogeneities, caused by an inhomogeneous bias field during data acquisition. This can have a negative effect on many different neuroimaging routines, most pronounced in brain extraction and image segmentation. To tackle this issue, fMRIflows uses ANTs' N4BiasFieldCorrection routine, which allows the analysis of datasets with even low image quality and strong image inhomogeneity. In the anatomical preprocessing pipeline, inhomogeneity correction is applied to improve the final output image. In the functional preprocessing pipeline, inhomogeneity correction is only applied to improve the estimation and extraction of different tissue types, but does not directly change the values in the final output image.

Brain extraction

Different brain extraction routines were explored to ensure: 1) that the extraction is sufficiently robust to handle different kinds of datasets, 2) that it is neither too conservative nor liberal with the removal of non-brain tissues, and 3) that it has an overall reasonably fast computation time. The best and most consistent results were achieved using SPM's image segmentation routine, followed by a specific thresholding and merging of the GM, WM and CSF probability maps. FSL's BET routine was not robust enough to lead to stable results on all tested datasets. While ANTs' Atropos routine led to comparably good results, we went with SPM because of the much faster computation time.

Image interpolation

For the single-shot spatial interpolation during normalization, we used ANTs and explored NearestNeighbor, BSpline and LanczosWindowedSinc (Lanczos, 1964) interpolation. NearestNeighbor

interpolation led to unnatural looking voxel-to-voxel value transitions. BSpline led in general to good results, but had issues especially with datasets that did not have full brain coverage and introduced some rippling low value fluctuations at the borders of non-zero voxels. LanczosWindowedSinc interpolation led to the best outcome by minimizing the smoothing effects and preventing the introduction of additional confounds reaffirming the observations from fMRIPrep (Esteban et al., 2019).

2.2.3 Results of phase 2: Performance validation

The performance validation of fMRIfloWS was conducted on three different task-based fMRI datasets, as described in Table 1. The preprocessing of fMRIfloWS was compared to other neuroimaging processing pipelines such as fMRIPrep, FSL and SPM. We tested fMRIfloWS' preprocessing pipeline with and without a temporal low-pass filter of 0.2 Hz to better understand performance differences between toolboxes and to stress the importance of adequate temporal filtering when processing fMRI datasets with high temporal resolution.

Estimated spatial smoothness after functional preprocessing

Each preprocessing step that resamples a functional image, such as slice time correction, motion correction, spatial or temporal interpolation has the potential to increase the spatial smoothness in the data. The less smoothness is introduced during preprocessing, the closer the data are to their initial version. We used AFNI's 3dFWHMx to estimate the average spatial smoothness (FWHM) of each functional image after preprocessing to compare the amount of data manipulation that was applied to the raw data (see Figure 8). As this FWHM value depends on the voxel resolution of a given dataset, we normalized it by the volume of the voxel to achieve a common FWHM value per 1mm^3 .

Overall, the estimated spatial smoothness after preprocessing with fMRIfloWS (without low-pass filter) is comparable to the one with fMRIPrep, while SPM's is in general significantly lower and FSL's is slightly higher. The differences with respect to SPM are probably due to the fewer numbers of resampling steps involved in SPM's preprocessing pipeline. The differences with respect to FSL are probably due to the interpolation method used during image resampling. While the FSL preprocessing pipeline uses the spline interpolation, fMRIfloWS and fMRIPrep use the LanczosWindowedSinc interpolation, which is known to minimize the smoothing during interpolation. The application of a temporal low-pass filter at 0.2 Hz during fMRIfloWS' preprocessing leads to a significantly higher spatial smoothness for the TR600 dataset when compared with the other approaches. This effect might also be present for the TR1000 dataset. However, there the difference between the fMRIfloWS preprocessing with and without low-pass filtering is not significant. This increased spatial smoothness for the approach that uses a low-pass filter makes sense, as the goal of the temporal low-pass filter itself is to smooth the time series values. This temporal smoothing forcibly also increases the spatial smoothness at each individual time point.

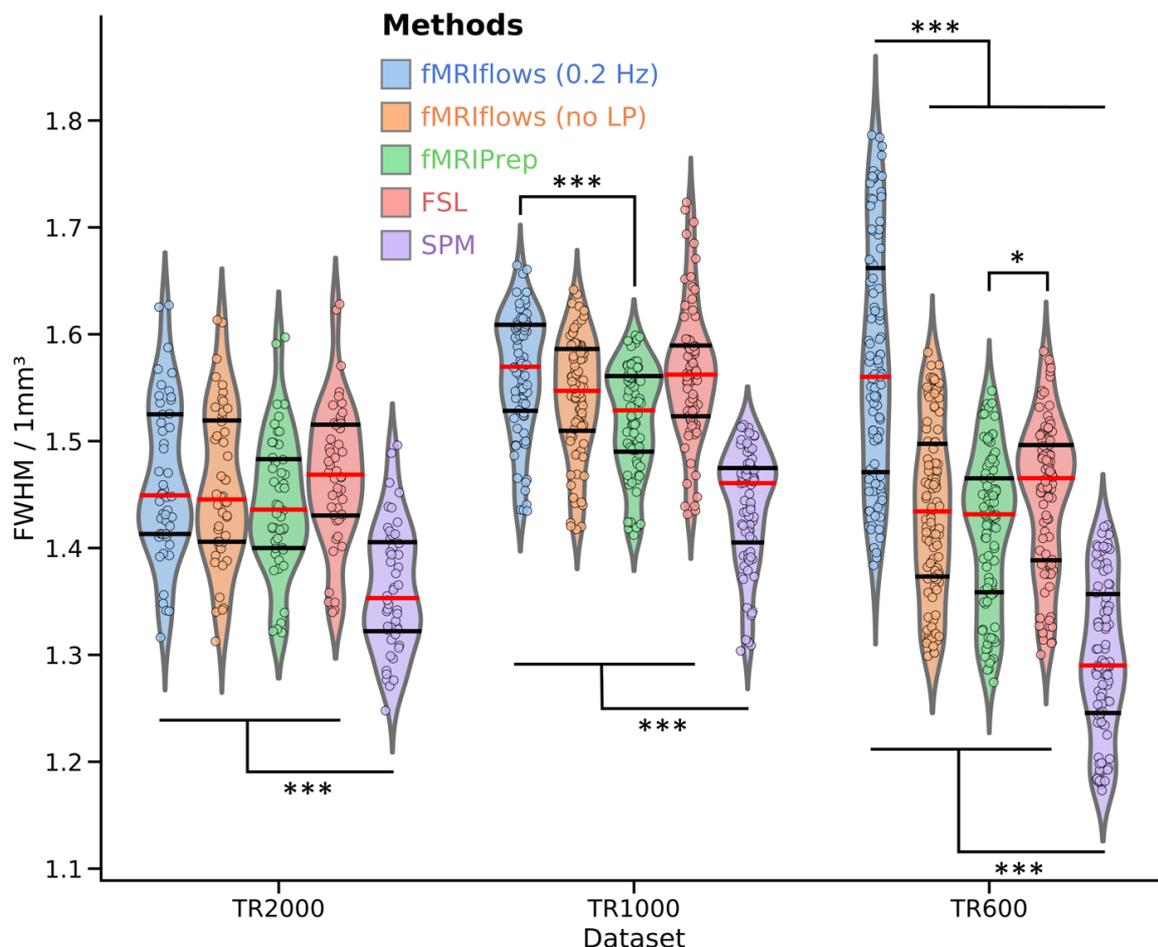


Figure 8: Investigation of estimated spatial smoothness after functional preprocessing of three different datasets, processed with varying approaches. The five different preprocessing approaches fMRIflows with (blue) and without (orange) a low-pass filter at 0.2 Hz, fMRIPrep (green), FSL (red) and SPM (violet) are plotted separately for the dataset TR2000 (left), TR1000 (middle) and TR600 (right). The violin plots indicate the overall distribution of the normalized smoothness estimates of each functional image (depicted in individual dots: TR2000=48 dots, TR1000=80 dots, TR600=102 dots). The red horizontal line represents the median value, while the horizontal black lines indicate the 25 and 75 percentiles of the value distribution respectively. Significant differences between groups are indicated with *: $p < 0.05$ and ***: $p < 0.001$.

Performance check of spatial normalization

We computed the standard deviation map for each population, based on the temporal average map of each preprocessed functional image, to compare the performance of spatial normalization of the different preprocessing methods on the three different datasets (see Figure 9).

The averaged standard deviation maps after fMRIflows’ and fMRIPrep’s preprocessing are very similar, which is not surprising as fMRIflows uses the same ANTs normalization routine with very similar parameters. The main difference lies in the fact that fMRIflows applies a brain extraction on the functional

images as well, which is not performed with fMRIPrep. Closer inspection reveals that the variability around the brain outline is slightly increased and more spread after normalization with fMRIflows or FSL, than with fMRIPrep. This effect was already observed in the original fMRIPrep paper, but seemed less pronounced in the current study. SPM performed comparably well, but a direct comparison was not possible as the spatial normalization with SPM is performed using SPM’s own tissue probability map template, while the other three methods used the normalization to the ICBM’s 2009c brain template. No clear performance differences have been observed between the three datasets.

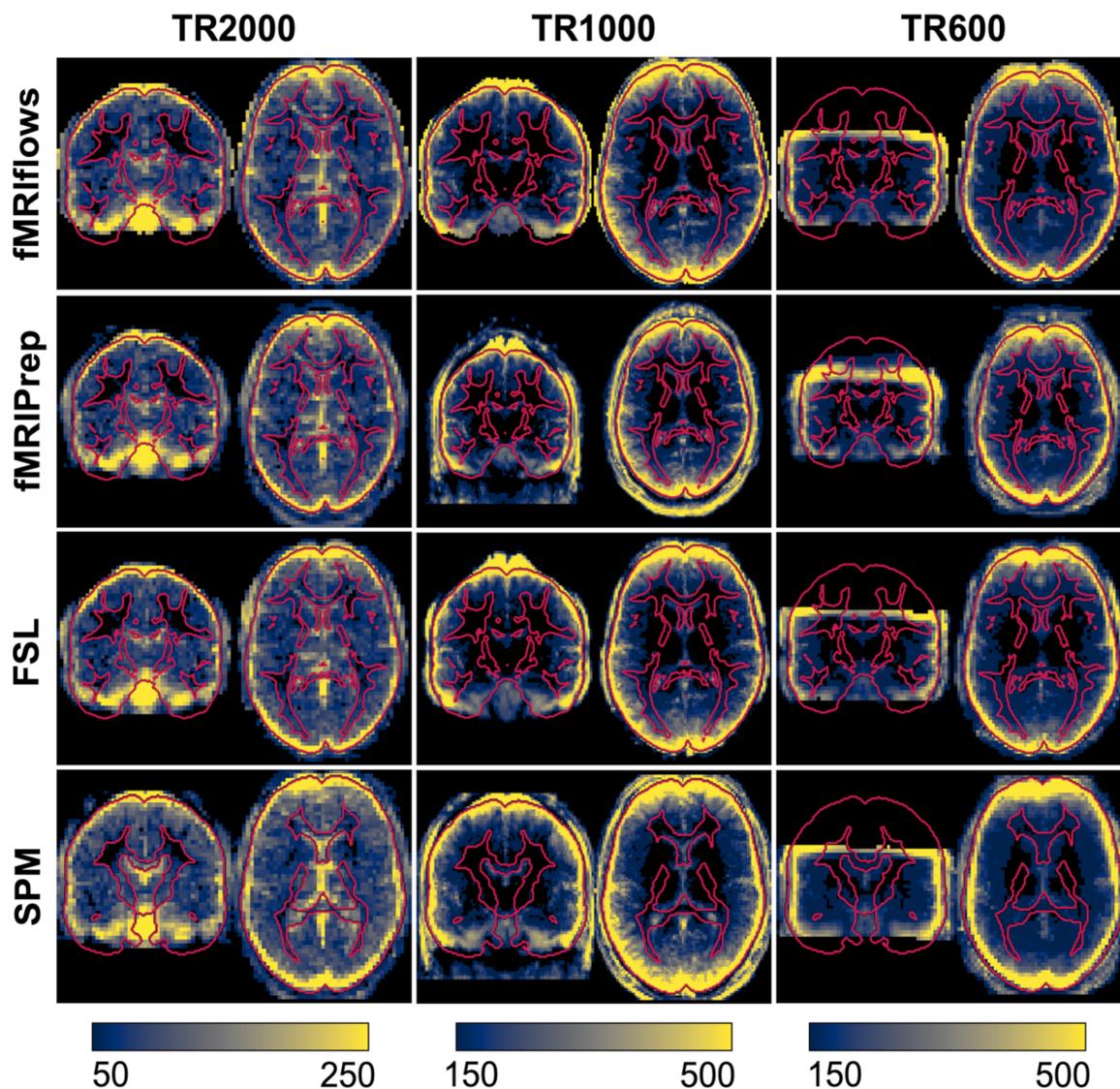


Figure 9: Depiction of standard deviation maps of the temporal averages of three different datasets, after multiple functional preprocessing approaches. Preprocessing was done with fMRIflows (with a temporal low-pass filter at 0.2 Hz; without low-pass filter looks identical), fMRIPrep, FSL and SPM (from top to bottom) separated for the TR2000 (left), TR1000 (middle) and TR600 (right) dataset. Color value represents the standard deviation value over all subjects. Color scale is the same within a dataset and was

set manually to highlight the border effects in gray matter regions. Regions with high inter-subject variability are shown in yellow, while regions with low inter-subject variability are shown in blue. Outline of the brain and subcortical white matter regions is delineated in red and is based on the ICBM 2009c brain template, except for the analysis with SPM where it is based on SPM's tissue probability map template.

Temporal signal-to-noise ratio (tSNR) after preprocessing

We computed the voxel-wise temporal SNR according to Smith et al. (2013) to assess the amount of informative signal contained in the data after preprocessing. This measurement serves as a rough estimate to compare different preprocessing methods, but did not allow a direct comparison between datasets, as the tSNR value is a relative measurement that depends highly on the paradigm presented, the initial spatial and temporal resolution of the functional images, as well as the MRI scan sequence specific parameters such as acceleration factors (Smith et al., 2013). Using Nipype's TSNR routine, we first removed 2nd-degree polynomial drifts in each functional image, and estimated tSNR maps by computing each voxel's temporal mean, dividing it by its temporal standard deviation, and multiplying it by the square root of the number of time points recorded in a given run. By averaging the tSNR maps over the population, we get a general tSNR map per preprocessing method for each dataset (see Figure 10).

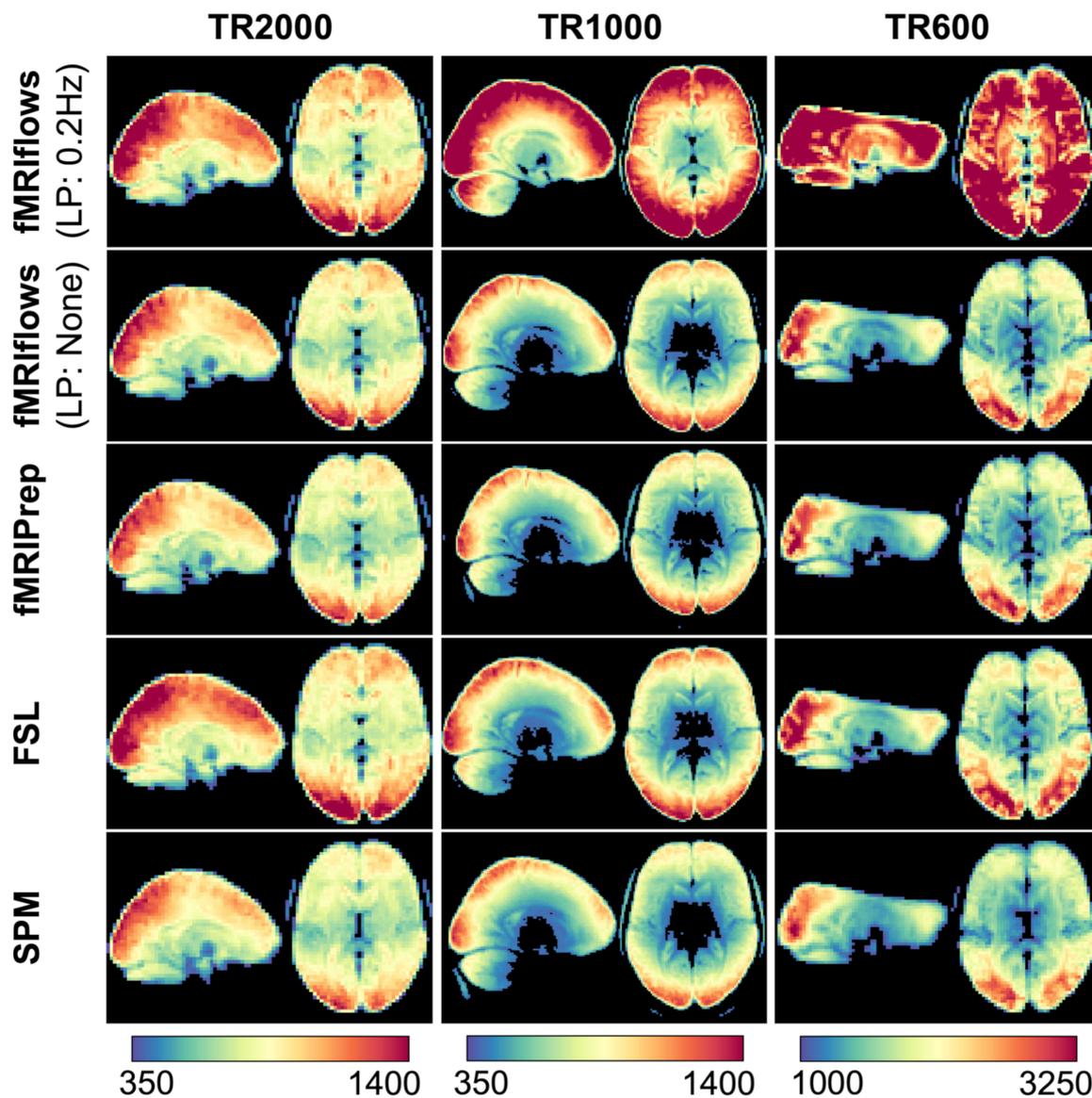


Figure 10: Depiction of temporal signal-to-noise ratio maps of three different datasets, after multiple functional preprocessing approaches. Preprocessing was done with fMRIflows (with and without a temporal low-pass filter at 0.2 Hz), fMRIPrep, FSL and SPM (from top to bottom) separated for the TR2000 (left), TR1000 (middle) and TR600 (right) dataset. Color value represents the tSNR value as computed with the Nipype routine TSNR. Color scale was set manually and differs between datasets, but is held constant between different preprocessing methods.

In general, preprocessing with fMRIflows without temporal low-pass filter led to similar average tSNR maps as preprocessing with fMRIPrep. Overall, preprocessing with FSL led to slightly increased average tSNR values, while preprocessing with SPM led to slightly decreased average tSNR maps. The additional application of a low-pass filter at 0.2 Hz in all three datasets led to increased tSNR values after preprocessing with fMRIflows. This effect was more pronounced for higher temporal resolution (as in

Dataset TR1000 and TR600). The color scales in Figure 10 were set manually so that the fMRIflows (without low-pass filter) approach shows comparable intensities for the three datasets.

Performance check after 1st-level analysis

To investigate the effect of the different preprocessing methods on the 1st-level analysis, we carried out within-subject statistical analysis using Nistats. The activation maps were estimated using a general linear model (GLM). The GLM included a constant term, the stimuli regressors convolved with a double-gamma canonical hemodynamic response function, six motion parameters (three translation and three rotation), and a high pass filter at 100Hz, represented by a set of cosine functions, and no temporal derivatives. The input data were smoothed using a kernel with a FWHM of 6mm, using a Nilearn routine. The analysis pipelines between the preprocessing methods and datasets were kept as identical as possible, and differed only in the number of time points contained in the dataset and the estimated motion parameters. The statistical map for each participant was binarized at $z = 3.09$, which corresponds to a one-sided test value of $p < 0.001$. The population average of these maps is shown in Figure 11.

The results show that the thresholded activation count maps between the fMRIflows approach without a low-pass filter, fMRIPrep, FSL and SPM do not differ too much between each other, for all three datasets. In contrast to the other preprocessing methods, however, the preprocessing with fMRIflows with a low-pass filter at 0.2 Hz drastically increased the size and fraction value of the thresholded activation count maps, for the datasets TR1000 and TR600. Thus, appropriate temporal filtering increased the statistics for datasets with higher temporal resolution remarkably. For a more detailed comparison between all the toolboxes, see Supplementary Note 7.

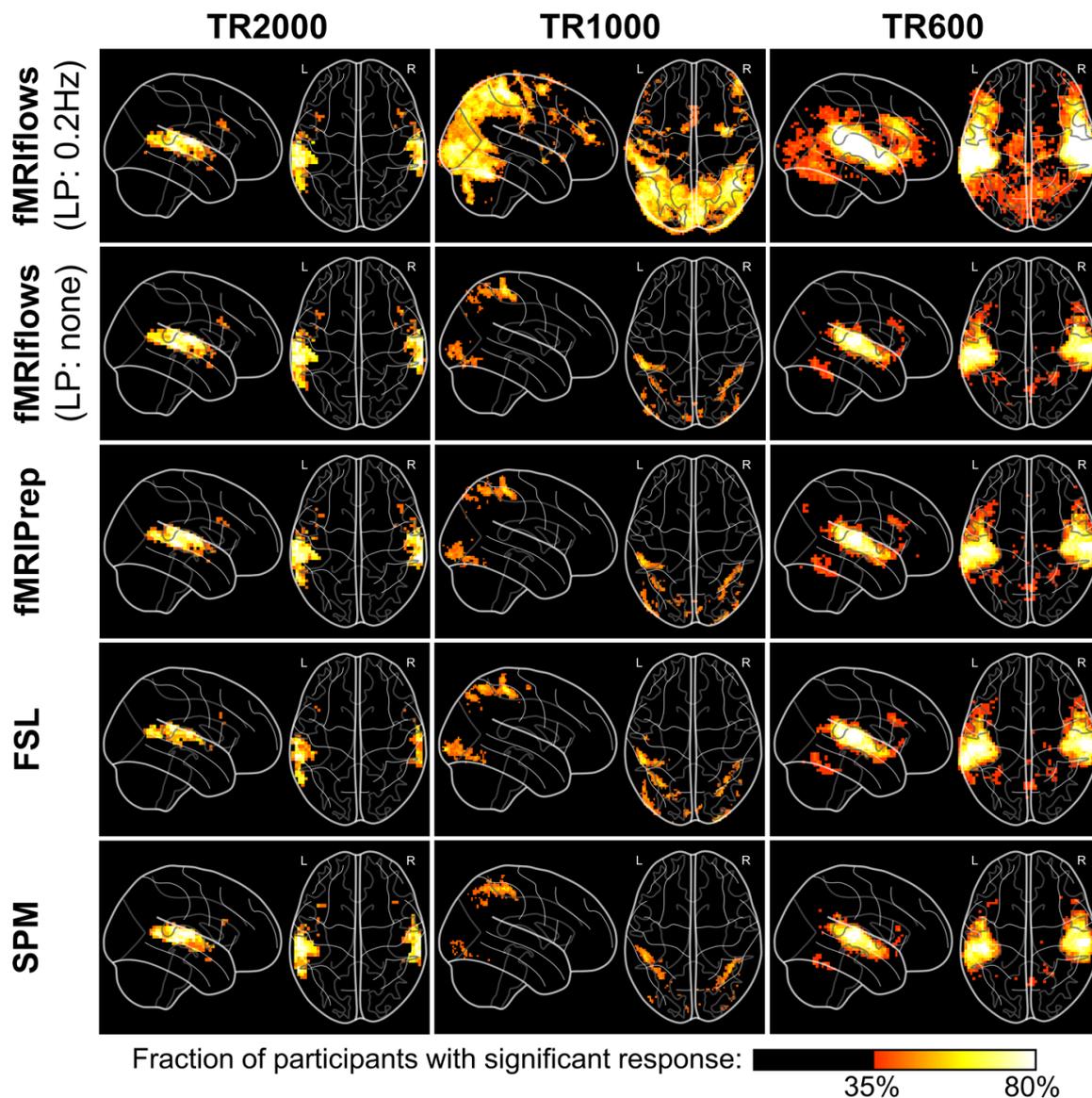


Figure 11: Depiction of binarized 1st-level activation count maps, thresholded at $p < 0.001$, after multiple functional preprocessing approaches. Preprocessing was done with fMRIflows (with and without a temporal low-pass filter at 0.2 Hz), fMRIPrep, FSL and SPM (from top to bottom) separated for the TR2000 (left), TR1000 (middle) and TR600 (right) dataset. Activation count maps were normalized to the ICBM 2009c brain template. Color code represents the fraction of participants that show significant activation above a p -value threshold at 0.001 and corrected for false positive rate (FPR).

Performance check after 2nd-level analysis

To investigate the effect of the different preprocessing methods on the 2nd-level analysis, we carried out between-subject statistical analysis using Nistats and computed one-sample t-test for each preprocessing method and dataset. The unthresholded group-level T-statistic maps of each analysis was

then compared to each other on a voxel-by-voxel level using Bland-Altman 2D histograms (Bowring et al., 2018), see Figure 12.

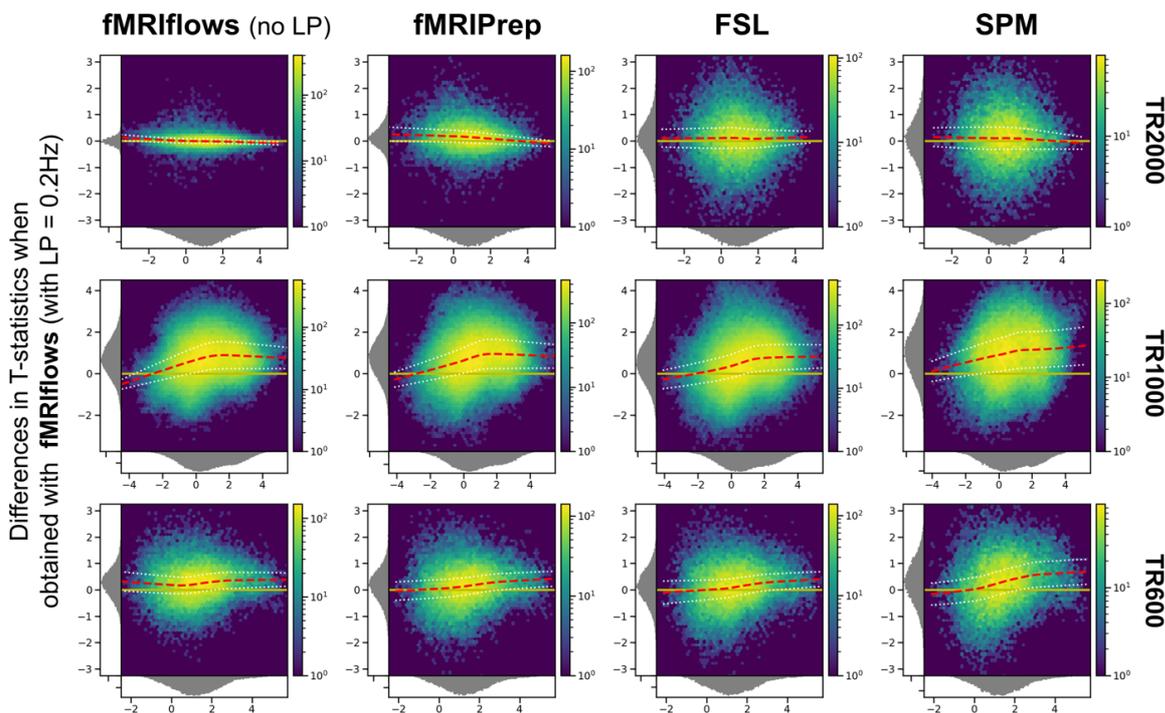


Figure 12: Bland-Altman 2D histograms of three different datasets, comparing unthresholded group-level T-statistic maps between multiple processing approaches. Datasets TR2000 (top), TR1000 (middle) and TR600 (bottom) were used for the comparison. Density plots show the relationship between average T-statistic value (horizontal) and difference of T-statistic values (vertical) at corresponding voxels for different pairwise combinations of toolboxes. The difference of T-statistics was always computed in contrast to a preprocessing with fMRIflows using a low-pass filter at 0.2 Hz, while the average T-statistics in horizontal direction investigated the preprocessing with (from left to right) fMRIflows without a low-pass filter, fMRIPrep, FSL and SPM. Distribution plots next to x- and y- axis depict occurrence of a given value in this domain. Color code within the figure indicates the number of voxels at this given overlap, from a few (blue) to many (yellow). Yellow horizontal line at zero indicates no value differences between corresponding voxels. Red dashed line depicts horizontal density average.

The results shown in Figure 12 indicate no pronounced differences between the preprocessing with fMRIflows with a low-pass filter at 0.2 Hz and the other four approaches for the analysis of the TR2000 dataset. An increased variability in the y-direction indicated a decrease in voxel-to-voxel correspondence, which might be explained by different spatial normalization implementations. The fact that the average horizontal density value (red dashed line) is close to the zero line (in yellow) indicated that the different preprocessing methods led to comparable group-level results with the TR2000 dataset. The Bland-Altman plots for the TR1000 and TR600 datasets showed a clear increase of t-statistic when the preprocessing was

done with fMRIflows with a low-pass filter at 0.2 Hz, compared to any other method. This effect was stronger for higher t-values. For a more detailed comparison between all the toolboxes, see Supplementary Note 8.

2.3 Discussion

fMRIflows is a fully automatic fMRI analysis pipeline, which can perform state-of-the-art preprocessing, 1st-level and 2nd-level univariate analyses, and multivariate analysis. The goal of such an autonomous approach is to improve objectiveness of the analysis, maximize transparency, facilitate ease of use, and provide newest analysis approaches to every researcher, including users outside the field of neuroimaging. While the predefined analysis pipelines help to reduce the number of error-prone manual interventions to a minimum, it also has the advantage of decreasing the number of analytical degrees of freedom available to a user to its minimum (Carp, 2012). This constraint in flexibility is important as it helps to control the variability in data processing and analysis (Botvinik-Nezer et al., 2020). fMRIPrep showed a clear need for such analysis-agnostic approaches and was therefore chosen to provide much of the groundwork for fMRIflows.

In comparison with other neuroimaging software like fMRIPrep, FSL and SPM, fMRIflows achieved comparable or improved results in (1) SNR after preprocessing, (2) within-subject t-statistics, and (3) between-subject t-statistics. These results were more obvious in the context of datasets that had a temporal resolution equal to or below 1000ms, and if a low-pass filter at 0.2 Hz was applied. The latter might also influence the overall outcome of the other software packages, but so far fMRIflows is the only such neuroimaging software that performs orthogonal filtering between the motion correction and temporal filtering, as proposed by Lindquist et al. (2019).

Being an open-source project, shared via GitHub, facilitates the transparency in the development of fMRIflows. Users will be able to inspect the complete history of the changes and have access to all discussion connected to the software. Code adaptations and additional support to new usage will be proposed by the user community, which will make the adaptation to newest standards easy and straightforward. In addition to the version-controlled system used on GitHub, a continuous integration scheme with CircleCi will ensure continuous functionality.

fMRIflows also improved the overall computation time needed to perform preprocessing and 1st and 2nd-level analysis. Indeed, Nipype provides a parallel execution feature of processing pipelines, which is not yet possible with FSL or SPM. fMRIPrep uses the same boost of parallelism, but is overall much slower if the default execution of FreeSurfer's recon-all routine is performed. However, fMRIflows does not yet support parallel computation via a job scheduler on a computation cluster, which is currently possible with fMRIPrep.

The inclusion of many informative visual reports allows a direct quality control and verification of the performed processing steps, as fMRIflows' outputs provide a general quality assessment even though it is not as detailed and rigorous as MRIQC (Esteban et al., 2017). In contrast to other software packages, fMRIflows uses an adapted visualization of the carpet plot proposed by Power (2017) to highlight underlying temporal structure and voxel-to-voxel correlations within different brain tissue regions and/or throughout the brain. Such approaches help to observe general signal trends and sudden abrupt signal

changes throughout the brain, but the exact implications of these modified carpet plots need to be further investigated.

Results of fMRIflows' validation phase 1 suggests that the software is capable of analyzing different types of datasets, independently of the extent of head coverage, original image orientation, spatial or temporal resolution. By increasing the user base and testing fMRIflows on many more datasets, new adaptations might be required and hidden bugs could emerge. Users can observe any changes done to the software in the future directly on GitHub and are encouraged to state any questions or comments in connection with the software on the community driven neuroinformatics forum NeuroStars (<https://neurostars.org>).

Further development of the software will involve (1) moving away from an SPM dependency for the 1st and 2nd-level modeling, (2) using the more flexible FitLins toolbox (<https://github.com/poldracklab/fitlins>), and (3) implementing an fMRIflows BIDS-App to further improve the toolbox's accessibility.

Acknowledgments

We thank the creator of fMRIPrep for providing an excellent starting point and inspiration in the development of fMRIflows. We also thank the developers of PyMVPA, Nilearn, Nibabel and Nistats for bringing the neuroimaging domain into the Python universe, and the developer of Nipype and BIDS for creating a clear framework to execute processing pipelines, as well as the whole neuroimaging open source and science community with its numerous contributors. The authors have no conflicts of interest to disclose. CRediT author statement: M.P.N.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing, Visualization, Project administration; P.H.: Conceptualization, Methodology, Software, Validation, Writing; S.D.C.: Methodology, Supervision, Writing - Review & Editing, Validation; O.F.G.: Writing- Reviewing and Editing, Validation; A.I.I.: Writing- Reviewing and Editing, Validation; M.M.M.: Supervision, Writing- Reviewing and Editing, Funding acquisition

Funding

This work was supported by the Swiss National Science Foundation (grant numbers 320030-169206 to M.M.M.) and research funds from the radiology service at the University Hospital in Lausanne (CHUV). This work was supported by the Centre d'Imagerie BioMédicale (CIBM) of the University of Lausanne (UNIL) and the Swiss Federal Institute of Technology Lausanne (EPFL). P.H. was supported in parts by funding from the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives initiative, the National Institutes of Health (NIH) NIH-NIBIB P41 EB019936 (ReproNim), the National Institute Of Mental Health of the NIH under Award Number R01MH096906, as well as by a research scholar award from Brain Canada, in partnership with Health Canada, for the Canadian Open Neuroscience Platform initiative.

2.4 Supplementary Material

Supplementary Note 1: Content examples of parameter specification files

Content example of file fmriflows_spec_preproc.json

```
{
  "subject_list_anat": ["01", "02", "03", "04", "05"],
  "session_list_anat": [],
  "T1w_id": "T1w",
  "res_norm": [1.0, 1.0, 1.0],
  "norm_accuracy": "precise",
  "subject_list_func": ["01", "02", "03", "04", "05"],
  "session_list_func": [],
  "task_list": ["multi"],
  "run_list": [1, 2, 3],
  "ref_timepoint": 500,
  "res_func": 2.0,
  "filters_spatial": [{"LP", 6.0}],
  "filters_temporal": [[null, 100.0 ], [5.0, 100.0]],
  "n_compcor_confounds": 5,
  "outlier_thresholds": [3.27, 3.27, 3.27, null, null, null],
  "n_independent_components": 10,
  "n_parallel_jobs": 7
}
```

Content example of file fmrfiflows_spec_analysis.json

```
{
  "tasks": {
    "multi": {
      "condition_names": ["cond_01", "cond_02", "cond_03", "cond_04"],
      "contrasts": [
        ["cond_01", [1.0, 0.0, 0.0, 0.0], "T"],
        ["cond_02", [0.0, 1.0, 0.0, 0.0], "T"],
        ["cond_03", [0.0, 0.0, 1.0, 0.0], "T"],
        ["cond_04", [0.0, 0.0, 0.0, 1.0], "T"],
        ["cond_01 > cond_02", [1.0, -1.0, 0.0, 0.0], "T"]
      ]
    }
  },
  "subject_list": ["01", "02", "03", "04", "05"],
  "session_list": [],
  "filters_spatial": [{"LP", 6.0}],
  "filters_temporal": [{"null", 100.0}, [5.0, 100.0]],
  "nuisance_regressors": ["Rotation", "Translation", "FD", "DVARs", "TV"],
  "use_outliers": true,
  "model_serial_correlations": "AR(1)",
  "model_bases": {"hrf": {"derivs": [0, 0]}},
  "estimation_method": {"Classical": 1},
  "normalize": true,
  "norm_res": [1, 1, 1],
  "con_per_run": true,
  "norm_res_multi": [3.0, 3.0, 3.0],
  "analysis_postfix": "",
  "gm_mask_thr": 0.1,
  "height_threshold": 0.001,
  "use_fwe_correction": false,
  "extent_threshold": 5,
  "use_topo_fdr": true,
  "extent_fdr_p_threshold": 0.05,
  "atlasreader_names": "default",
  "atlasreader_prob_thresh": 5,
  "n_parallel_jobs": 7
}
```

Content example of file fmriflows_spec_multivariate.json

```
{
  "subject_list": [ "01", "02", "03", "04", "05"],
  "session_list": [],
  "filters_spatial": [{"LP", 6.0}],
  "filters_temporal": [null, 100.0], [5.0, 100.0]],
  "multivariate_postfix": "",
  "clf_names": ["LinearNuSVMC"],
  "sphere_radius": 3,
  "sphere_steps": 3,
  "n_chunks": 6,
  "tasks": {
    "hrf": [
      [{"cond_01", "cond_02"}, {"cond_01", "cond_02"}],
      [{"cond_01", "cond_02"}, {"cond_03", "cond_04"}]
    ]
  },
  "n_perm": 100,
  "n_bootstrap": 100000,
  "block_size": 1000,
  "threshold": 0.001,
  "multicomp_correction": "fdr_bh",
  "fwe_rate": 0.05,
  "atlasreader_names": "default",
  "atlasreader_prob_thresh": 5,
  "n_parallel_jobs": 7
}
```

Supplementary Note 2: Description of anatomical preprocessing pipeline steps

Image reorientation: To make sure that all images throughout the processing have the same orientation, images are first reoriented with Nipype according to the neurological convention RAS (right, anterior, superior).

Image cropping: To make sure that the focus in the anatomical image is on the brain, we use FSL's `robustfov` function to remove irrelevant portions of the neck. This is particularly relevant for the later brain extraction step, and helps to ensure that the segmentation algorithm focuses on the brain and not on additional body sections.

Image inhomogeneity correction: To correct for intensity non-uniformities caused by the inhomogeneity of the bias field during data acquisition, we use ANTs' `N4BiasFieldCorrection` algorithm. This step improves the quality of the following image segmentation and is crucial for anatomical images of lower image quality, as they would otherwise fail during the image segmentation.

Anatomy segmentation: The image segmentation uses SPM12's standard image segmentation and provides probability maps for five tissue segments: gray matter (GM), white matter (WM), cerebrospinal fluid (CSF), skull and head.

Brain extraction: The GM, WM and CSF probability maps are combined using Nilearn to create a binary mask which is used to extract the brain. We chose this approach over others as it proved to be more robust and provided the best balance between restriction and inclusion than other algorithms, especially in the context of low image quality.

Spatial normalization: As a final step, the extracted brain image is spatially normalized to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2011) with ANTs' `antsRegistration` algorithm, using nonlinear image registration with a b-spline interpolation.

Supplementary Note 3: Description of functional preprocessing pipeline steps

Image reorientation: As a first preprocessing step, functional images are reoriented to the neurological convention RAS, using Nipype, to make sure that they have the same orientation as the anatomically preprocessed images.

Non-steady-state detection: Afterwards, the first few volumes of each functional run are investigated for non-steady-state volumes using Nipype. If non-steady-state volumes are detected, they are removed before the motion correction is applied.

Creation of brain masks: To remove unwanted tissue from functional images a three-step approach was chosen. First, the mean image of the functional run is corrected for intensity inhomogeneity using ANTs' N4BiasFieldCorrection algorithm. Second, FSL's BET algorithm is applied to create a binary brain mask. Third, the binary brain mask is dilated by two iterations and holes are filled. This procedure has proven to be optimal in removing non-brain tissue in almost all cases that it was tested on, while ensuring that all brain tissue types are included within the mask. This brain mask is then applied to the functional images before the motion correction step, to make sure that the estimation of the motion parameter is in relation to the brain and not the whole head. The mask is additionally used to restrict the coregistration of the functional images to the anatomical images to the brain tissue and not to the whole head. However, this binary mask is not used to mask the functional images at any time.

Slice-timing correction: If specified, slice time correction (Sladky et al., 2011) is performed on the reoriented functional images using SPM, according to the slice onset parameters specified in the BIDS information file and the reference time point mentioned in the fMRIflows JSON specification file. Slice-time correction is applied after the estimation of the head motion as recommended by (Power, Plitt, Kundu, Bandettini, & Martin, 2017).

Head motion estimation: Estimation of the motion parameters is performed using FSL's MCFLIRT algorithm, on the reoriented functional images, after non-steady-states volumes are removed and the brain is masked with the binary mask computed during the previous step. If the user specified a low-pass filter, an additional step is included in the estimation of the motion parameters. This step takes the estimated motion parameters (three rotation and three translation) and applies a Butterworth (Stephen Butterworth, 1930) low-pass filter to each of the six components individually. This step is crucial to guarantee that the motion correction and the temporal filter are orthogonal to each other. Otherwise, previously filtered confounds might be reintroduced with a later step (Lindquist et al., 2019). We are using custom-written Python code to perform this step, using routines from Nilearn, output files from FSL's MCFLIRT routine and FSL's avscale routine. To our knowledge, our implementation is the first openly available routine that transforms the affine motion parameters for each volume (here applying a low-pass filter) before they are applied to the functional images.

Intra-subject registration: The coregistration of the functional image to the anatomical image is based on FSL's FEAT pipeline and fMRIPrep and uses a two-step co-registration. Both steps use FSL's FLIRT algorithm. The first step uses the anatomical image to pre-align the mean image from the estimation

of the motion parameters, followed by the second coregistration step where the white matter probabilistic image computed with SPM's segmentation routine is used together with the anatomical image in a boundary-based registration (BBR) approach. The first step uses six degrees of freedom (three rotations and three translations), while the second step uses nine degrees of freedom, adding three scaling degrees. The addition of these scaling degrees was copied from fMRIPrep's approach (Esteban et al., 2019) and allows the image to be stretched in the direction of the recording. For example, functional images that are recorded in the A-P axis are often squeezed a bit in this direction. Keep in mind that this scaling/stretching is only used for an optimal image coregistration, so that the functional images overlay optimally with the white matter boundaries.

Spatial interpolation: Once we have the slice-time corrected functional images, computed the correct motion parameters that we want to apply, have the coregistration matrix between the functional and anatomical images, and have the transformation matrix to normalize the anatomical image onto the ICBM 2009c template brain, we can go to the next step and apply all spatial interpolations in a single-shot. The spatial normalization is optional and can also be applied in the later 1st-level notebook. But we recommend doing all of those spatial transformations at once to keep the number of spatial interpolations as low as possible. First we use C3 to combine all the different transformation matrices and apply them to each volume individually with ANTs' ApplyTransforms routine, using the LanczosWindowedSinc interpolation (Lanczos, 1964). We chose this interpolation over others, as it minimizes the smoothing effect of the interpolation and creates the least artifacts outside the brain volume.

Creation and application of warped brain masks: Once the functional images are spatially transformed either into each subject's structural space using a user-defined voxel resolution, or directly normalized into template space, also using a user-specified voxel resolution, two masks are created to remove irrelevant voxels outside of the brain, such as skull, eyes and head, as well as voxels that do not have values in most volumes. Those voxels are mostly seen in functional images that only cover part of the brain (slabs) and are introduced through motion at top or bottom slices of the slab. The first mask is later used to mask the functional image before the application of the temporal filters, while the second mask (confound brain mask) is applied before the extraction of confound signals. For this reason, the first mask should be sensitive enough to keep all voxels within the brain, while the second mask should be specific enough to only keep brain voxels, so that extracted confound curves are only based on nuisance sources within the brain, i.e. the region that we want to clean. The starting point of both masks is again the binary brain mask. For this we first compute the functional mean image using Nilearn, second correct for intensity non-uniformities using ANTs N4BiasFieldCorrection routine and third apply FSL's BET routine to create a binary brain mask. This mask is then dilated by two iteration steps, holes are filled up and the mask is then applied to the spatially transformed functional image. The steps to create the first brain mask are as follows: the initial brain mask is first dilated by two voxels, holes are filled up and any voxels that have no signal in more than 1% of volumes are removed. The second brain mask is created as follows: the

initial brain mask is first dilated by one voxel, holes are filled, the binary image is again eroded by two iteration steps and afterwards, voxels with no signal in more than 5% of volumes are removed.

Temporal filtering: The next preprocessing step applies temporal filters if they were specified. The user can decide if they want to apply low-pass, high-pass or band-pass filters, or none. The temporal filtering is performed with AFNI's 3dBandpass routine. In this step we also apply the first mask from the previous step to remove voxels that are clearly non-brain tissues. After temporal filtering of the data, the image intensity is normalized in such a way that the white matter distribution peak throughout the functional image is at a value of 10'000.

Spatial filtering: The final preprocessing step applies a spatial filter to the functional images, if the user specified this. Contrary to other toolboxes, our software allows the application of spatial low-pass, high-pass and band-pass filters. The last one can be especially useful for the preprocessing of data that is later used in a multivariate analysis. Sengupta, Pollmann, & Hanke (2018) have shown that the correct application of a band-pass filter can drastically improve the prediction accuracy in a multivariate approach. Independently of whether or not a spatial filter was applied, functional images are checked for absolute values above 30'000. If this is the case, images are rescaled to have a maximum absolute value at 30'000. After this, images are stored in integer16 data format to reduce their footprint on the database.

Supplementary Note 4: Description of 1st-level analysis pipeline steps

Data collection: Preprocessed functional images and model relevant parameters are collected and prepared for the 1st-level analysis.

Model Specification and estimation: Model relevant parameters are combined to create the design matrix needed for the general linear model (GLM) analysis.

Univariate contrast estimation: 1st-level contrasts are computed for each subject individually, according to the input parameters specified in the fMRIflows JSON parameter file.

Optional multivariate contrast estimation: To create multiple beta contrasts that then can be used for multivariate analysis, fMRIflows computes one beta contrast per condition per run. These beta contrasts are based on the same design matrix as the one for the univariate analysis. This step also creates a CSV-file containing a list of condition identifiers which can be used in the multivariate analysis to label the contrast maps.

Optional spatial normalization: The user can spatially normalize the estimated contrasts, if they have not been normalized during functional preprocessing. Contrasts need to be normalized, should the user want to use them in a 2nd-level analysis.

Supplementary Note 5: Description of 2nd-level univariate analysis pipeline steps

Model Specification and estimation: See description in Supplementary Note 4.

Univariate contrast estimation: See description in Supplementary Note 4.

Topological thresholding: Once the 2nd-level contrast is estimated, fMRIflows uses Nipype's Thresholding routine to apply first, a voxel-wise threshold, followed by a cluster-wise topological False Discovery Rate (FDR) correction. The user can decide which parameters to use and if topological thresholding should be applied or not.

Supplementary Note 6: Description of 2nd-level multivariate analysis pipeline steps

Data preparation: To prepare the data for the multivariate pattern analysis (MVPA) with PyMVPA the β -maps from the first level analysis are normalized by voxel-wise z-scoring them. A binary mask is then created to only include voxels in the searchlight analysis that have a value in at least one of the β -maps. This binary mask is then dilated by two voxels and eroded by one to make sure that the mask does not include single voxel holes. The z-scored and masked images are then saved together with a list of corresponding labels in a PyMVPA conform dataset.

Searchlight classification: The searchlight classification is performed for each subject individually and is based on the beta contrasts created during the 1st-level analysis. The user can specify which binary classification identifiers to use for the training and testing of the classifier. This approach allows the user to look for patterns that distinguish two classes (i.e. when the two classes are the same in training and testing) or to look for patterns that are identifying class differences (i.e. when the two classes are different during training and testing). The second approach allows the investigation of recurrent patterns, even if the stimuli are not the same. For example, training on the differences between visual stimuli of cats and dogs and testing on auditory stimuli of cats and dogs would reveal regions with distinct brain patterns for cats and dogs, independent of the modality of the stimuli. Natively, fMRIflows supports the following classifiers: C- and Nu-SVM classifiers with a linear or radial basis function kernel, Sparse Multinomial Logistic Regression (SMLR) classifier, Gaussian Naive Bayes classifier and k-Nearest-Neighbor classifier. The addition of any other classifiers from PyMVPA or Scikit-Learn is straightforward and can be done very quickly, if needed. Searchlight specific parameters, such as radius of the sphere and step size can be specified by the user. For each sphere, an N -fold leave-one-out cross-validation is performed, where N is set by the user, but usually represents the number of runs per subject. The result is one accuracy value per sphere. To account for holes caused by a searchlight step size bigger than one, results between different searchlight spheres are aggregated so that each voxel in the brain mask represents the average accuracy of all searchlight spheres that contained this voxel in their classification. This approach has the additional beneficial effect of increasing the SNR by smoothing the results before the group analysis.

Group analysis using T-test: The application of a simple T-test to test the classification accuracies against chance level is not recommended (Stelzer et al., 2013), and users should choose the appropriate group analysis as proposed by Stelzer et al. (2013). Nonetheless, fMRIflows contains this approach, as it can give some preliminary insights into the group results, using extremely reduced computation time.

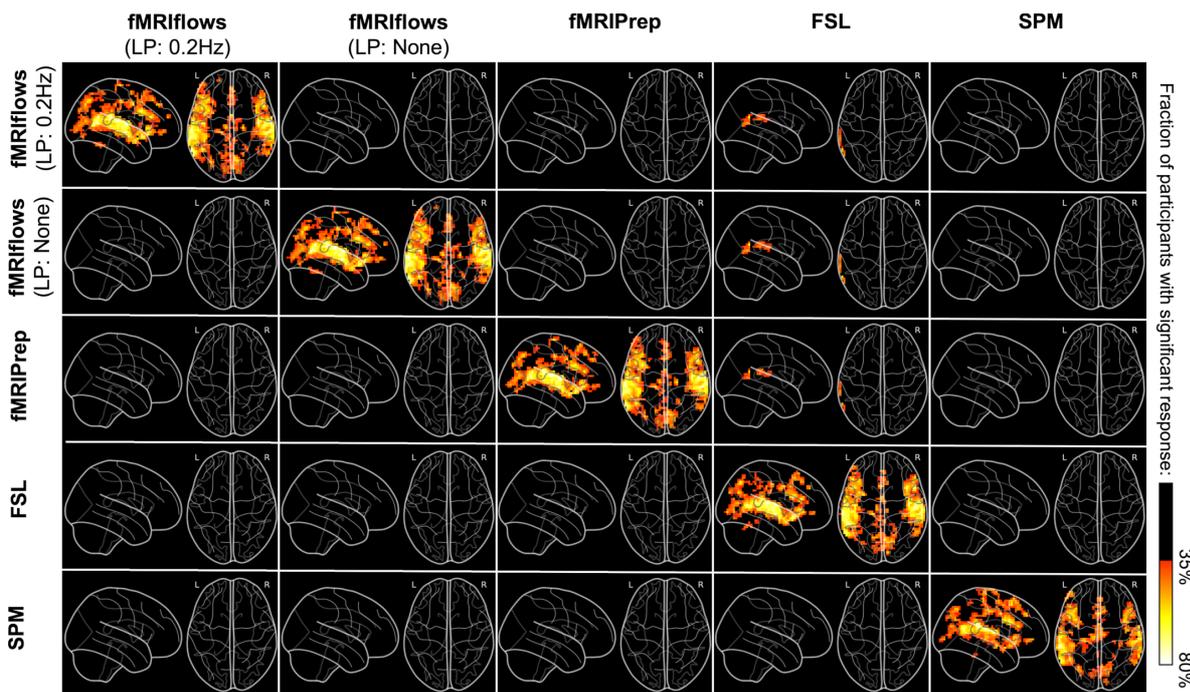
Group analysis using random permutations and cluster size control: The group analysis approach according to Stelzer et al. (2013) includes a permutation step where subject-specific classifications are rerun up to 100 times with randomized labels. In a later step, those randomized control accuracy maps are then combined up to 100'000 times into "false" group averages and help to estimate a voxel-wise null distribution and expected cluster size of the classification accuracy maps, given the dataset and specified classes. The subject specific searchlight accuracy maps are averaged over all subjects to

obtain a group classification accuracy map, which then is tested against the estimated null distribution. The user can additionally specify the strategy for the multiple comparison correction.

Topological thresholding: This step is identical to the one used in the 2nd-level univariate analysis.

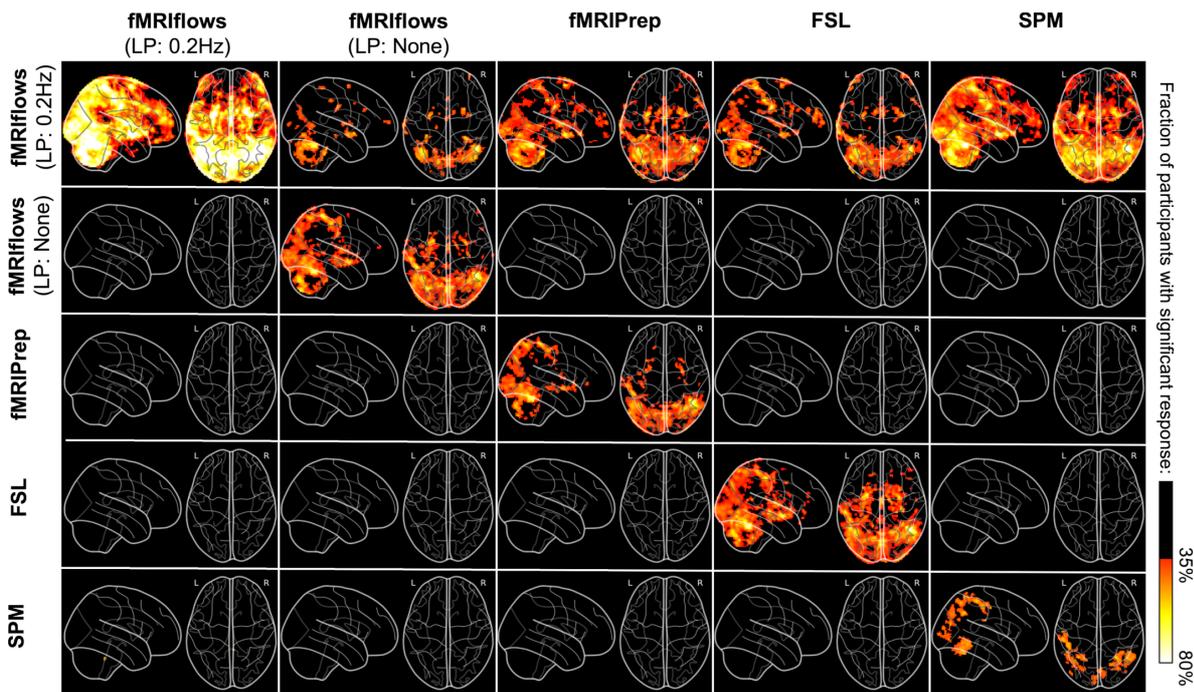
Supplementary Note 7: Complete 1st-level activation count maps results comparison between toolboxes, separated by dataset.

Dataset TR2000: The comparison between the binarized 1st-level activation count maps computed on the TR2000 dataset shows no clear differences between the toolboxes. The images on the diagonal all seem comparable. The diagonal offset images only show significantly increased activation counts for maps generated with fMRIflows and fMRIPrep, over activation count maps generated with FSL.



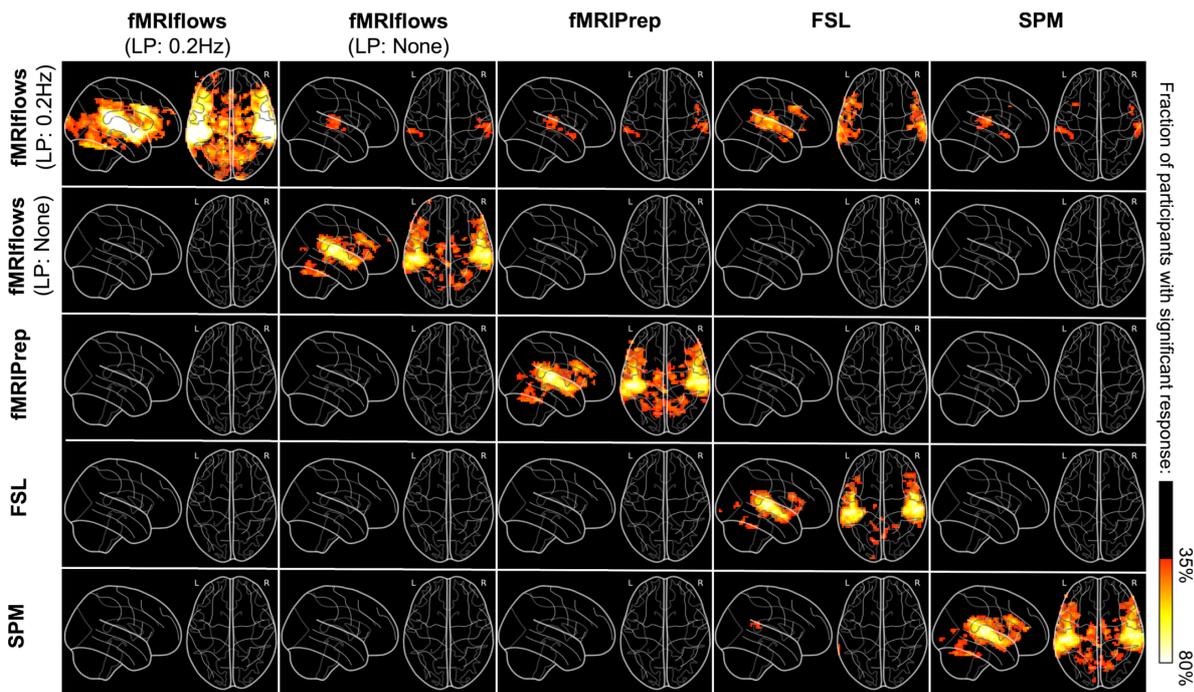
Supplementary Figure 1. Investigation of differences between binarized 1st-level activation count maps, thresholded at $p < 0.001$, after multiple functional preprocessing approaches analyzing dataset TR2000. Preprocessing was done with fMRIflows (with and without a temporal low-pass filter at 0.2 Hz), fMRIPrep, FSL and SPM (from top to bottom). The diagonal images represent the original activation count map of the toolbox. The diagonal offset images represent the difference between the horizontal and vertical toolbox. Activation count maps were normalized to the ICBM 2009c brain template. Color code represents the fraction of participants that show significant activation above a p-value threshold at 0.001 and corrected for false positive rate (FPR).

Dataset TR1000: The comparison between the binarized 1st-level activation count maps computed on the TR1000 dataset shows clear differences between fMRIfloWS with a low-pass filter at 0.2 Hz and the other approaches. The images on the diagonal seem overall similar. However, fMRIfloWS with a low-pass filter at 0.2 Hz shows a clearly increased (and SPM shows a clearly decreased) activation count map. The significantly increased activation count map values between fMRIfloWS with a low-pass filter at 0.2 Hz and the other approaches, seem to be centered around locations with already increased overlap, indicating that the low-pass filtering improves the overall statistical sensitivity.



Supplementary Figure 2: Investigation of differences between binarized 1st-level activation count maps, thresholded at $p < 0.001$, after multiple functional preprocessing approaches analyzing dataset TR1000. Preprocessing was done with fMRIfloWS (with and without a temporal low-pass filter at 0.2 Hz), fMRIPrep, FSL and SPM (from top to bottom). The diagonal images represent the original activation count map of the toolbox. The diagonal offset images represent the difference between the horizontal and vertical toolbox. Activation count maps were normalized to the ICBM 2009c brain template. Color code represents the fraction of participants that show significant activation above a p-value threshold at 0.001 and corrected for false positive rate (FPR).

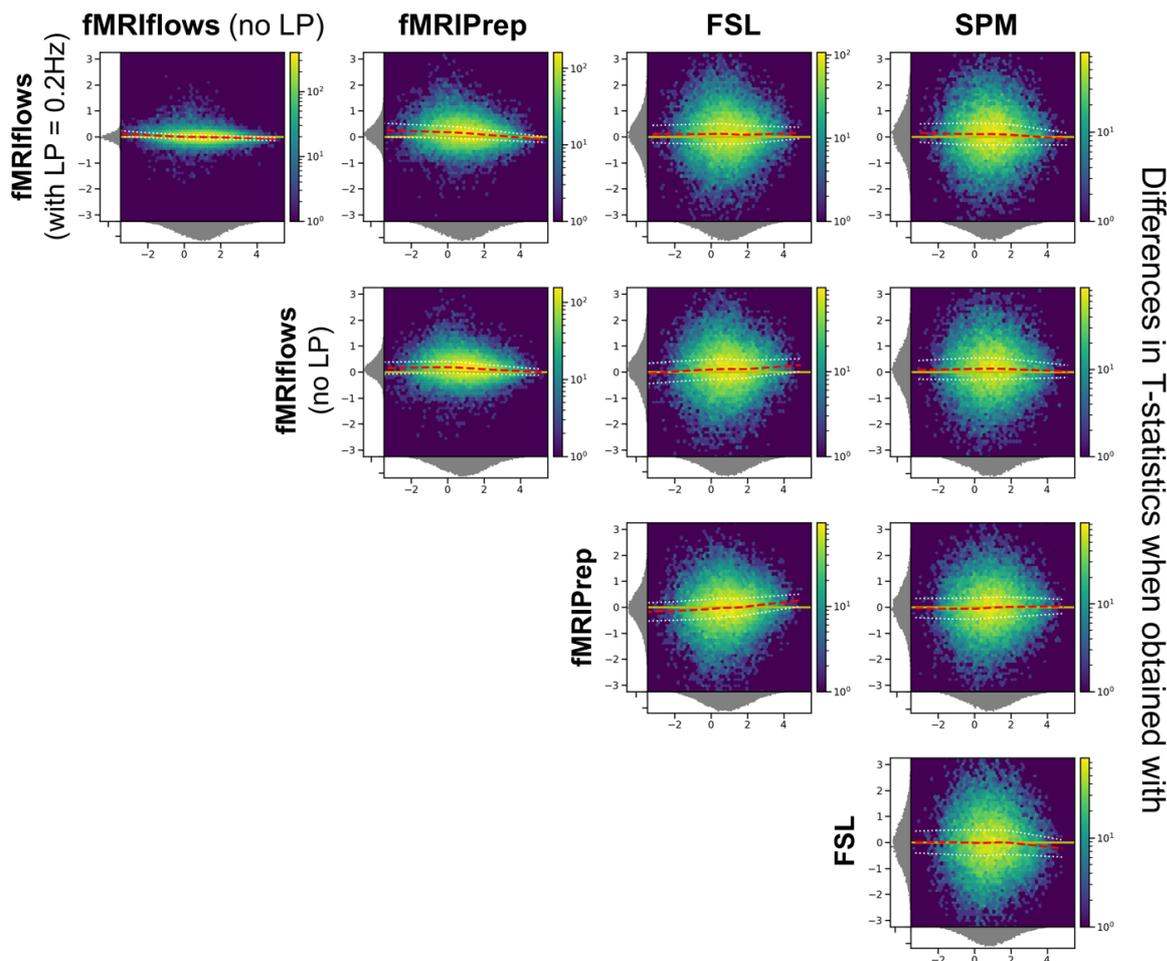
Dataset TR600: The comparison between the binarized 1st-level activation count maps computed on the TR600 dataset show clear differences between fMRIflores with a low-pass filter at 0.2 Hz and the other approaches. The images on the diagonal seem overall comparable. However, fMRIflores with a low-pass filter at 0.2 Hz shows a clearly increased activation count map. The significantly increased activation count map values between fMRIflores with a low-pass filter at 0.2 Hz and the other approaches, seem to be centered around locations with already increased overlap, indicating that the low-pass filtering improves the overall statistical sensitivity.



Supplementary Figure 3: Investigation of differences between binarized 1st-level activation count maps, thresholded at $p < 0.001$, after multiple functional preprocessing approaches analyzing dataset TR600. Preprocessing was done with fMRIflores (with and without a temporal low-pass filter at 0.2 Hz), fMRIPrep, FSL and SPM (from top to bottom). The diagonal images represent the original activation count map of the toolbox. The diagonal offset images represent the difference between the horizontal and vertical toolbox. Activation count maps were normalized to the ICBM 2009c brain template. Color code represents the fraction of participants that show significant activation above a p-value threshold at 0.001 and corrected for false positive rate (FPR).

Supplementary Note 8: Complete group-level T-statistic maps comparison between toolboxes, separated by dataset.

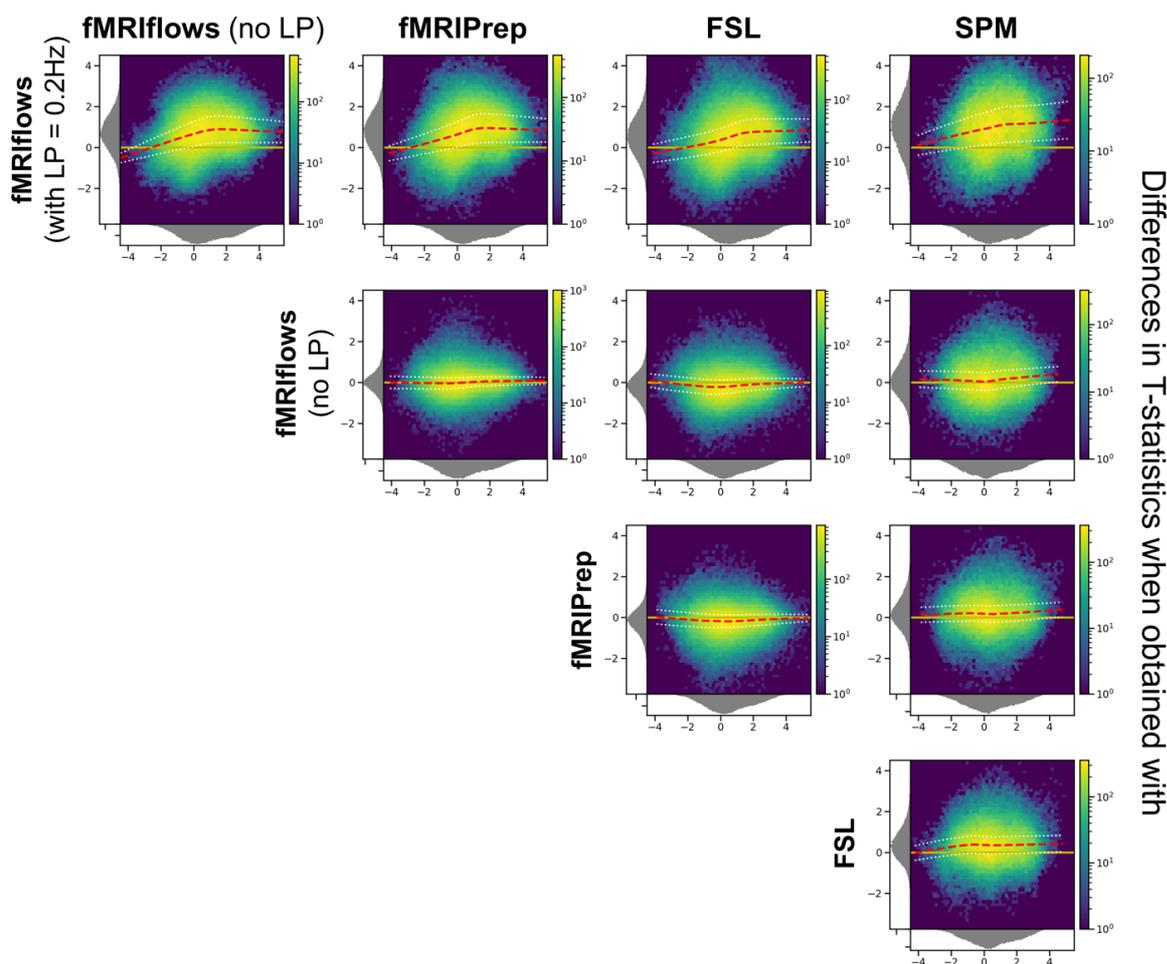
Dataset TR2000: The investigation of group-level T-statistic maps differences due to toolbox specific preprocessing pipelines on the TR2000 dataset does not show any clear differences between the results. The Bland-Altman 2D histograms show no particular trend in the x-direction and are centered around the horizontal zero line.



Supplementary Figure 4: Bland-Altman 2D histograms of datasets TR2000, comparing unthresholded group-level T-statistic maps between multiple processing approaches. Density plots show the relationship between average T-statistic value (horizontal) and difference of T-statistic values (vertical) at corresponding voxels for different pairwise combinations of toolboxes. The difference in T-statistics was computed in contrast to a preprocessing with (from top to bottom) fMRIflores with and without a low-pass filter at 0.2 Hz, fMRIPrep and FSL in respect to a preprocessing with (from left to right) fMRIflores without a low-pass filter at 0.2 Hz, fMRIPrep, FSL and SPM. Distribution plots next to x- and y- axis depict occurrence of a given value in this domain. Color code within figure indicates number

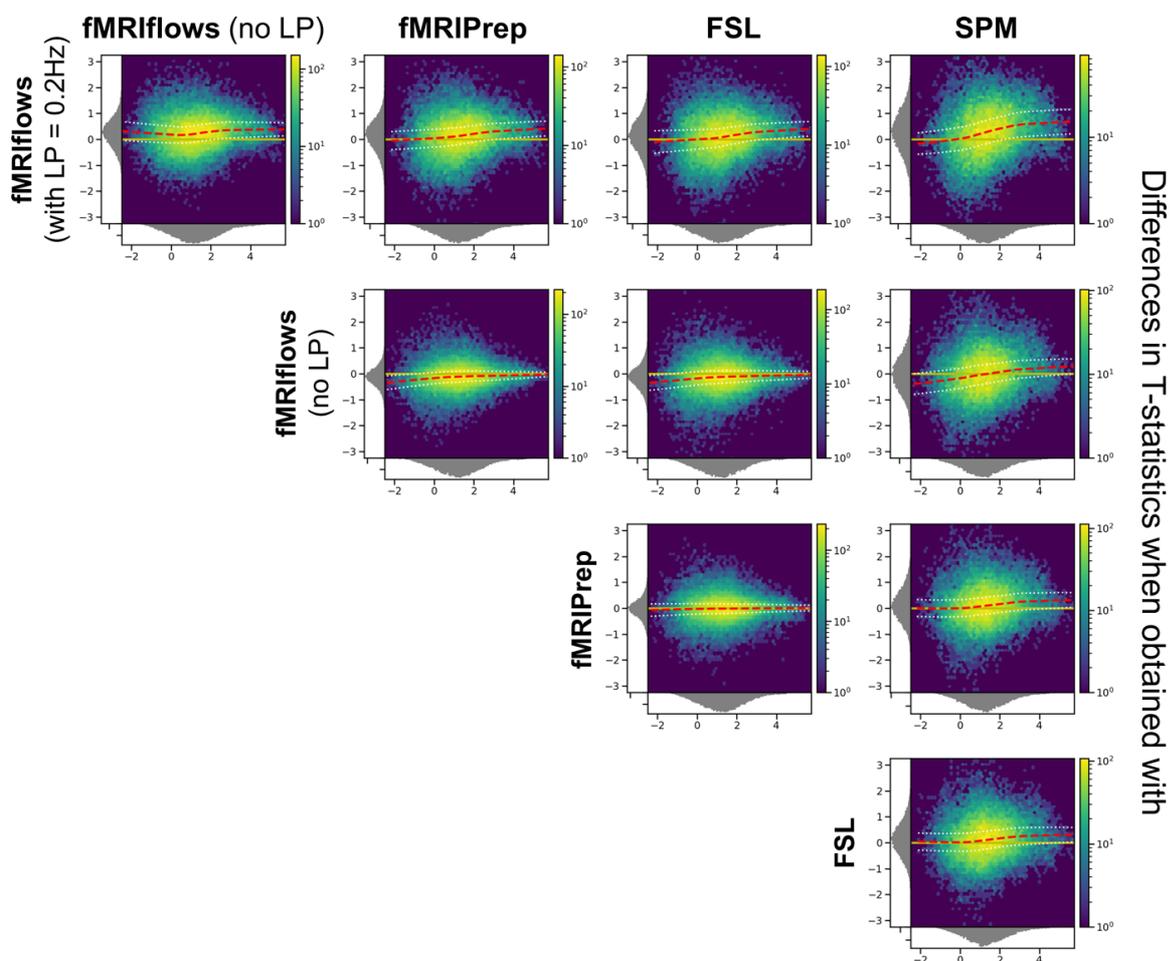
of voxels at this given overlap, from a few (blue) to many (yellow). Yellow horizontal line at zero indicates no value differences between corresponding voxels. Red dashed line depicts horizontal density average.

Dataset TR1000: The investigation of group-level T-statistic maps differences due to toolbox specific preprocessing pipelines on the TR1000 dataset shows clear differences between the results obtained with fMRIflores with a low-pass filter at 0.2 Hz and all the other approaches. The Bland-Altman 2D histograms show clearly increased t-statistic value differences in the first column of the following figure, especially for higher t-values. No clear differences can be seen between the analysis approach fMRIflores without a low-pass filter, fMRIPrep, FSL and SPM.



Supplementary Figure 5: Bland-Altman 2D histograms of datasets TR1000, comparing unthresholded group-level T-statistic maps between multiple processing approaches. Density plots show the relationship between average T-statistic value (horizontal) and difference of T-statistic values (vertical) at corresponding voxels for different pairwise combinations of toolboxes. The difference in T-statistics was computed in contrast to a preprocessing with (from top to bottom) fMRIflores with and without a low-pass filter at 0.2 Hz, fMRIPrep and FSL in respect to a preprocessing with (from left to right) fMRIflores without a low-pass filter at 0.2 Hz, fMRIPrep, FSL and SPM. Distribution plots next to x- and y- axis depict occurrence of a given value in this domain. Color code within figure indicates number of voxels at this given overlap, from a few (blue) to many (yellow). Yellow horizontal line at zero indicates no value differences between corresponding voxels. Red dashed line depicts horizontal density average.

Dataset TR600: The investigation of group-level T-statistic maps differences due to toolbox specific preprocessing pipelines on the TR600 dataset shows clear differences between the results obtained with fMRIflows with a low-pass filter at 0.2 Hz and all the other approaches. The Bland-Altman 2D histograms show increased t-statistic value differences in the first column of the following figure, especially for high t-values. No clear differences can be seen between the analysis approach fMRIflows without a low-pass filter, fMRIPrep, FSL and SPM.



Supplementary Figure 6: Bland-Altman 2D histograms of datasets TR600, comparing unthresholded group-level T-statistic maps between multiple processing approaches. Density plots show the relationship between average T-statistic value (horizontal) and difference of T-statistic values (vertical) at corresponding voxels for different pairwise combinations of toolboxes. The difference in T-statistics was computed in contrast to a preprocessing with (from top to bottom) fMRIflows with and without a low-pass filter at 0.2 Hz, fMRIPrep and FSL in respect to a preprocessing with (from left to right) fMRIflows without a low-pass filter at 0.2 Hz, fMRIPrep, FSL and SPM. Distribution plots next to x- and y- axis depict occurrence of a given value in this domain. Color code within figure indicates number of voxels at this given overlap, from a few (blue) to many (yellow). Yellow horizontal line at zero indicates no value differences between corresponding voxels. Red dashed line depicts horizontal density average.

References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8(February), 14. <https://doi.org/10.3389/fninf.2014.00014>
- Ashburner, J. (2009). Preparing fMRI data for statistical analysis. In *fMRI techniques and protocols* (pp. 151–178). Springer.
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, 54(3), 2033–2044. <https://doi.org/10.1016/j.neuroimage.2010.09.025>
- Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., ... Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, 582(7810), 84–88. <https://doi.org/10.1038/s41586-020-2314-9>
- Botvinik-Nezer, R., Iwanir, R., Holzmeister, F., Huber, J., Johannesson, M., Kirchler, M., ... Schonberg, T. (2019). fMRI data of mixed gambles from the Neuroimaging Analysis Replication and Prediction Study. *Scientific Data*, 6(1), 106. <https://doi.org/10.1038/s41597-019-0113-7>
- Bowring, A., Maumet, C., & Nichols, T. (2018). Exploring the impact of analysis software on task fMRI results. *BioRxiv*, 285585.
- Brett, M., Hanke, M., Markiewicz, C., Côté, M.-A., McCarthy, P., Ghosh, S., ... Basile. (2018). *nibabel: Access a cacophony of neuro-imaging file formats, version 2.3.0*. <https://doi.org/10.5281/zenodo.1287921>
- Caballero-Gaudes, C., & Reynolds, R. C. (2016). Methods for cleaning the BOLD fMRI signal. *NeuroImage*, (December), 0–1. <https://doi.org/10.1016/j.neuroimage.2016.12.018>
- Caballero-Gaudes, C., & Reynolds, R. C. (2017). Methods for cleaning the BOLD fMRI signal. *NeuroImage*, 154(December 2016), 128–149. <https://doi.org/10.1016/j.neuroimage.2016.12.018>
- Carp, J. (2012). The secret lives of experiments: methods reporting in the fMRI literature. *NeuroImage*, 63(1), 289–300. <https://doi.org/10.1016/j.neuroimage.2012.07.004>
- Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, 10(4–5), 171–178. [https://doi.org/10.1002/\(SICI\)1099-1492\(199706/08\)10:4/5<171::AID-NBM453>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L)
- Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLoS One*, 12(9), e0184661. <https://doi.org/10.1371/journal.pone.0184661>
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., ... Gorgolewski,

- K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1), 111–116. <https://doi.org/10.1038/s41592-018-0235-4>
- Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Gunther, M., ... Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS One*, 5(12), e15710. <https://doi.org/10.1371/journal.pone.0015710>
- Feinberg, D. A., & Setsompop, K. (2013). Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *Journal of Magnetic Resonance (San Diego, Calif. : 1997)*, 229, 90–100. <https://doi.org/10.1016/j.jmr.2013.02.002>
- Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>
- Fonov, V., Evans, A. C., Botteron, K., Almlí, C. R., McKinstry, R. C., & Collins, D. L. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1), 313–327. <https://doi.org/10.1016/j.neuroimage.2010.07.033>
- Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996). Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3), 346–355. <https://doi.org/10.1002/mrm.1910350312>
- Friston, K., Penny, W., Ashburner, J., Kiebel, S., & Nichols, T. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (Vol. 8). <https://doi.org/10.1016/B978-012372560-8/50002-4>
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5(August), 13. <https://doi.org/10.3389/fninf.2011.00013>
- Gorgolewski, K., Esteban, O., Schaefer, G., Wandell, B., & Poldrack, R. (2017). OpenNeuro - a free online platform for sharing and analysis of neuroimaging data. *Organization for Human Brain Mapping. Vancouver, Canada*, 1677. <https://doi.org/10.1038/sdata.2016.44.3>
- Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., ... Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3, 160044. <https://doi.org/10.1038/sdata.2016.44>
- Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., ... Margulies, D. S. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in Neuroinformatics*, 9(April), 8. <https://doi.org/10.3389/fninf.2015.00008>
- Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., ... Haase, A. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6), 1202–1210. <https://doi.org/10.1002/mrm.10171>
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, 7(1),

- 37–53. <https://doi.org/10.1007/s12021-008-9041-y>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage*, 62(2), 782–790.
- Jones, E., Oliphant, T., Peterson, P., & others. (2001). *{SciPy}: Open source scientific tools for {Python}*. Retrieved from <http://www.scipy.org/>
- Kluyver, T., Ragan-kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., ... Willing, C. (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, 87–90. <https://doi.org/10.3233/978-1-61499-649-1-87>
- Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 1(1), 76–85. <https://doi.org/10.1137/0701007>
- Lindquist, M. A., Geuter, S., Wager, T. D., & Caffo, B. S. (2019). Modular preprocessing pipelines can reintroduce artifacts into fMRI data. *Human Brain Mapping*, (January), 407676. <https://doi.org/10.1002/hbm.24528>
- McKinney, W., & others. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, 445, 51–56. <https://doi.org/10.3389/fninf.2014.00014>
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Ugurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic Resonance in Medicine*, 63(5), 1144–1153. <https://doi.org/10.1002/mrm.22361>
- Notter, M., Gale, D., Herholz, P., Markello, R., Notter-Bielsler, M.-L., & Whitaker, K. (2019). AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI images. *Journal of Open Source Software*, 4(34), 1257. <https://doi.org/10.21105/joss.01257>
- Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science & Engineering*, 9(3). <https://doi.org/10.1109/MCSE.2007.58>
- Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.
- Power, J. D. (2017). A simple but useful way to assess fMRI scan qualities. *NeuroImage*, 154(August 2016), 150–158. <https://doi.org/10.1016/j.neuroimage.2016.08.009>
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>
- Power, J. D., Plitt, M., Kundu, P., Bandettini, P. A., & Martin, A. (2017). Temporal interpolation alters motion in fMRI scans: Magnitudes and consequences for artifact detection. *PloS One*, 12(9), e0182939. <https://doi.org/10.1371/journal.pone.0182939>

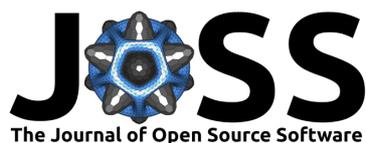
- Power, J. D., Plitt, M., Laumann, T. O., & Martin, A. (2017). Sources and implications of whole-brain fMRI signals in humans. *NeuroImage*, *146*(May 2016), 609–625. <https://doi.org/10.1016/j.neuroimage.2016.09.038>
- Sengupta, A., Pollmann, S., & Hanke, M. (2018). Spatial band-pass filtering aids decoding musical genres from auditory cortex 7T fMRI. *F1000Research*, *7*(0), 142. <https://doi.org/10.12688/f1000research.13689.2>
- Sladky, R., Friston, K. J., Tröstl, J., Cunnington, R., Moser, E., & Windischberger, C. (2011). Slice-timing effects and their correction in functional MRI. *NeuroImage*, *58*(2), 588–594. <https://doi.org/10.1016/j.neuroimage.2011.06.078>
- Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., ... WU-Minn HCP Consortium. (2013). Resting-state fMRI in the Human Connectome Project. *NeuroImage*, *80*, 144–168. <https://doi.org/10.1016/j.neuroimage.2013.05.039>
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., ... Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*(SUPPL. 1), S208-19. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
- Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. *NeuroImage*, *65*, 69–82. <https://doi.org/10.1016/j.neuroimage.2012.09.063>
- Stephen Butterworth. (1930). On the Theory of Filter Amplifiers. *Experimental Wireless and the Wireless Engineer*, Vol. 7, pp. 536–541.
- Strother, S. C. (2006). Evaluating fMRI preprocessing pipelines. *IEEE Engineering in Medicine and Biology Magazine : The Quarterly Magazine of the Engineering in Medicine & Biology Society*, *25*(2), 27–41. <https://doi.org/10.1109/MEMB.2006.1607667>
- Viessmann, O., Möller, H. E., & Jezzard, P. (2018). Dual regression physiological modeling of resting-state EPI power spectra: Effects of healthy aging. *NeuroImage*, *187*(December 2017), 68–76. <https://doi.org/10.1016/j.neuroimage.2018.01.011>
- Yarkoni, T., Markiewicz, C., de la Vega, A., Gorgolewski, K., Salo, T., Halchenko, Y., ... Blair, R. (2019). PyBIDS: Python tools for BIDS datasets. *Journal of Open Source Software*, *4*(40), 1294. <https://doi.org/10.21105/joss.01294>

Annex 2

Study B – AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI images

Michael P. Notter, Dan Gale, Peer Herholz, Ross Markello, Marie-Laure Notter-Bielsler,
and Kirstie Whitaker

Published on *2019 Feb 24*, in *Journal of Open Source Software*, 4(34), 1257.



AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI images

Michael Philipp Notter¹, Dan Gale², Peer Herholz^{3, 4, 5}, Ross Markello⁵, Marie-Laure Notter-Bielsler⁶, and Kirstie Whitaker⁷

1 The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology and Department of Clinical Neurosciences, Lausanne, Switzerland; Center for Biomedical Imaging (CIBM), Lausanne, Switzerland **2** Centre for Neuroscience Studies, Queen's University, Kingston, Canada **3** Laboratory for Multimodal Neuroimaging, Philipps-University Marburg, Hesse, Germany **4** International Laboratory for Brain, Music and Sound Research, Université de Montréal & McGill University, Montréal, Canada **5** McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, Canada **6** Centre Leenaards de la Mémoire, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland **7** Alan Turing Institute, London, UK; Department of Psychiatry, University of Cambridge, Cambridge, UK

DOI: [10.21105/joss.01257](https://doi.org/10.21105/joss.01257)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 05 February 2019

Published: 24 February 2019

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

A major advantage of magnetic resonance imaging (MRI) over other neuroimaging methods is its capability to noninvasively locate a region of interest (ROI) in the human brain. For example, using functional MRI, we are able to pinpoint where in the brain a cognitive task elicits higher activation relative to a control. But just knowing the Cartesian coordinate of such a ROI is not useful if we cannot assign it a neuroanatomical label. For this reason, MRI images are usually normalized into a common template space (Fonov et al., 2011), where well-established atlases can be used to associate a given coordinate with the label of a brain region. Most major neuroimaging software packages provide some functionality to locate the main peaks of an ROI but this functionality is often restricted to a few atlases, frequently requires manual intervention, does not give the user much flexibility in the output creation process, and never considers the full extent of the ROI.

To tackle those shortcomings, we created AtlasReader, a Python interface for generating coordinate tables and region labels from statistical MRI images. With AtlasReader, users can use any of the freely and publicly available neuroimaging atlases, without any restriction to their preferred software package, to create publication-ready output figures and tables that contain relevant information about the peaks and clusters extent of each ROI. To our knowledge, providing atlas information about the full extent of a cluster, i.e. over which atlas regions does a ROI extent, is a new feature that is not available in any other, comparable neuroimaging software package.

Executing AtlasReader on an MRI image will create the following four outputs:

1. An **overview figure** showing all ROIs throughout the whole brain (Fig. 1).
2. For each ROI, an **informative figure** showing the sagittal, coronal and transversal plane centered on the main peak of the ROI (Fig. 2).
3. A **table** containing information about the main **peaks** in each ROI (Fig. 3).
4. A **table** containing information about the **cluster extent** of each ROI (Fig. 4).

Users have many parameters available to guide the creation of these outputs. For example, with `cluster_extent` a user can specify the minimum number of contiguous voxels

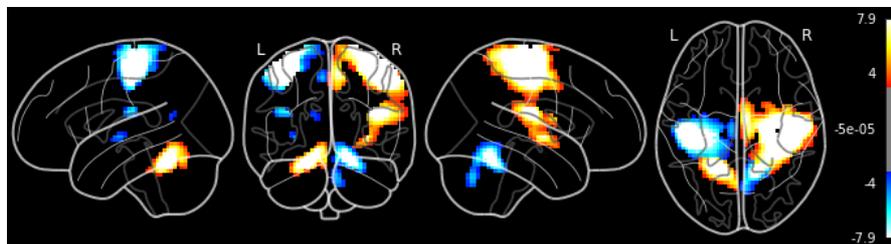
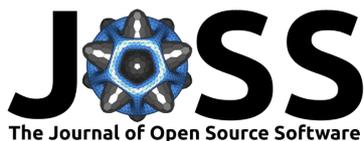


Figure 1: Overview figure showing the ROIs throughout the whole brain at once.

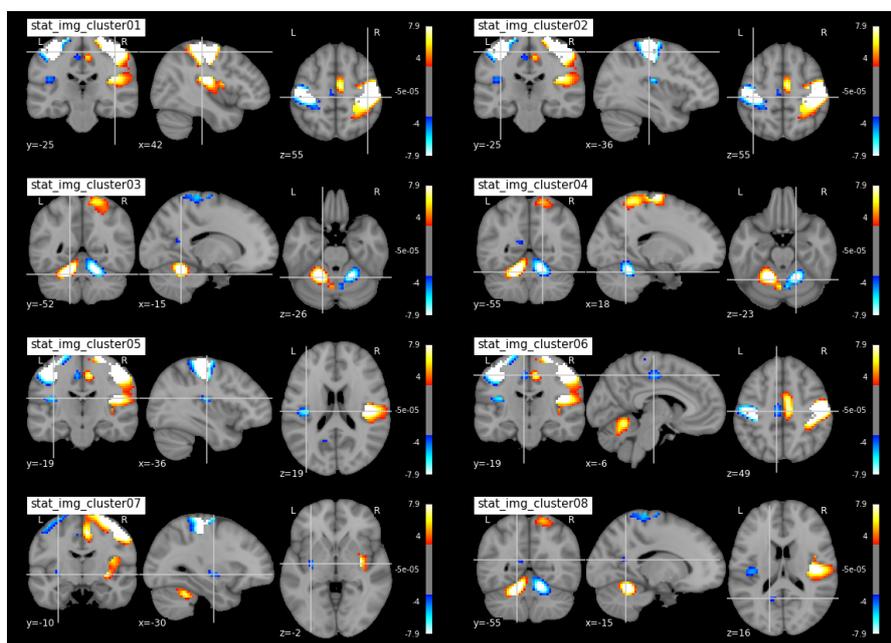


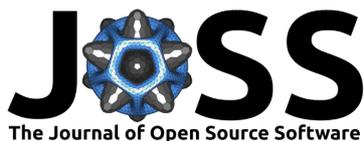
Figure 2: Eight cluster figures, each centered on the main peak of the ROI, showing the sagittal, coronal and transversal plane of the ROI.

required for a ROI to be shown in the output, `min_distance` can be used to extract information from multiple peaks within a given ROI, and `atlas` can be used to specify which atlases should be used for the output creation. By default, AtlasReader uses the AAL, the Desikan-Killiany, and the Harvard-Oxford atlases (Fig. 5). In the current version, users also have access to the Aicha, the Destrieux, the Juelich, the Marsatlas, the Neuro-morphometrics, and the Talairach atlas. Further details about the individual atlases, how to acknowledge them, and their license requirements are detailed in the [atlasreader/data](#) directory.

AtlasReader is licensed under the BSD-3 license and depends on the following python libraries: `matplotlib` (Hunter, 2007), `nibabel` (Brett et al., 2018), `nilearn` (Abraham et al., 2014), `numpy` (Oliphant, 2007), `scipy` (Jones, Oliphant, Peterson, & others, 2001), `scikitlearn` (Pedregosa et al., 2011) and `scikitimage` (Van der Walt et al., 2014).

For a more detailed explanation about how AtlasReader works and instructions on how to install the software on your system, see <https://github.com/miykael/atlasreader>.

Notter et al., (2019). AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI 2 images. *Journal of Open Source Software*, 4(34), 1257. <https://doi.org/10.21105/joss.01257>



cluster_id	peak_x	peak_y	peak_z	peak_value	volume_mm	aal	desikan_killiany	harvard_oxford
1	6	-10	52	7.94135	58563	Supp_Motor_Area_R	ctx-rh-paracentral	53% Right_Juxtapositional_Lobule_Cortex_(forme...
1	45	-19	16	7.94135	58563	Rolandic_Oper_R	Unknown	43% Right_Central_Opercular_Cortex; 40% Right_...
1	42	-25	58	7.94135	58563	Postcentral_R	ctx-rh-postcentral	43% Right_Postcentral_Gyrus; 17% Right_Precen...
1	33	-7	-2	7.90531	58563	Putamen_R	Right-Putamen	71% Right_Putamen
1	42	-1	13	5.47070	58563	Rolandic_Oper_R	Unknown	74% Right_Central_Opercular_Cortex; 7% Right_I...
1	9	2	73	3.56015	58563	Supp_Motor_Area_R	ctx-rh-superiorfrontal	49% Right_Superior_Frontal_Gyrus; 6% Right_Jux...
2	-30	-19	67	-7.94144	19089	Precentral_L	Left-Cerebral-White-Matter	46% Left_Precentral_Gyrus
3	-15	-52	-26	7.94135	9612	no_label	Left-Cerebellum-Cortex	0% no_label
4	18	-55	-23	-7.94144	8505	Cerebellum_6_R	Right-Cerebellum-Cortex	0% no_label
4	6	-70	-38	-5.30572	8505	Vermis_8	Right-Cerebellum-Cortex	0% no_label
5	-36	-19	19	-6.21808	1161	Insula_L	Unknown	37% Left_Central_Opercular_Cortex; 37% Left_In...
6	-6	-19	49	-5.03538	1134	Cingulate_Mid_L	ctx-lh-paracentral	50% Left_Precentral_Gyrus; 9% Left_Juxtapositi...
7	-30	-10	-2	-4.65454	378	Putamen_L	Left-Putamen	98% Left_Putamen
8	-15	-55	16	-3.57240	243	Precuneus_L	Left-Cerebral-White-Matter	35% Left_Precuneous_Cortex

Figure 3: Example of a peak table showing relevant information for the main peaks of each ROI. This table contains the cluster association and location of each peak, its signal value at this location, the cluster extent (in mm, not in number of voxels), as well as the membership of each peak, given a particular atlas.

cluster_id	peak_x	peak_y	peak_z	cluster_mean	volume_mm	aal	desikan_killiany	harvard_oxford
1	42	-25	55	5.80230	58563	29.09% Postcentral_R; 15.17% Precentral_R; 9.1...	31.21% Unknown; 27.43% Right-Cerebral-White-Ma...	28.54% Right_Postcentral_Gyrus; 19.59% Right_P...
2	-36	-25	55	-5.96750	19089	60.82% Postcentral_L; 26.45% Precentral_L; 5.9...	47.81% Left-Cerebral-White- Matter; 19.09% ctx-...	61.10% Left_Postcentral_Gyrus; 35.06% Left_Pre...
3	-15	-52	-26	5.42533	9612	44.10% Cerebellum_6_L; 33.99% Cerebellum_4_5_L; ...	75.84% Left-Cerebellum- Cortex; 19.94% Left-Cer...	78.37% no_label; 10.39% Left_Lingual_Gyrus; 5...
4	18	-55	-23	-5.04111	8505	32.70% Cerebellum_6_R; 32.70% Cerebellum_4_5_R; ...	76.51% Right-Cerebellum- Cortex; 12.38% Right-C...	81.90% no_label; 16.19% Right_Lingual_Gyrus
5	-36	-19	19	-4.36624	1161	72.09% Rolandic_Oper_L; 27.91% Insula_L	48.84% Unknown; 30.23% ctx- lh-supramarginal; 1...	58.14% Left_Central_Opercular_Cortex; 25.58% L...
6	-6	-19	49	-3.82011	1134	50.00% Cingulate_Mid_L; 40.48% Supp_Motor_Area...	40.48% ctx-lh-paracentral; 30.95% Unknown; 14...	71.43% Left_Precentral_Gyrus; 16.67% Left_Juxt...
7	-30	-10	-2	-3.67586	378	92.86% Putamen_L; 7.14% no_label	100.00% Left-Putamen	100.00% Left_Putamen
8	-15	-55	16	-3.28974	243	77.78% Precuneus_L; 11.11% Cuneus_L; 11.11% Ca...	44.44% Left-Cerebral-White- Matter; 33.33% Unkn...	100.00% Left_Precuneous_Cortex

Figure 4: Example of a cluster table showing relevant information for the cluster extent of each ROI. This table contains the cluster association and location of each peak, the mean value within the cluster, the cluster extent (in mm, not in number of voxels), as well as the membership of each cluster, given a particular atlas.

Notter et al., (2019). AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI 3 images. *Journal of Open Source Software*, 4(34), 1257. <https://doi.org/10.21105/joss.01257>

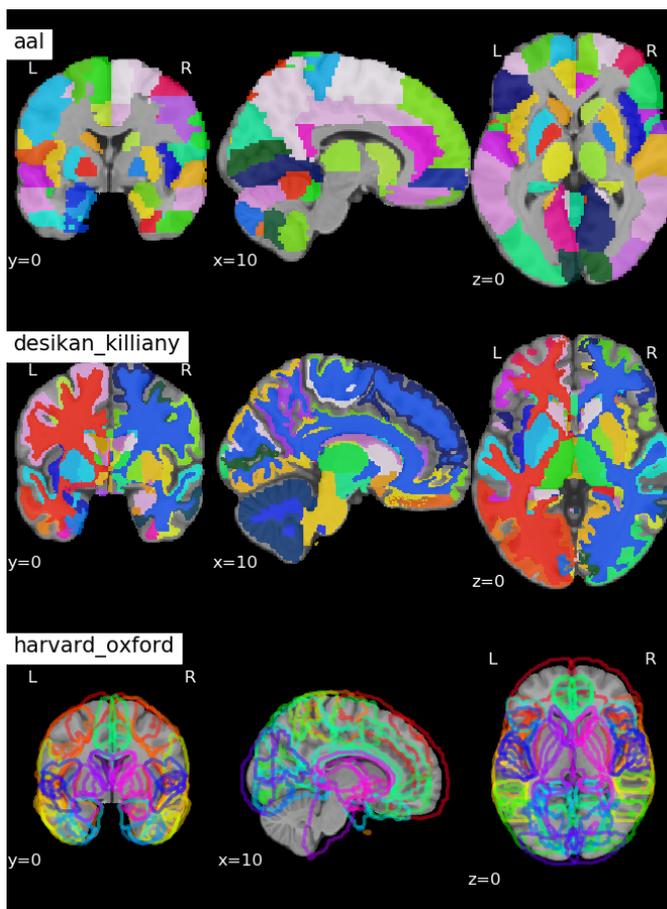
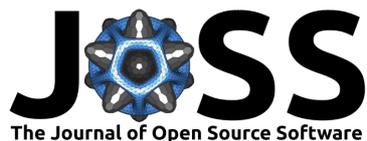
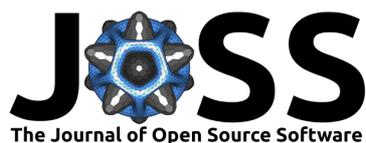


Figure 5: Depiction of AtlasReader’s default atlases. Individually colored label of the three default atlases, AAL, Desikan-Killiany and Harvard-Oxford, overlaid on the ICBM 2009c nonlinear asymmetric atlas. The Harvard-Oxford atlas is visualized differently because it is a probability atlas and therefore has overlapping regions.

Notter et al., (2019). AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI 4 images. *Journal of Open Source Software*, 4(34), 1257. <https://doi.org/10.21105/joss.01257>



Acknowledgement

Michael Philipp Notter is grateful for the support of the Swiss National Science Foundation grants, awarded to M. Murray (320030-149982, 320030-169206, and 51NF40-158776). Kirstie Whitaker is funded through an Alan Turing Institute Research Fellowship under EPSRC Research grant (TU/A/000017). This project was started at Neurohackademy 2018 which is funded by the National Institute of Mental Health through a grant to Ariel Rokem and Tal Yarkoni (R25MH112480).

References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., et al. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in neuroinformatics*, 8, 14. doi:[10.3389/fninf.2014.00014](https://doi.org/10.3389/fninf.2014.00014)
- Brett, M., Hanke, M., Markiewicz, C., Côté, M.-A., McCarthy, P., Ghosh, S., Wassermann, D., et al. (2018). Nibabel: Access a cacophony of neuro-imaging file formats, version 2.3.0. doi:[10.5281/zenodo.1287921](https://doi.org/10.5281/zenodo.1287921)
- Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., Group, B. D. C., et al. (2011). Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*, 54(1), 313–327. doi:[10.1016/j.neuroimage.2010.07.033](https://doi.org/10.1016/j.neuroimage.2010.07.033)
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, 9(3), 90–95. doi:[10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- Jones, E., Oliphant, T., Peterson, P., & others. (2001). SciPy: Open source scientific tools for Python. Retrieved from <http://www.scipy.org/>
- Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science & Engineering*, 9(3). doi:[10.1109/MCSE.2007.58](https://doi.org/10.1109/MCSE.2007.58)
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., et al. (2011). Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct), 2825–2830. Retrieved from <https://scikit-learn.org/>
- Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., et al. (2014). Scikit-image: Image processing in python. *PeerJ*, 2, e453. doi:[10.7717/peerj.453](https://doi.org/10.7717/peerj.453)

Annex 3

Study C – Decoding of multisensory semantics and memories in low-level visual cortex

Michael P. Notter, Anna Gaglianese, Sandra Da Costa, and Micah M. Murray

Revision for *Neuroimage*, in preparation.

1 **Decoding of multisensory semantics and memories in low-level visual cortex**

2

3 Michael P. Notter*¹, Sandra Da Costa², and Micah M. Murray^{1,3-5}

4

5 The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, University
6 Hospital Center and University of Lausanne, Switzerland

7 ²Center for Biomedical Imaging (CIBM), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausan-
8 ne, Switzerland

9 ³Sensory, Perceptual and Cognitive Neuroscience Section, Center for Biomedical Imaging (CIBM),
10 Lausanne, Switzerland

11 ⁴Department of Ophthalmology, Fondation Asile des aveugles and University of Lausanne, Lausanne,
12 Switzerland

13 ⁵Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

14

15

16 ***Corresponding Author**

17 Michael Notter

18 University Hospital Center and University of Lausanne

19 Radiology Research Center, RC7, 04/021

20 Rue du Bugnon 46

21 CH-1011 Lausanne, Switzerland

22 Tel: +41 (0)79 786 47 17

23 Email: michael.notter@epfl.ch

24

1 **Abstract**

2 Ample evidence indicates that recognition memory performance can be facilitated by multisensory
3 stimuli at the time of encoding. While it is not necessary that multisensory information is present dur-
4 ing the decoding of a memory, a beneficial effect of the multisensory stimulation only takes place if
5 the initial encoding happened in a semantically congruent audio-visual context. The goal of this study
6 was to investigate the brain mechanisms involved during the encoding and subsequent retrieval of
7 semantically congruent multisensory objects. In a functional MRI paradigm, participants performed a
8 continuous recognition task in which they discriminated between initially and previously seen objects.
9 The task was independent of the semantic congruence of the sound with the initially shown image. We
10 performed a univariate analysis to identify regions involved in the information processing of semantic
11 context dependent multisensory memory traces. Next, a multivariate pattern analysis (MVPA) located
12 where the representational content of these traces is encoded and later again retrieved. We show that
13 low-level visual cortex can reliably decode whether incoming visual stimuli had been previously en-
14 countered on a single-trial with either a semantically congruent or incongruent sound. Aside from fur-
15 ther reinforcing the notion that low-level visual cortex is fundamentally multisensory in its architec-
16 ture, our findings suggest that these functions extend to include both semantically-related and
17 memory-related functions.

18

19 **Keywords:** multisensory integration, multivariate pattern analysis, memory, semantic context, audito-
20 ry & visual perception, fMRI

21

1 **Introduction**

2 Most studies investigating how our brain creates and retrieves memories focus predominantly
3 on unisensory stimuli. However, in naturalistic environments, stimuli are rarely experienced in only
4 one sensory modality. Multisensory information is differentially processed by our brains and can posi-
5 tively impact our behavior (Calvert, Spence, & others, 2004; Molholm, Ritter, Javitt, & Foxe, 2004;
6 Murray et al., 2004; Murray & Wallace, 2012; Thomas, Nardini, & Mareschal, 2017). This has been
7 extended from perception of current information to the manner in which multisensory information
8 promotes distinctly different object representations in the brain that can affect memory-related pro-
9 cesses. Research has shown that brain regions involved in encoding multisensory memories are also
10 again active during subsequent retrieval, even if this retrieval is based on unisensory information
11 (Gibson & Maunsell, 1997; Nyberg, Habib, McIntosh, & Tulving, 2000; Wheeler, Petersen, &
12 Buckner, 2000).

13 Several studies have characterized how memory traces generated from previous multisensory
14 experiences differed from those created from exclusively unisensory experiences. In particular, the
15 lateral occipital cortex (LOC) and the superior temporal cortex (STC) appear to play an important role
16 in the retrieval of multisensory memories (Doehrmann & Naumer, 2008; Lu, Zhang, Xu, & Liu, 2018;
17 Matusz, Wallace, & Murray, 2017). In the case of visual objects that had been presented initially with
18 semantically congruent sounds (versus those only encountered visually), both electroencephalography
19 (EEG) and functional magnetic resonance imaging (fMRI) have shown that the LOC responds more
20 strongly (Murray, Foxe, & Wylie, 2005; Murray et al., 2004) reviewed in (Matusz et al., 2017) see
21 also (Matusz et al., 2015) for the case of auditory objects impacted by semantically congruent images).

22 Our group has worked to identify determinant factors for differential effect of multisensory ver-
23 sus unisensory experiences on memory performance (reviewed in (Matusz et al., 2017; Thelen &
24 Murray, 2013)). In particular, semantic context appears to play a critical role. Visual objects initially
25 experienced with a semantically congruent sound, even if task-irrelevant and unpredictable, are better
26 remembered when repeated in a unisensory visual context than those objects initially experienced
27 without a sound or with a sound that was either meaningless or semantically incongruent (Lehmann &
28 Murray, 2005; Murray et al., 2005; Thelen, Talsma, & Murray, 2015). This pattern also holds when

1 participants perform the task in the auditory modality and visual information is task-irrelevant (Matusz
2 et al., 2015; Thelen et al., 2015). We have contended that such effects might rely on multisensory
3 stimuli either themselves creating or accessing an enhanced memory trace that is, in turn, activated
4 more readily or efficiently upon repetition of just the task-relevant unisensory component several trials
5 later (Matusz et al., 2017). This conjecture is largely based, at present, on psychophysical results.

6 The effect of multisensory semantic context on current stimulus processing has been investigat-
7 ed in prior fMRI studies. In particular, the superior temporal cortex (STC) seems to have higher sensi-
8 tivity to semantically congruent stimuli (van Atteveldt, Formisano, Goebel, & Blomert, 2004), while
9 the inferior frontal cortex (IFC) is responsive to incongruent stimuli (Jung, Larsen, & Walther, 2018;
10 Noppeney, Josephs, Hocking, Price, & Friston, 2008). These findings support the notion that the acti-
11 vation of cortical object representations varies with semantic congruence. However and to our
12 knowledge, no investigation has described the brain mechanisms involved during the encoding and
13 subsequent retrieval of semantically congruent multisensory objects. Such is necessary to provide evi-
14 dence regarding where in the brain the creation and reactivation of those memory traces take place.

15 Thus, the present fMRI study aimed (1) to identify brain networks that are involved in the pro-
16 cess of creating and retrieving memory traces of semantically congruent episodic multisensory experi-
17 ences; and (2) to identify where in the brain these mechanisms take place. To answer the first question,
18 we used a standard univariate analysis approach to analyze fMRI data of an experiment with multisen-
19 sory stimuli in different semantic contexts, which will provide information about the overall activation
20 differences between conditions and therefore will highlight which brain regions are involved during
21 the processing of the stimuli. While this type of analysis can show the degree of involvement of a giv-
22 en region, it does not provide information about the representational content that is computed at this
23 location. For this purpose, we performed a multivariate pattern analysis (MVPA) on the same dataset.
24 Determining the distinct activation pattern during the encoding of semantically congruent and incon-
25 gruent multisensory memory objects and comparing this pattern to the activation pattern during the
26 subsequent retrieval of those memory traces based on exclusively unisensory visual inputs, will allow
27 us to identify where in the brain the multisensory object representations are processed during encoding
28 and retrieval.

1

1 **Materials and Methods**

2 **Participants**

3 Twelve healthy adults (8 female; aged 22-35 years mean \pm SD = 28.06 \pm 3.29) participated in
4 this study. All reported normal hearing and normal (or corrected-to-normal) vision, and provided writ-
5 ten informed consent in accordance with procedures approved by the cantonal ethics committee of
6 Vaud, Switzerland and according to the World Medical Association Helsinki Declaration (WMA
7 General Assembly, 2008). Subjects were compensated a fixed sum for their participation and were free
8 to end their participation at any time. All subjects were right-handed according to the Edinburgh hand-
9 edness inventory (Oldfield, 1971) and had no history of neurological or psychiatric illnesses.

10

11 **Behavioral procedures and task**

12 Subjects performed a continuous recognition task on equal numbers of initial and repeated
13 presentations of line drawings of common objects, which were pseudo-randomized within a session of
14 trials. The stimuli used in this experiment were the same as in Lehmann & Murray (2005). On each
15 trial, subjects indicated whether the visual stimulus was appearing for the first time or had appeared
16 previously. They were instructed to ignore the sounds and to focus on the visual information. The ex-
17 perimental paradigm is schematized in Figure 1.

18 Visual stimuli were comprised of 108 line drawings selected from either a standardized set
19 (Snodgrass & Vanderwart, 1980) or obtained from an online library (dgl.microsoft.com) and modified
20 to stylistically resemble those from the standardized set. Images were equally likely coming from a
21 multitude of semantic categories (e.g., animals, miscellaneous household items, musical instruments,
22 tools, vehicles, etc.) and their occurrence was balanced across experimental conditions. Images were
23 centrally presented for 500ms and appeared black on a white background. On initial presentations,
24 visual stimuli were subdivided into 3 groups: visual presentation only (V), which were appearing 50%
25 of the initial presentations; visual with semantically *congruent* sound (AVc), which were presented
26 25% of the initial presentations; and visual with semantically *incongruent* sound (AVi), which were
27 shown on the remaining 25% of initial presentations. In this way, the amount of unisensory and multi-
28 sensory initial presentations were equal.

1 Auditory stimuli were complex, meaningful sounds (16 bit stereo; 44100 Hz digitization; 500
2 ms duration) and were either semantically congruent (e.g. a “dong” sound with the image of a bell) or
3 semantically incongruent (e.g. a “woof” sound with the image of a gun) to one of the visual stimuli.
4 Sounds were obtained from an online library (dgl.microsoft.com) and modified using audio editing
5 software (Adobe Audition version 1.0) to be 500ms in duration. The volume was adjusted to a com-
6 fortable and comprehensible level for each subject, such that sounds were clearly, but not uncomforta-
7 bly, audible in the scanner environment.

8 On repeated presentations, only the visual stimuli from initial presentations were displayed.
9 Subjects’ task was to indicate as quickly and as accurately as possible, via a right-hand button press,
10 whether the image had been seen before. Thus, there were three classes of repeated presentations: (1)
11 initially presented as visual alone; (2) initially presented with a semantically *congruent* sound; and (3)
12 initially presented with a semantically *incongruent* sound. These conditions differ in (1) whether there
13 was a visual unisensory or an auditory-visual multisensory experience associated with the image, and
14 (2) whether this multisensory experience had semantically congruent or incongruent stimuli. To sim-
15 plify, we refer to repetitions of the images from the V condition as V-, to repetitions of images from
16 the AVc condition as V+c, and to repetitions of images from the AVi condition as V+i. Subjects were
17 not asked to judge if the semantic stimuli were incongruent or not, so the context (i.e., whether the
18 initial encounter with the image during the experiments was unisensory or multisensory) was com-
19 pletely orthogonal to the task.

20 Stimuli were presented for 500ms. The inter-trial interval (ITI) ranged from 5000ms to 12000ms
21 in steps of 1000ms, varying randomly from one trial to the next though evenly distributed within each
22 experimental condition to provide adequate temporal sampling of the blood oxygen level dependent
23 (BOLD) response. Stimulus delivery and the recording of behavioral data (reaction time and accuracy)
24 were controlled by E-prime in conjunction with their serial response box (www.pstnet.com; Psycholo-
25 gy Software Tools). Subjects were instructed to respond as fast as possible, whilst still being accurate.
26 Button presses longer than 1300ms after stimuli presentation were not recorded and the trial was la-
27 beled as a miss. Misses were considered as incorrect responses in the behavioral analysis.

1 Stimuli were presented in sessions of 64 trials, with equal likelihood of initial and repeated
2 presentations as well as balanced trials between initial unisensory and multisensory conditions (i.e., 16
3 trials of V-, 8 trials of V+c and 8 trials of V+i per session). Within each session, the conditions were
4 pseudo-randomized, and each image was repeated once, independently of how the image was initially
5 presented. The average number of trials between the initial and repeated presentation of any given
6 stimulus was 8.58 images (range = 3-20 images). Each subject completed 4 sessions. No object was
7 repeated more than once for any subject - that is, each experiment was comprised of distinct stimuli.
8 Likewise, sounds/images used for the incongruent condition were neither previously nor later used for
9 other conditions. There were four new multisensory events during each quarter of trials in each of the
10 four sessions, with an average of 10.25 new images during the first quarter and 7.0 new images during
11 the final quarter. Thus, neither experiment had a clear bias in the distribution of multisensory vs.
12 unisensory events nor in terms of old/new images that would readily explain performance differences
13 between images with multisensory vs. unisensory pasts.

14

15 **MRI Data acquisition**

16 Structural and functional images were collected on a 3 Tesla Siemens TrioTim scanner
17 equipped with an 8-channel head coil at the Center for Biomedical Imaging (CIBM) at the University
18 Hospital Lausanne (CHUV). A 3-dimensional high-resolution isotropic T1-weighted sequence provid-
19 ed 160 contiguous, anatomical slices (MPRAGE; TR/TE/flip angle = 1480ms/3.42ms/15°; 256 x
20 256mm in-plane resolution; a slice thickness of 1 mm; voxel size = 1x1x1 mm). Functional MRI im-
21 ages were continuously acquired using a standard gradient echo sequence (TR/TE/flip angle =
22 2007.5ms/30ms/90°) with 36 axial functional images (224x224mm in-plane resolution; 3 mm slice
23 thickness; 0.30 mm inter-slice gap; voxel size = 3.50 × 3.50 × 3.30mm) acquired in descending order
24 covering the whole brain. For each subject four separate sessions were recorded, each having 275 vol-
25 umes.

26 The raw data has been curated and organized according to the Brain Imaging Data Structure
27 (BIDS) standards (Gorgolewski et al., 2016) and is freely available via the OpenNeuro repository at
28 <https://openneuro.org/datasets/ds001345>.

1

2 **Behavioral analyses**

3 *Response accuracy* and *reaction time* were analyzed using the SPSS statistical software (IBM
4 Corporation, version 25). Two two-way repeated measures ANOVAs over the factors *Condition* (se-
5 mantically *congruent*, semantically *incongruent*, no semantic context) and *Presentation time* (*initial*
6 and *repeated* stimuli presentation) were conducted to identify main effects and interactions, once to
7 investigating response accuracy and once to investigate reaction time. When necessary, Greenhouse-
8 Geisser correction was applied when data violated assumptions of sphericity. Missed trials were con-
9 sidered as errors. Only correct responses were used for the investigation of reaction time.

10

11 **fMRI analysis**

12 Imaging data processing and analysis was carried out with fMRIfloWS v0.1
13 (<https://github.com/miykael/fmriflows>), a python-based neuroimaging toolbox that combines the fol-
14 lowing software packages: Nipype v1.1.9 (K. Gorgolewski et al., 2011), integrating algorithms of the
15 software packages FSL v5.0.9 (Smith et al., 2004), ANTs v2.2.0 (Avants et al., 2011), SPM12 v7219
16 (Penny, Friston, Ashburner, Kiebel, & Nichols, 2011), AFNI v18.0.5 (Cox & Hyde, 1997), Nilearn
17 v0.5 (Abraham et al., 2014), PyMVPA v2.6.5 (Hanke et al., 2009), AtlasReader v0.1 (Notter et al.,
18 2019), and Docker v18.09.2 (<https://docker.com>).

19 *Preprocessing*: The preprocessing of the functional images was done for each session individu-
20 ally and was based on FSL's FEAT workflow. As a first step, non-steady state volumes at the begin-
21 ning of each functional image were removed with Nipype and FSL and a brain extraction process was
22 applied to the images, using Nilearn. This was followed by motion correction using FSL, where all
23 images of a session were realigned to the mean of the session. Images were then slice-time corrected
24 relative to the 18th (i.e. middle) slice using FSL. The mean image of the slice-time corrected images
25 was then used to co-register the functional images to the subject specific anatomical image. This co-
26 registration was based on FSL's FEAT pipeline and uses a two-step co-registration. The first step uses
27 the anatomical image to pre-align the mean image, followed by the second co-registration step where
28 the white matter probabilistic image computed with SPM's segmentation function is used together

1 with the anatomical image in a boundary-based registration (BBR) approach to co-register the slice-
2 time corrected mean image to the anatomy. This co-registration is then applied to all functional images
3 of the session. In this step, images were also resampled to an isometric voxel resolution of $3.5 \times 3.5 \times$
4 3.5mm . Following this transformation, images were temporally filtered with a high pass filter of
5 100Hz , using AFNI. Finally, functional images were spatially smoothed with a smoothing kernel with
6 FWHM of 3.5mm and 7mm using Nilearn, to create the data for the multivariate and univariate analy-
7 sis respectively. In order to preserve small patterns of activation in the signal, a smaller smoothing
8 kernel of 3.5mm was applied for the multivariate analysis compared to the univariate analysis. Prior to
9 the spatial smoothing, functional images were used to compute confound regressors, such as
10 Framewise Displacement (FD), DVARS to identify motion-affected frames (Power, Barnes, Snyder,
11 Schlaggar, & Petersen, 2012), as well as average signal within the brain volume using Nilearn. Specific
12 volumes were labeled as outliers if their value was above 99%, or 3.27 standard deviation from the
13 whole time-series in either FD, DVARS or average brain volume signal. Outlier volumes were not
14 deleted from the preprocessed functional images, but each outlier volume was added as a nuisance
15 vector in the general linear model (GLM). On average, 2.64% of all volumes were identified as outli-
16 ers.

17 *Univariate analysis:* The functional images used for the univariate analysis were the ones with a
18 smoothing kernel of FWHM = 7.0mm . Individual statistics were based on a least-square estimation
19 using the GLM for serially autocorrelated observations and were performed separately on each voxel
20 in the individual participant's space (Friston et al., 1995) with SPM12. Six covariates of interest were
21 calculated, representing the stimulus onset of the six experimental conditions AVc, AVi, V, V+c, V+i
22 and V-. Covariates were convolved with the canonical hemodynamic response function. For each co-
23 variate of interest, a temporal derivative was added to the GLM. The serial autocorrelation of the
24 BOLD time series was modeled using a first-order autoregressive model. The four experimental runs
25 were treated as separate sessions. No global intensity normalization was applied. The following
26 regressors of no interest were added to the statistical model: six realignment-parameters (3 rotational
27 and 3 translational), one regressor for FD and one for DVARS to identify motion-affected frames
28 (Power et al., 2012), and, finally, one regressor containing the average activation in the whole brain. A

1 temporal high-pass filter of 100s was applied to remove low-frequency drifts over a timescale longer
2 than this cutoff. For the random-effects analysis, contrasts of interest on the individual level were cal-
3 culated and normalized to the ICBM 2009c Nonlinear Asymmetric template (Fonov et al., 2011), us-
4 ing ANTs. Spatial resolution after the normalization was $1 \times 1 \times 1 \text{mm}^3$. One-sample t-tests were per-
5 formed on the second level. The analysis was restricted to the gray matter voxels by the use of a binary
6 mask. This mask was obtained by thresholding the gray matter probability map of the ICBM 2009c
7 Nonlinear Asymmetric template at 10%, meaning that voxels were included in the mask if they had a
8 chance higher than 10% to be classified as gray matter. T-maps were thresholded on the voxel level at
9 $p < 0.001$, and topological False Discovery Rate (FDR)-correction for multiple comparisons, as im-
10 plemented in Nipype, was applied on the whole-brain activation cluster-extent level at $p < 0.05$.

11 *Multivariate pattern analysis (MVPA) preparation:* The functional images used for the multi-
12 variate analysis had a smoothing kernel of FWHM = 3.5 mm. In order to preserve small patterns of
13 activation in the signal, a smaller smoothing kernel was applied to the data for the multivariate analy-
14 sis than the univariate analysis. The multivariate analysis was performed within each subject. The
15 individual statistics were performed in the same way as for the univariate analysis (i.e. same setup of
16 GLM, usage of motion regressors and specification of outliers). The only differences were the reduced
17 smoothing kernel and that the β -maps were normalized to an isometric voxel resolution of
18 $3.5 \times 3.5 \times 3.5 \text{mm}^3$. To prepare for the multivariate analysis, β -maps from the GLM were normalized by
19 voxel-wise z-scoring per run across all six conditions and sessions. The classification was based on 16
20 β -maps (4 conditions x 4 sessions), as the two unisensory conditions V and V- were not used for the
21 MVPA classification. A binary mask was created to selectively include voxels in the searchlight anal-
22 ysis that have a value in at least one of the β -maps. This binary mask was then dilated by two voxels
23 and eroded by one voxel in order to ensure that the mask does not include single voxel “islets”. For the
24 classification analysis, a linear support vector machine (SVM) from the LIBSVM package
25 (<https://www.csie.ntu.edu.tw/~cjlin/libsvm>) as implemented in PyMVPA v2.6.5 (Hanke et al., 2009)
26 was used. A four-fold leave-one-run-out cross-validation (i.e. training on the beta maps of three ses-
27 sions and testing on the beta maps of the remaining session) was performed (Pereira, Mitchell, &

1 Botvinick, 2009). The SVM Nu hyper-parameter was automatically scaled according to the data norm
2 for each training data fold.

3 *Searchlight analysis:* This type of analysis was performed due to its sensitivity to functional
4 brain organization, and because the searchlight analysis is an approach to localize brain regions with
5 high discriminative information (Kriegeskorte, Goebel, & Bandettini, 2006). A binary classifier was
6 trained to discriminate between semantically congruent and incongruent stimuli, during the encoding
7 as well as during the retrieval phase (see Figure 2). To detect reoccurring and discriminative activation
8 patterns during encoding and retrieval, the classifier was once trained on “AVc vs. AVi” and tested on
9 “V+c vs. V+i”, and once trained on “V+c vs V+i” and tested on “AVc vs AVi”. The two searchlight
10 output maps were then averaged to receive a new accuracy map. With this approach, we can distin-
11 guish brain regions where the stored information is identical during encoding and retrieval, from brain
12 regions that are discriminative between the two semantic conditions.

13 The searchlight analysis was performed across the brain in template space and for each subject
14 and classification, separately. Every voxel in the brain mask was taken as a center point for a sphere
15 with a radius of 3.5 voxels (i.e. 12.25mm radius, up to 123 voxels). For each of those spheres, a four-
16 fold leave-one-out cross-validation was performed, resulting in one accuracy value per sphere. The
17 classification accuracy of a given voxel was the average accuracy of all spheres that included this
18 voxel. The resulting subject specific accuracy maps per classification pair were averaged over all par-
19 ticipants to obtain the group classification accuracy map. Significance testing in the searchlight analy-
20 sis was performed by using permutation and bootstrap sampling methods, followed by cluster
21 thresholding with correction for multiple comparisons as suggested by Stelzer, Chen, & Turner (2013)
22 and implemented by PyMVPA. The permutation was achieved by conducting the previously men-
23 tioned searchlight for each subject an additional 99 times, each time with randomly permuted data
24 labels resulting in one condition relevant accuracy map and 99 “noise” accuracy maps per subject. For
25 the bootstrapping process, new group-level “noise” accuracies maps were generated by selecting ran-
26 domly one of the 100 accuracy maps per subject and averaging them into a new group-level chance
27 accuracy map. This process was repeated 100’000 times to create a voxel-wise null distribution of the

1 classification accuracy map on the group level. All volumes were thresholded with a voxel-wise
2 threshold of $p < 0.002$. Only clusters with size bigger than 214mm^3 were retained.

3 *Visualization & cluster information:* Extraction of cluster information of the univariate and mul-
4 tivariate analysis were obtained with the freely available AtlasReader (Notter et al., 2019), which de-
5 pends on Nilearn (v0.5.0; nilearn.github.io; Abraham et al., 2014). Only clusters bigger than 100mm^3
6 and 214mm^3 were retained for the univariate and multivariate analysis, respectively. This represents a
7 minimum cluster size of 100 voxels for the univariate and 5 voxels for the multivariate analysis.

8

9 **Results**

10 **Behavior**

11 Participants performed at near-ceiling levels, with an average accuracy rate across conditions
12 above 95%. Mean (\pm standard error) accuracy and reaction time for each condition are shown in Fig-
13 ure 3. The two-way repeated measures ANOVA investigating accuracy showed a significant main
14 effect of *Condition* ($F_{(2, 22)} = 5.36$; $p = 0.013$; $\eta_p^2 = 0.33$), but none for *Presentation Time* ($F_{(1, 11)} = 1.71$;
15 $p = 0.22$; $\eta_p^2 = 0.13$), nor their *interaction* ($F_{(1, 27, 13.98)} = 2.27$; $p = 0.15$; $\eta_p^2 = 0.17$). The main effect of
16 condition was due to generally lower performance when no multisensory stimulus was or had been
17 presented (93.8%), independently of semantic congruence (95.9%; $t_{(11)} = 2.9$; $p = 0.021$) or incongruence
18 (96.5%; $t_{(11)} = 3.22$; $p = 0.008$). There was no evidence for a general accuracy difference between condi-
19 tions that were or had been presented in a semantically congruent vs. semantically incongruent multi-
20 sensory context ($p = 0.56$).

21 The two-way repeated measures ANOVA investigating *reaction time* showed a significant ef-
22 fect of *Condition* ($F_{(2, 22)} = 3.53$; $p = 0.047$; $\eta_p^2 = 0.24$) and of the interaction between *Condition* and
23 *Presentation Time* ($F_{(2, 22)} = 4.10$; $p = 0.031$; $\eta_p^2 = 0.27$), but none for *Presentation Time* alone ($F_{(1, 11)} =$
24 0.620 ; $p = 0.45$; $\eta_p^2 = 0.05$). Given this interaction, separate ANOVAs were computed for initial and
25 repeated presentation times. For initial presentations, there was a main effect of condition ($F_{(2, 22)} = 4.95$;
26 $p = 0.017$; $\eta_p^2 = 0.31$) that was due to the generally faster reaction times to visual than either multisensory

1 ry condition, independent of semantic congruency. For repeated presentations, there was no main ef-
2 fect of condition ($F(2,22) < 1$; $p > 0.65$; $\eta_p^2 = 0.04$).

3 **Univariate fMRI analysis**

4 The univariate analysis investigated the following four contrasts: “AVc vs. AVi”, “V+c vs.
5 V+i”, “V+c vs. V-” and “V+i vs. V-”. Significant results are shown in Figure 4. The contrast “AVc vs.
6 AVi” highlighted that the first presentation of a congruent multisensory stimuli induced a lower acti-
7 vation in the left occipital pole and LOC, and in the right superior temporal gyrus and planum
8 temporale, in contrast to an incongruent multisensory stimuli (Figure 4A, Table 1). Comparing the
9 second presentation of a visual stimulus, if the first presentation was paired with a congruent sound in
10 contrast to no sound showed an increased activation in the left LOC and a decreased activation in the
11 left precentral gyrus and the right middle frontal gyrus (Figure 4B, Table 1). No statistically signifi-
12 cant group activation differences were observed when comparing “V+c vs. V+i” or “V+i vs. V-”, when
13 the voxel threshold was set to $p < 0.001$. Increasing the voxel threshold to $p < 0.005$, while keeping the
14 cluster threshold at $p < 0.05$ (FDR) reveals bilateral increased activation in the LOC during the percep-
15 tion of V+c in comparison to V+i (Figure 4C, Table 1). No group differences were observed in the
16 contrast “V+i vs. V-”, even at a voxel threshold of $p < 0.01$.

17

18 **Multivariate fMRI analysis**

19 To identify brain regions that encode and retrieve semantically congruent multisensory stimuli,
20 a searchlight analysis was performed using SVM Nu classifiers, that investigated the binary classifica-
21 tions “AVc vs. AVi” and “V+c vs. V+i” (i.e. a cross-classification analysis). The clusters depicted in
22 Figure 5 show areas where a classifier trained on one contrast (e.g. “AVc vs. AVi”) could predict the
23 class differences in the other contrast (e.g. “V+c vs. Vi”) and vice versa. This represents brain regions
24 where the activation pattern difference between congruent and incongruent multisensory stimuli dur-
25 ing the encoding phase was informative for the discrimination of repeated visual stimuli differing only
26 in whether they had been previously encountered with a semantically congruent or incongruent sound.

1 Significantly increased prediction accuracy was observed in the left inferior frontal gyrus, frontal pole
2 and occipital pole, and in the right superior frontal gyrus (see Figure 5, Table 2).

3

4

1 **Discussion**

2 We show, for the first time, that decoding of semantically congruent from incongruent multisensory information not only occurs within low-level visual cortices, but is also effective for the decoding
3 of responses to unisensory, visual repetitions of these contexts several trials later. This decoding appears to be independent from subjects explicitly attending to the multisensory context or its semantic
4 congruence. As a reminder, the task here required the discrimination of initial from repeated presentation of visual objects, with the auditory information being task-irrelevant. One possibility for why
5 these effects were observed here within visual cortices might be due to the visually-based task. Future studies should therefore explore stimuli paradigms with a task based on other or multiple modalities.
6 Our results cannot be explained by behavioural differences across conditions, as there were no reliable accuracy differences between the AVc and AVi conditions nor between the V+c and V+I conditions.
7 In contrast to prior works, where stimuli were presented repeatedly across trials and/or for a relatively long time, participants' exposure to any multisensory stimulus in this study was limited to a single
8 500ms trial. This short, single-trial exposure was sufficient to engender distinct representations both during encoding and later unisensory-based retrieval. The fact that there was overlap in representations
9 of semantic congruence (i.e. the decoding of the AVc from AVi condition) and representations of prior multisensory context (i.e. the decoding of V+c from V+i) is consistent with the notion of
10 redintegration (Hamilton, 1859) that postulates that a component part suffices to activate a consolidated memory representation (cf. Thelen and Murray, 2013 for discussion).
11

12 We identified brain networks and specific regions involved in the encoding and retrieval of multisensory memories during a continuous recognition task, where visual stimuli were presented either
13 alone or with semantically congruent or incongruent auditory stimuli. Our behavioral results highlighted that multisensory pairing during the initial presentation improved memory performance by
14 increasing accuracy. These behavioral results were linked to the activation of several brain regions involved differently in the encoding and retrieval of the multisensory memories. Using an univariate
15 analysis, our results indicate different involvement of brain networks in the left occipital pole, the right superior temporal gyrus and the left LOC, while multivariate analysis revealed that semantically con-
16

1 gruent multisensory information is encoded and retrieved in specific brain regions, such as the left
2 inferior frontal gyrus, left frontal and occipital poles and the right superior frontal gyrus.

3 Our univariate results showed significant differences in the primary auditory and visual cortices
4 for the comparison of the initial presentation of multisensory stimuli in a congruent vs. an incongruent
5 context (see Figure 4A), which were consistent with findings of previous studies (Beauchamp, Lee,
6 Argall, & Martin, 2004; Lu et al., 2018; Martuzzi et al., 2007; Murray, Lewkowicz, Amedi, &
7 Wallace, 2016; van Atteveldt et al., 2004) and supported their involvement during functional integra-
8 tion. Additional to the primary cortices, we also observed an involvement of the LOC. This latter acti-
9 vation was not surprising as many studies have already shown the involvement of the LOC in line with
10 object representation and multisensory integration (Lacey, Tal, Amedi, & Sathian, 2009; Matusz et al.,
11 2017).

12 Memory traces of natural and well-established objects, which were different from those of new-
13 ly established modality pairings, seem to require a different memory strategy to be reactivated at a
14 later stage. This was highlighted by the univariate analysis shown in Figure 4B. Participants showed
15 an increased activation during V+c in the left LOC, which has been shown to process episodic
16 memory, memory encoding and retrieval (Hassabis, Kumaran, & Maguire, 2007; Qin et al., 2007;
17 Ranganath, Johnson, & D'Esposito, 2003), while the increased activation during V- in the right middle
18 frontal gyrus is known to be involved in working memory tasks (Sabb, Bilder, Chou, & Bookheimer,
19 2007; Yeh, Kuo, & Liu, 2007). The lack of significant differences between “V+i and V-“, even at a
20 lenient voxel threshold of $p < 0.01$, might indicate that the brain processes during the perception of both
21 conditions were too similar.

22 Using a multivariate approach, we investigated if the representational content in a given area
23 differed between conditions, even if the overall activation within this region was the same (Mur,
24 Bandettini, & Kriegeskorte, 2009). This multivariate analysis revealed that the primary visual cortex
25 was not just involved during the encoding of semantic congruent stimuli, but also during the retrieval
26 of memory traces thereof. While our findings did not provide information about the direction of the
27 involvement of a region, they matched previous reports by Doehrmann & Naumer (2008), who pro-

1 posed that the STC was more sensitive to semantically congruent stimuli, while the inferior frontal
2 gyrus was more responsive to semantically incongruent stimuli, and that those effect were more often
3 lateralized in the left hemisphere. Several studies have shown that the inferior frontal cortex was
4 linked to semantic memory retrieval (Chou, Chen, Wu, & Booth, 2009; Forgács et al., 2012; Raposo,
5 Mendes, & Marques, 2012), which supports the hypothesis that semantically congruent memory trac-
6 es, in contrast to incongruent ones, require episodic memory to be stored and reactivated.

7 Our study introduced a novel approach to investigate the differences in brain regions for the en-
8 coding and retrieval of the representations of multisensory induced and semantic context dependent
9 objects, by looking at reoccurring activation patterns of specific memory traces. While compelling for
10 the specific modality pairing, these results should be tested with other paradigms before being further
11 generalized. Future studies should therefore investigate if the location of stable activation pattern dur-
12 ing encoding and retrieval of specific memory traces might take part in the same brain regions, and/or
13 if the modality pairings might be different from the ones used in this study. We might expect that simi-
14 lar, but slightly weaker results will be observed if the repeated stimuli modality was auditory instead
15 of visual (as used in Noppeney et al., 2008). Another way to test the generalization of these multisen-
16 sory integration regions will be to pair different sensorial modalities, such as touch or olfactory, or
17 move away from well-established natural objects and investigate the encoding and decoding of new
18 memory traces by creating newly primed object pairings, such as used in Noppeney et al. (2008). Such
19 a paradigm would clarify our results, and their explicit causality with well-established natural object
20 representation and episodic memory, but also if the same networks are involved in newly created
21 memory traces.

22 Previous work has demonstrated that categories of sounds either presented in a unisensory man-
23 ner or alternatively imagined in response to a heard word can be successfully decoded by patterned
24 responses within low-level visual cortices, including V1 (Vetter, Smith, & Muckli, 2014). The impli-
25 cation is that low-level visual cortices represent not only the occurrence of multisensory events (i.e.
26 that there is something multisensory in one's real or imagined environment), but also the quality of
27 these events in terms of semantic categories. Our results provide an important extension to this claim.
28 First, we showed that it is possible to decode the semantic congruence itself. This suggests that visual

1 cortices represent qualitative information about experiences; namely whether or not information from
2 different senses refer to the same object. Second and extending beyond the above, we furthermore
3 show that the visual cortex can decode unisensory visual information according to how it had been
4 previously encountered several trials earlier. Visual repetitions of images previously encountered in a
5 semantically congruent or incongruent context could be reliably decoded by primary visual cortex.

6

7

1 Acknowledgements

2 The computation time of the searchlight analysis with permutation testing is very time consum-
3 ing. The total computation time for the 1'200 accuracy maps (12 subjects x 100 permutations) took
4 approximately 900 hours. We want to thank Michael Hanke and the Psychoinformatics lab at the Otto
5 von Guericke University in Magdeburg for allowing us to use their computation cluster Medusa,
6 which allowed us to finish the complete computation in less than a day. We also want to thank
7 Raphaël V. Meylan for his initial involvement in this study and recording of the data.

8

9 Funding

10 This work was supported by Swiss National Science Foundation (grant numbers 169206,
11 149982, 133136, 118419, and 105680 to M.M.M.). Infrastructural support for the Centre d'Imagerie
12 BioMédicale (CIBM) was provided by the University and University Hospital Center of Lausanne, the
13 University and University Hospital Center of Geneva, and the Swiss Federal Institute of Technology
14 of Lausanne, as well as the Leenards and Jeantet Foundations.

15

1 **Tables**

2

3 **Table 1.** Cluster information corresponding to the univariate fMRI results shown in Figure 4.

MNI coordinate of peak in (x,y,z)	Mean t- value	Cluster extent (in mm ³)	Cluster extent according to Harvard-Oxford probabilistic atlas
A) AVc vs. AVi – voxel threshold at $p < 0.001$ and cluster threshold at $p < 0.05$ (FDR)			
-16, -95, 10	-5.16	1283	100.00% Left_Occipital_Pole
68, -28, 9	-4.76	1100	74.00% Right_Superior_Temporal_Gyrus_posterior_division 17.18% Right_Planum_Temporale 5.73% Right_Postcentral_Gyrus
-32, -89, 14	-5.04	795	84.53% Left_Lateral_Occipital_Cortex_superior_division 12.96% Left_Occipital_Pole
B) V+c vs. V- – voxel threshold at $p < 0.001$ and cluster threshold at $p < 0.05$ (FDR)			
-33, -79, 37	5.06	1826	99.07% Left_Lateral_Occipital_Cortex_superior_division
-36, -11, 66	-4.56	953	95.91% Left_Precentral_Gyrus
39, 6, 54	-5.15	488	100.00% Right_Middle_Frontal_Gyrus
C) V+c vs. V+i – voxel threshold at $p < 0.005$ and cluster threshold at $p < 0.05$ (FDR)			
42, -70, 42	3.77	1884	75.58% Right_Lateral_Occipital_Cortex_superior_division; 16.35% Right_Angular_Gyrus
47, -66, 20	3.89	1787	53.44% Right_Lateral_Occipital_Cortex_superior_division; 29.66% Right_Angular_Gyrus 8.90% Right_Lateral_Occipital_Cortex_inferior_division 8.00% Right_Middle_Temporal_Gyrus_temporooccipital_part
-36, -61, 39	3.58	1711	91.88% Left_Lateral_Occipital_Cortex_superior_division; 6.78% Left_Angular_Gyrus

4

5

6 **Table 2.** Cluster information corresponding to the multivariate fMRI results shown in Figure 5.

MNI coordinate of peak in (x,y,z)	Mean accuracy	Cluster extent (in mm ³)	Cluster extent according to Harvard-Oxford probabilistic atlas
Trained on AVc vs. AVi - tested on V+c vs. V+i – voxel threshold at $p < 0.002$ and cluster extent at 214mm³			
-61, 22, 6	0.590	2187	41.2% Left_Inferior_Frontal_Gyrus_pars_opercularis 17.7% Left_Precentral_Gyrus 17.7% Left_Inferior_Frontal_Gyrus_pars_triangularis 15.7% Left_Temporal_Pole
-29.5, 50, 41	0.594	2058	85.4% Left_Frontal_Pole; 14.58% no_label
16, 11.5, 72.5	0.599	1501	100.0% Right_Superior_Frontal_Gyrus
-12, -107.5, 6	0.599	1415	100.0% Left_Occipital_Pole

7

8

1 **Figure Legends**

2

3 **Figure 1.** Experimental paradigm: The recognition task comprised of initial and repeated image
4 presentations. 50% of the initial presentations were only visually (V), 25% of initial presentations
5 were simultaneously appearing with semantically congruent sounds (AVc) and 25% of initial presenta-
6 tions were simultaneously appearing with semantically incongruent sounds (AVi). All repeated
7 presentations contained only images but were subdivided for analyses into three conditions - those that
8 initially had been presented with congruent sounds (V+c trials), with incongruent sounds (V+i trials)
9 or only visually (V- trials).

10

11 **Figure 2.** Illustration of the searchlight classification. Fictitious brain activation is shown during the
12 perception of semantically congruent (*top*; i.e. AVc and V+c) and incongruent (*bottom*; i.e. AVi and
13 V+i) stimuli, during the initial encoding (*left*; yellow box) and subsequent retrieval (*right*; green box)
14 phase. Red-white-blue colors indicate different relative activation levels within a voxel. In black out-
15 line, the searchlight sphere in 2D is shown, with a radius of 3.5 voxels around the center (outlined in
16 green). Each classifier will use the information from all voxels within a searchlight sphere to try to
17 discriminate between the two semantic conditions, while looking for a stable activation pattern during
18 encoding and retrieval.

19

20 **Figure 3.** Behavioral results for accuracy (*left*) and reaction time (*right*). Group averages for the initial
21 stimuli presentation for the condition AVc, AVi and V are shown in *blue* and group averages for the
22 repeated stimuli presentation for the condition V+c, V+i and V- are shown in *orange*. Error bars repre-
23 sent standard errors.

24

25 **Figure 4.** Univariate fMRI results showing significant difference between (A) the initial presentation
26 of congruent and incongruent multisensory stimuli; (B) the repeated presentation of visual stimuli if
27 they were previously seen with congruent sounds or no sounds and (C) the repeated presentation of
28 visual stimuli if they were previously seen with congruent or incongruent sounds. Colors represent
29 stronger (red-yellow) or lower (blue-white) activation during the first comparison object, i.e. AVc or
30 V+c. Shown values represent t-values from the group statistics.

31

32 **Figure 5.** Multivariate fMRI results showing regions with significant prediction accuracy of a classifi-
33 er which looked for stable activation pattern between the conditions AVc and V+c, but also differed
34 significantly from the stable activation pattern between the conditions AVi and V+i. Shown values
35 represent prediction accuracies.

36

37

1 **References**

- 2 Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... Varoquaux, G.
3 (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*,
4 8(February), 14. <https://doi.org/10.3389/fninf.2014.00014>
- 5 Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible
6 evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*,
7 54(3), 2033–2044. <https://doi.org/10.1016/j.neuroimage.2010.09.025>
- 8 Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual
9 information about objects in superior temporal sulcus. *Neuron*, 41(5), 809–823.
10 [https://doi.org/10.1016/S0896-6273\(04\)00070-4](https://doi.org/10.1016/S0896-6273(04)00070-4)
- 11 Calvert, G. A., Spence, C., & others. (2004). *The Handbook of multisensory processes*, MIT Press.,
12 Cambridge, MA.
- 13 Chou, T.-L., Chen, C.-W., Wu, M.-Y., & Booth, J. R. (2009). The role of inferior frontal gyrus and
14 inferior parietal lobule in semantic processing of Chinese characters. *Experimental Brain*
15 *Research*, 198(4), 465–475. <https://doi.org/10.1007/s00221-009-1942-y>
- 16 Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in*
17 *Biomedicine*, 10(4–5), 171–178. [https://doi.org/10.1002/\(SICI\)1099-1492\(199706/08\)10:4/5<171::AID-NBM453>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L)
- 19 Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: how meaning
20 modulates processes of audio-visual integration. *Brain Research*, 1242, 136–150.
21 <https://doi.org/10.1016/j.brainres.2008.03.071>
- 22 Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., & Collins, D. L. (2011).
23 Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1), 313–327.
24 <https://doi.org/10.1016/j.neuroimage.2010.07.033>
- 25 Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M. J., Pléh, C., & Jacobs, A. M. (2012). Neural
26 correlates of combinatorial semantic processing of literal and figurative noun noun compound
27 words. *NeuroImage*, 63(3), 1432–1442. <https://doi.org/10.1016/j.neuroimage.2012.07.029>
- 28 Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., & Turner,

- 1 R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, 2(1), 45–53.
2 <https://doi.org/10.1006/nimg.1995.1007>
- 3 Gibson, J. R., & Maunsell, J. H. (1997). Sensory modality specificity of neural activity related to
4 memory in visual cortex. *Journal of Neurophysiology*, 78(3), 1263–1275.
- 5 Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh,
6 S. S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing
7 framework in python. *Frontiers in Neuroinformatics*, 5(August), 13.
8 <https://doi.org/10.3389/fninf.2011.00013>
- 9 Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., ... Poldrack, R. A.
10 (2016). The brain imaging data structure, a format for organizing and describing outputs of
11 neuroimaging experiments. *Scientific Data*, 3, 160044. <https://doi.org/10.1038/sdata.2016.44>
- 12 Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009).
13 PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*,
14 7(1), 37–53. <https://doi.org/10.1007/s12021-008-9041-y>
- 15 Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis
16 of episodic memory. *The Journal of Neuroscience : The Official Journal of the Society for*
17 *Neuroscience*, 27(52), 14365–14374. <https://doi.org/10.1523/JNEUROSCI.4549-07.2007>
- 18 Jung, Y., Larsen, B., & Walther, D. B. (2018). Modality-Independent Coding of Scene Categories in
19 Prefrontal Cortex. *The Journal of Neuroscience : The Official Journal of the Society for*
20 *Neuroscience*, 38(26), 5969–5981. <https://doi.org/10.1523/JNEUROSCI.0272-18.2018>
- 21 Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping.
22 *Proceedings of the National Academy of Sciences of the United States of America*, 103(10),
23 3863–3868. <https://doi.org/10.1073/pnas.0600244103>
- 24 Lacey, S., Tal, N., Amedi, A., & Sathian, K. (2009). A putative model of multisensory object
25 representation. *Brain Topography*, 21(3–4), 269–274. [https://doi.org/10.1007/s10548-009-0087-](https://doi.org/10.1007/s10548-009-0087-4)
26 4
- 27 Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object
28 discrimination. *Brain Research. Cognitive Brain Research*, 24(2), 326–334.

- 1 <https://doi.org/10.1016/j.cogbrainres.2005.02.005>
- 2 Lu, L., Zhang, G., Xu, J., & Liu, B. (2018). Semantically Congruent Sounds Facilitate the Decoding of
3 Degraded Images. *Neuroscience*, 377, 12–25. <https://doi.org/10.1016/j.neuroscience.2018.01.051>
- 4 Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J.-P., Maeder, P. P., Clarke, S., & Meuli, R. A.
5 (2007). Multisensory Interactions within Human Primary Cortices Revealed by BOLD
6 Dynamics. *Cerebral Cortex*, 17(7), 1672–1679. <https://doi.org/10.1093/cercor/bhl077>
- 7 Matusz, P. J., Thelen, A., Amrein, S., Geiser, E., Anken, J., & Murray, M. M. (2015). The role of
8 auditory cortices in the retrieval of single-trial auditory-visual object memories. *European*
9 *Journal of Neuroscience*, 41(5), 699–708. <https://doi.org/10.1111/ejn.12804>
- 10 Matusz, P. J., Wallace, M. T., & Murray, M. M. (2017). A multisensory perspective on object
11 memory. *Neuropsychologia*, (April), 1–10.
12 <https://doi.org/10.1016/j.neuropsychologia.2017.04.008>
- 13 Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object
14 recognition in humans: a high-density electrical mapping study. *Cerebral Cortex (New York,*
15 *N.Y. : 1991)*, 14(4), 452–465. <https://doi.org/10.1093/cercor/bhh007>
- 16 Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with
17 pattern-information fMRI--an introductory guide. *Social Cognitive and Affective Neuroscience*,
18 4(1), 101–109. <https://doi.org/10.1093/scan/nsn044>
- 19 Murray, M. M., Foxe, J. J., & Wylie, G. R. (2005). The brain uses single-trial multisensory memories
20 to discriminate without awareness. *NeuroImage*, 27(2), 473–478.
21 <https://doi.org/10.1016/j.neuroimage.2005.04.016>
- 22 Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory Processes: A
23 Balancing Act across the Lifespan. *Trends in Neurosciences*, 39(8), 567–579.
24 <https://doi.org/10.1016/j.tins.2016.05.003>
- 25 Murray, M. M., Michel, C. M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., &
26 Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by
27 electrical neuroimaging. *NeuroImage*, 21(1), 125–135.
28 <https://doi.org/10.1016/j.neuroimage.2003.09.035>

- 1 Murray, M. M., & Wallace, M. T. (2012). *The Neural Bases of Multisensory Processes*. (M. M.
2 Murray & M. T. Wallace, Eds.). Boca Raton, FL: CRC Press.
- 3 Noppeney, U., Josephs, O., Hocking, J., Price, C. J., & Friston, K. J. (2008). The effect of prior visual
4 information on recognition of speech and sounds. *Cerebral Cortex (New York, N.Y. : 1991)*,
5 *18*(3), 598–609. <https://doi.org/10.1093/cercor/bhm091>
- 6 Notter, M., Gale, D., Herholz, P., Markello, R., Notter-Bielser, M.-L., & Whitaker, K. (2019).
7 AtlasReader: A Python package to generate coordinate tables, region labels, and informative
8 figures from statistical MRI images. *Journal of Open Source Software*, *4*(34), 1257.
9 <https://doi.org/10.21105/joss.01257>
- 10 Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain
11 activity during memory retrieval. *Proceedings of the National Academy of Sciences of the United*
12 *States of America*, *97*(20), 11120–11124. Retrieved from
13 <http://www.pnas.org/cgi/doi/10.1073/pnas.97.20.11120>
- 14 Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory.
15 *Neuropsychologia*, *9*(1), 97–113. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/5146491>
- 16 Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical*
17 *parametric mapping: the analysis of functional brain images*. Elsevier.
- 18 Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial
19 overview. *NeuroImage*, *45*(1 Suppl), S199-209.
20 <https://doi.org/10.1016/j.neuroimage.2008.11.007>
- 21 Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but
22 systematic correlations in functional connectivity MRI networks arise from subject motion.
23 *NeuroImage*, *59*(3), 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>
- 24 Qin, S., Piekema, C., Petersson, K. M., Han, B., Luo, J., & Fernández, G. (2007). Probing the
25 transformation of discontinuous associations into episodic memory: an event-related fMRI study.
26 *NeuroImage*, *38*(1), 212–222. <https://doi.org/10.1016/j.neuroimage.2007.07.020>
- 27 Ranganath, C., Johnson, M. K., & D'Esposito, M. (2003). Prefrontal activity associated with working
28 memory and episodic long-term memory. *Neuropsychologia*, *41*(3), 378–389.

- 1 [https://doi.org/10.1016/S0028-3932\(02\)00169-0](https://doi.org/10.1016/S0028-3932(02)00169-0)
- 2 Raposo, A., Mendes, M., & Marques, J. F. (2012). The hierarchical organization of semantic memory:
3 executive function in the processing of superordinate concepts. *NeuroImage*, *59*(2), 1870–1878.
4 <https://doi.org/10.1016/j.neuroimage.2011.08.072>
- 5 Sabb, F. W., Bilder, R. M., Chou, M., & Bookheimer, S. Y. (2007). Working memory effects on
6 semantic processing: priming differences in pars orbitalis. *NeuroImage*, *37*(1), 311–322.
7 <https://doi.org/10.1016/j.neuroimage.2007.04.050>
- 8 Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H.,
9 ... Matthews, P. M. (2004). Advances in functional and structural MR image analysis and
10 implementation as FSL. *NeuroImage*, *23 Suppl 1*(SUPPL. 1), S208-19.
11 <https://doi.org/10.1016/j.neuroimage.2004.07.051>
- 12 Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name
13 agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental*
14 *Psychology: Human Learning & Memory*, *6*(2), 174–215. [https://doi.org/10.1037/0278-](https://doi.org/10.1037/0278-7393.6.2.174)
15 [7393.6.2.174](https://doi.org/10.1037/0278-7393.6.2.174)
- 16 Sperdin, H. F., Cappe, C., Foxe, J. J., & Murray, M. M. (2009). Early, low-level auditory-
17 somatosensory multisensory interactions impact reaction time speed. *Frontiers in Integrative*
18 *Neuroscience*, *3*(1), 2. <https://doi.org/10.3389/neuro.07.002.2009>
- 19 Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in
20 classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size
21 control. *NeuroImage*, *65*, 69–82. <https://doi.org/10.1016/j.neuroimage.2012.09.063>
- 22 Thelen, A., Cappe, C., & Murray, M. M. (2012). Electrical neuroimaging of memory discrimination
23 based on single-trial multisensory learning. *NeuroImage*, *62*(3), 1478–1488.
24 <https://doi.org/10.1016/j.neuroimage.2012.05.027>
- 25 Thelen, A., & Murray, M. M. (2013). The efficacy of single-trial multisensory memories.
26 *Multisensory Research*, *26*(5), 483–502. <https://doi.org/10.1163/22134808-00002426>
- 27 Thelen, A., Talsma, D., & Murray, M. M. (2015). Single-trial multisensory memories affect later
28 auditory and visual object discrimination. *Cognition*, *138*, 148–160.

- 1 <https://doi.org/10.1016/j.cognition.2015.02.003>
- 2 Thomas, R. L., Nardini, M., & Mareschal, D. (2017). The impact of semantically congruent and
3 incongruent visual information on auditory object recognition across development. *Journal of*
4 *Experimental Child Psychology*, *162*, 72–88. <https://doi.org/10.1016/j.jecp.2017.04.020>
- 5 van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech
6 sounds in the human brain. *Neuron*, *43*(2), 271–282.
7 <https://doi.org/10.1016/j.neuron.2004.06.025>
- 8 Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding sound and imagery content in early visual
9 cortex. *Current Biology : CB*, *24*(11), 1256–1262. <https://doi.org/10.1016/j.cub.2014.04.020>
- 10 Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory’s echo: vivid remembering
11 reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences of the*
12 *United States of America*, *97*(20), 11125–11129. Retrieved from
13 <http://www.ncbi.nlm.nih.gov/pubmed/11005879>
- 14 WMA General Assembly. (2008). World Medical Association Declaration of Helsinki: Ethical
15 Principles for Medical Research Involving Human Subjects (as amended by the 59th WMA
16 General Assembly, Seoul, October 2008). *World Medical Association*.
- 17 Xie, Y., Xu, Y., Bian, C., & Li, M. (2017). Semantic congruent audiovisual integration during the
18 encoding stage of working memory: An ERP and sLORETA study. *Scientific Reports*, *7*(1), 1–
19 10. <https://doi.org/10.1038/s41598-017-05471-1>
- 20 Yeh, Y.-Y., Kuo, B.-C., & Liu, H.-L. (2007). The neural correlates of attention orienting in
21 visuospatial working memory for detecting feature and conjunction changes. *Brain Research*,
22 *1130*(1), 146–157. <https://doi.org/10.1016/j.brainres.2006.10.065>
- 23

Figure 1

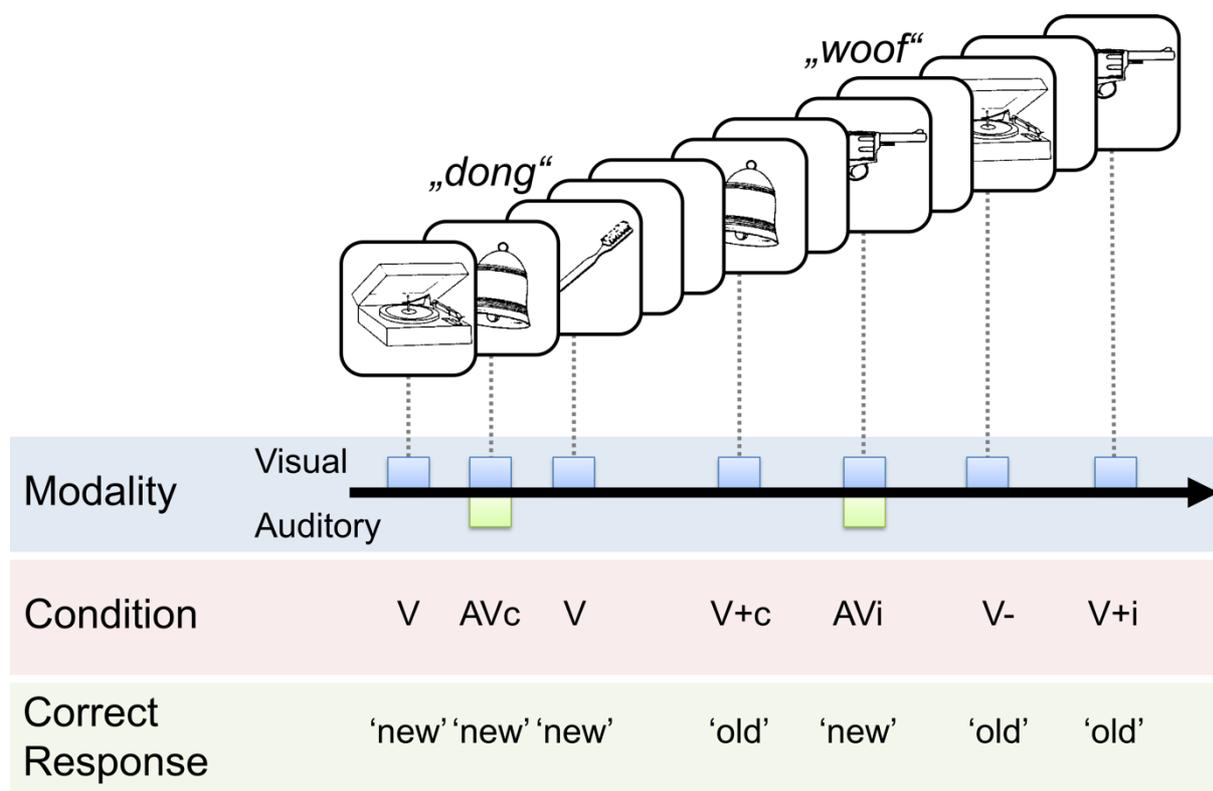
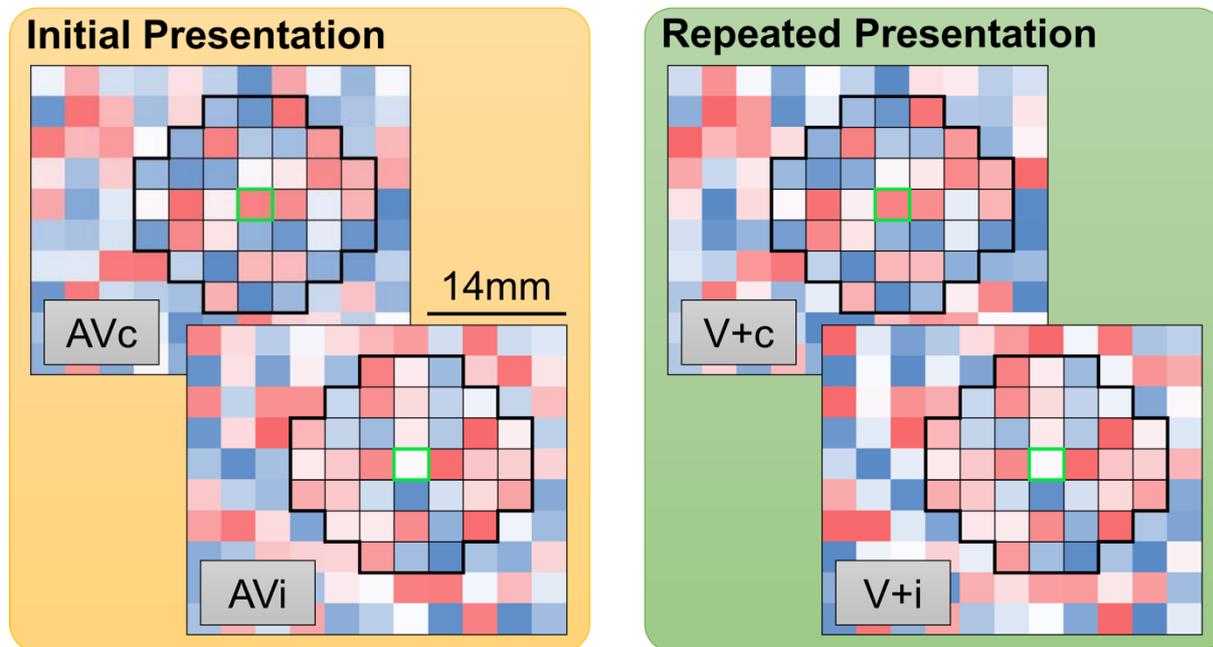


Figure 2

Illustration of the searchlight classification



Training 1: **Train** classifier on *initial* and **test** on *repeated* presentation

Training 2: **Train** classifier on *repeated* and **test** on *initial* presentation

Figure 3

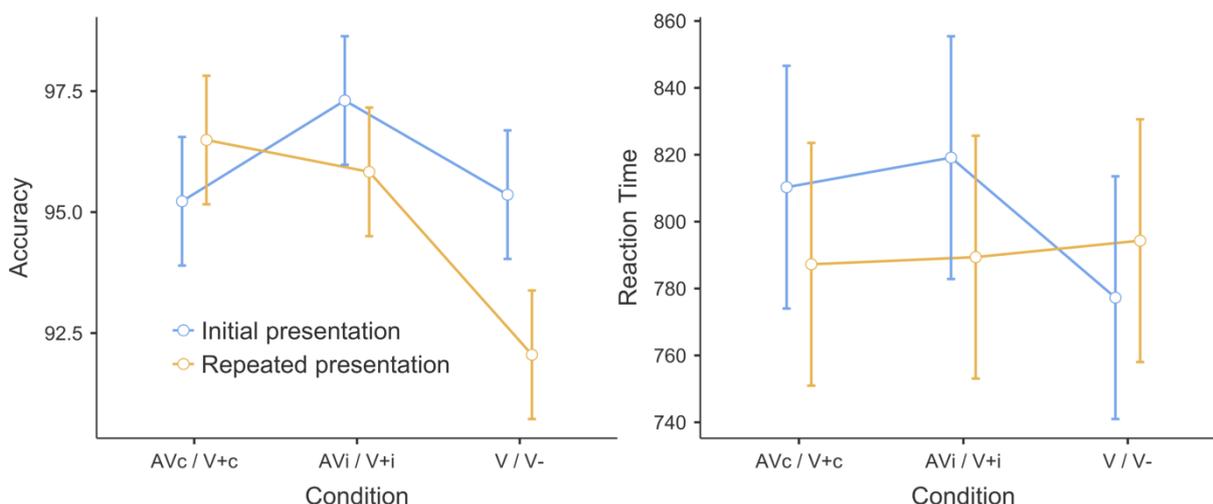


Figure 4

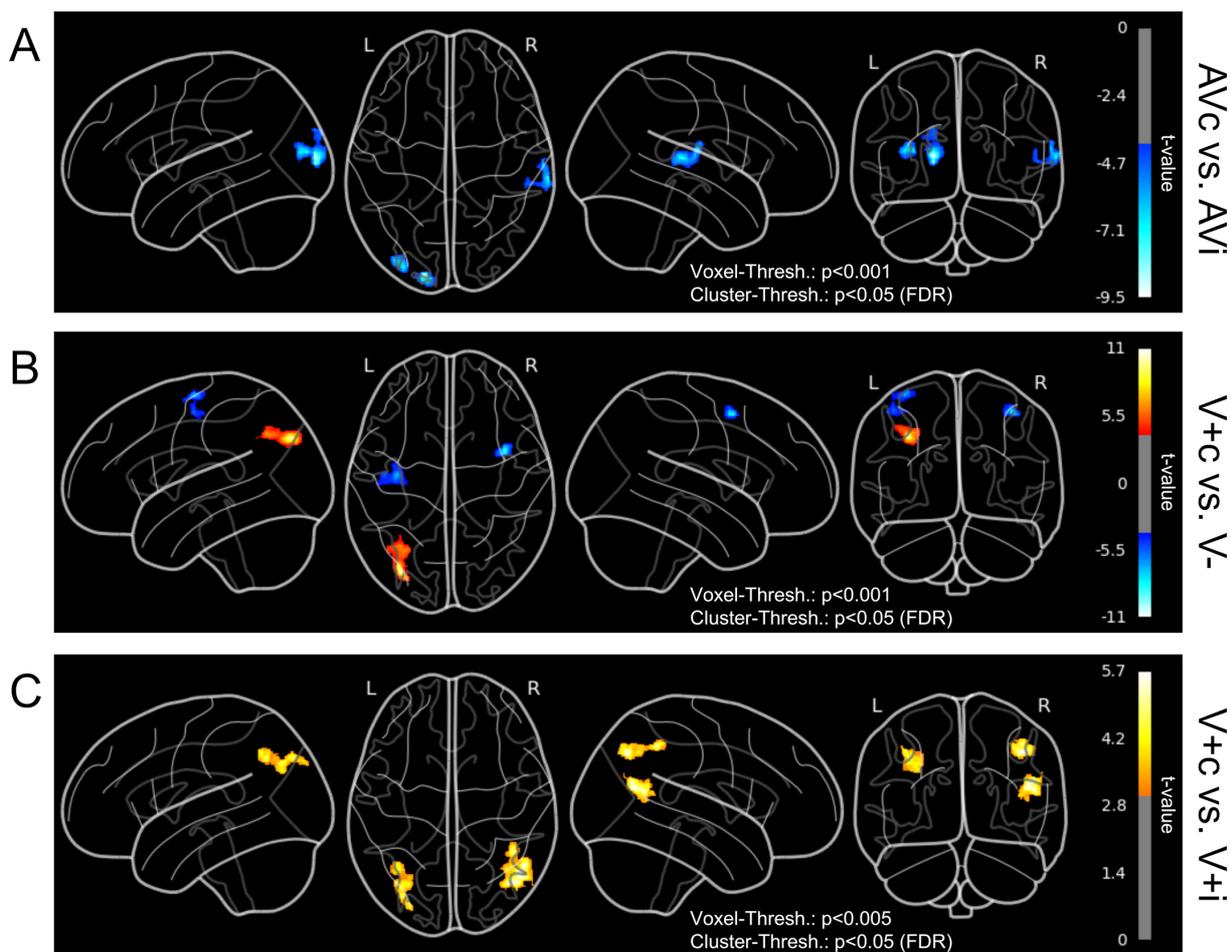
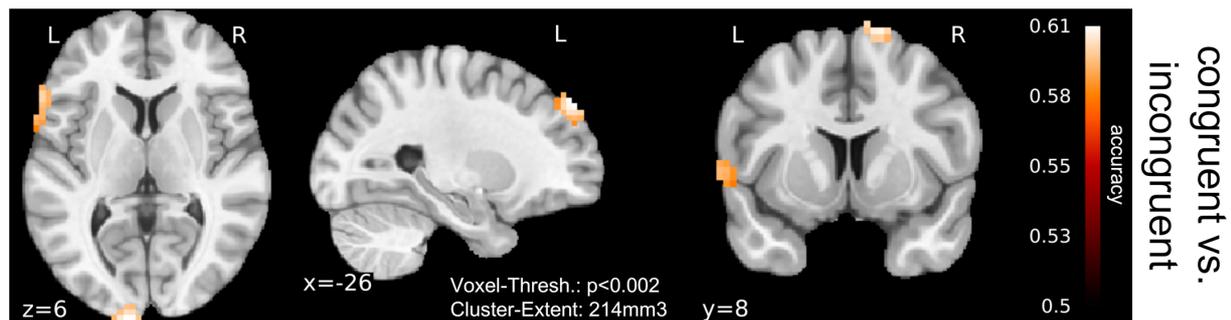


Figure 5



Closing words

Dear reader,

Thank you for going through my thesis until the end! I hope it was interesting and gave you some new insights.

Let me now finish the thesis with where this journey started more than ten years ago, the Nipype logo, of which I changed the color scheme and used it for my first [Nipype Beginner's Guide](#). Image credit goes to Arno Klein (License CCO).



To this day and into the future, the Nipype logo, depicting a Python in the shape of a human brain, perfectly symbolizes my fascination and passion for programming, data analysis, open-source software and neuroscience!