Proceedings of the Master's Programme Cognitive Neuroscience of the Radboud University

Contact Information:
Journal CNS
Radboud University
Postbus 9104
6500 HE Nijmegen
The Netherlands
nijmegencns@gmail.com

Table of Contents

From the Editors-in-Chief

Dear reader,

We proudly present to you 14.2 of the CNS journal. This marks the last CNS Journal issue that will be published with us as Editors-in-Chief and is thus quite special to us.

Time is a strange thing. A week can feel like an eternity if you are awaiting something eagerly. Then again, two years can feel like an eyeblink. That's what it is like for us when thinking about our time as CNS students. Even though the time was filled with so many events, it does feel like it just flew by. Recapping our time in the CNS master program, we think about some stress and frustration (mostly about non-running codes), but primarily we think about the fun that investigating the human brain brings us (yup, nerd talk), about the excitement when that darn code finally did run and about all the great people we got to meet and work with along the way.

The latter of course includes our amazing journal team without whom we would have never been able to put together the journal issues in such short periods of time. Studying is a full-time job that demands a great deal of you. Our team members were willing, however, to sacrifice parts of their sparse free time to the CNS journal. We want to thank them for their commitment, hard work and for really trying to meet the deadlines. ;-)

We enjoyed our time as CNS students and as Editors-in-Chief but it's now time to move on. We are sure that Ilaria and Anne will do an amazing job taking over all Editor-in-Chief duties as of now.
Enjoy reading our latest issue.

Nijmegen, May 2019

Antonia & Katharina

Editors-in-Chief

Dear readers,

We all know what it feels like to be curious about something. We are curious when we are reading a book and cannot wait to find out how the story will end. We are curious when we are catching up with friends and want to know what they did last weekend. And perhaps you are curious to find out about the work that is presented in this journal. The work of many curious CNS master students, who worked very hard to complete their theses and to start the rest of their lives.

After I graduated from the CNS master's programme back in 2015, I started a PhD on the topic of curiosity at the Donders Institute. On the first day of my PhD, I googled "curiosity" to see what the World Wide Web could tell me about it. The first thing that popped up was a famous quote by Albert Einstein saying: "I have no special talent. I am only passionately curious." For some reason this quote always stuck with me. How could one of the greatest scientists ever alive, of who we all believe that he had MANY extraordinary talents, say this about himself?

At the same time, I realized that I also felt like I had approximately zero special talents. I was just one of the many students that graduated that year. One thing I was very good at, however, was telling you exactly in what aspects every single fellow student was more skilled or talented than I was. This made me worry quite a bit about whether I would be able to do everything that needed to be done to complete a PhD. How the hell was I supposed to do this?

Over the years, I talked to many colleagues about these worries. I was amazed when I found out that many of them, even the ones who I really admired for their skills and talents, told me that they have had that exact same feeling at some point during their careers. Apparently, it is very easy to focus on what other people excel at, while only seeing your own failures.

At some point, I realized that something had to change - I had to stop worrying and start focusing on the aspects of my work I really wanted to know more about. I decided to take more courses that I was interested in, I tried to learn news skills and I focused as much as I could on the aspects of my work that sparked my curiosity. In doing so, I came up with new research questions that I really wanted to know the answer to and also the curiosity of my supervisors and others around me became more and more contagious.

Of course it was not always easy to let go and I have to admit that I still sometimes miserably fail in not worrying about failure. However, I realize every day that following my curiosity helps me to be happy and to enjoy the pretty cool life I'm living right now. What else could I have wished for when I started my PhD at the Donders Institute?

I'm not sure about this, but if Albert Einstein meant that focusing on your curiosity will help you to find pleasure in whatever it is you want to do next, I could not agree more. Therefore, my piece of advice to all of you would be: "Stop worrying and embrace your curiosity!"

Best wishes,
Lieke van Lieshout

White Matter Changes in the Perforant Path in ALS: Providing Evidence for ALS as a Multisystem Disease

M. Hiemstra[1], J. Mollink[1,2], M. Pallabage-Gamarallage[3], I.N. Huszar[2], K.L. Miller[2], M. Jenkinson[2], O. Ansorge[3], A.M. van Cappellen van Walsum[1]

[1] *Department of Anatomy, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, Netherlands.*
[2] *FMRIB, Wellcome Centre for Integrative NeuroImaging, Oxford, United Kingdom.*
[3] *Department of Neuropathology, John Radcliffe Hospital, Oxford, United Kingdom.*

Amyotrophic lateral sclerosis (ALS) is a severe, progressive and incurable motor disease. Roughly 20% of the ALS patients are affected by a level of cognitive decline that meets the criteria for behavioural frontotemporal lobe dementia (bvFTD). ALS and bvFTD share some clinical and pathological features, for example, the deposition of TAR DNA binding protein 43(pTDP-43) in several brain regions that are part of the circuit of Papez. Previous literature suggests involvement of the perforant path, a white matter tract in the hippocampus that is part of the circuit of Papez, in patients with both ALS and bvFTD. We hypothesize that white matter degeneration in the perforant pathway is a key feature of ALS, providing a neuronal correlate for ALS as a multisystem disorder. To verify our hypothesis we studied white matter changes in ex-vivo hippocampal blocks from patients with known ALS (n=13) and controls (n=5) using diffusion MRI. The dMRI results were evaluated using polarised light imaging (PLI), a microscopy technique sensitive to density and orientation of myelinated axons. From the same hippocampal blocks, sections were cut and stained for myelin, pTDP-43, neurofilaments and activated microglia. The dMRI results show a significant decrease in fractional anisotropy (p=0.018) and an increase in mean diffusivity (p=0.0017), axial diffusivity (p=0.023) and radial diffusivity (p=0.028) in the perforant path in ALS patients compared to controls, likely indicating a loss of fibres. The PLI retardance values within the perforant path were lower in ALS cases compared to controls, however, not significantly (p=0.16). The retardance correlates with the fractional anisotropy (p=0.04). Furthermore, an increase in dispersion was observed in ALS specimens (p=0.04), implying a less organised axonal structure. Histology data showed a non-significant increase in myelin (PLP) (p=0.11) and an increase in neurofilaments (SMI-312) (p=0.03) in ALS cases compared to controls. No differences were found in the amount of inflammation and two out of the 13 ALS cases exhibited pTDP-43 pathology in the hippocampus. These results demonstrate degradation of the perforant path in ALS patients, providing a potential neuronal correlate for the cognitive symptoms observed in ALS and substantiating the hypothesis that ALS and bvFTD are part of the same spectrum of diseases. Future research should focus on correlating the degree of clinically observed cognitive decline to the amount of white matter atrophy in the perforant path.

Keywords: Amyotrophic Lateral Sclerosis, Frontotemporal Lobe Dementia, Perforant Pathway, Polarised Light Imaging, diffusion Magnetic Resonance Imaging

Corresponding author: Marlies Hiemstra; E-mail: marlieshiemstra@hotmail.com

List of abbreviations

AD	Axial diffusion
ALS	Amyotrophic lateral sclerosis
bvFTD	Behavioral Frontotemporal lobe dementia
CD68	Cluster of differentiation 68
DG	Dentate gyrus
dMRI	Diffusion magnetic resonance imaging
EC	Entorhinal cortex
FA	Fractional anisotropy
FOM-HS	Fibre orientation map hue saturation value
FOV	Field of view
LGN	Lateral geniculate nucleus
MD	Mean diffusivity
PBS	Phosphate buffered saline
PLI	Polarized light imaging
PLP	Proteolipid protein
pTDP-43	Phosphorylated TAR DNA binding protein
RD	Radial diffusion
Sub	Subicular cortices, pre-subiculum and subiculum

Introduction

Amyotrophic lateral sclerosis (ALS) is a severe, progressive and incurable motor disease that is characterized by death of motor neurons in the brain and spinal cord, resulting in a loss of voluntary movement. It is the most common form of adult onset motor neuron degeneration and has an incidence of 1-2 per 100,000 per year in the world population and about 6-7 per 100,000 in the European population (Bonafede & Mariotti, 2017; Ferrari, Kopogiannis, Huey & Momeni, 2011; Logroscino et al., 2010; Tan, Ke, Ittner & Halliday, 2017). ALS can occur both sporadically (90%) and familially (10%) and has a median survival of three years after disease onset (Tan et al., 2017). ALS has mainly been described as a neurological disorder that affects the motor system, but more and more evidence is suggesting ALS to be a multisystem neurodegenerative disease because other areas besides motor areas of the brain also undergo degeneration. One of the most prominent bodies of evidence of ALS being a multisystem disorder is that 50% of the ALS patients suffer from a form of cognitive impairment and 15% of the ALS patients are affected by a level of cognitive decline that meets the criteria for behavioural frontotemporal lobe dementia (bvFTD) (Cykowski et al., 2017). We examined the whole brains and spinal cords of 57 patients (35 men; 22 women; mean age 63.3 years; 15 patients with c9orf72-associated ALS [c9ALS]; Phukan et al., 2012; Tan et al., 2017) the population-based frequency, clinical characteristics and natural history of cognitive impairment in amyotrophic lateral sclerosis (ALS). However, a clear neural correlate of these cognitive symptoms has been lacking to date. Looking into the resemblance between bvFTD and ALS provides information about the neural substrate of the cognitive problems observed in ALS.

Frontotemporal dementia (FTD) is a heterogeneous form of dementia that is characterized by progressive neurodegeneration in the temporal and frontal lobes. Different symptoms can prevail in FTD, resulting in three different subtypes: bvFTD, semantic FTD and progressive non-fluent FTD (Gao, Almeida, 2017; Ghosh & Lippa, 2015; Warren et al., 2013).

ALS and bvFTD share neuropathological, genetic and clinical features providing lending to the hypothesis that bvFTD and ALS are part of the same spectrum of diseases and providing further evidence for ALS being a multisystem disorder (Ferrari et al., 2011). The strongest neuropathological resemblance between ALS and bvFTD pathology is the deposition of ubiquitinated and hyperphosphorylated TAR DNA binding protein 43 (pTDP-43) in select neurons and glial cells of the central nervous system (Ferrari et al., 2011; Scotter, Chen & Shaw 2015). Normal TAR DNA binding protein 43(TDP-43) is mainly located in the cell-nucleus where it plays a role in inhibition of splicing and in exon skipping. In ALS and FTD pTDP-43 is mainly located in the cytoplasm where it forms potentially neurotoxic aggregations (Hu & Grossman, 2009; Mackenzie & Rademakers, 2008). In ALS, pTDP-43 pathology is present in different brain regions and develops over time via a specific pattern. Lesions spread from the motor neurons in the brainstem and spinal cord and also from the frontal neocortex to the parietal and temporal neo- and sub-cortical regions (Brettschneider et al., 2014; Brettschneider et al., 2013).

In bvFTD sequential patterns of spreading of the pTDP-43 inclusions are also observed. In bvFTD, the inclusions spread from the orbitofrontal cortex and amygdala towards the frontal and temporal cortices before progressing to the visual cortex, motor system and cerebellum (Brettschneider et al., 2014; Tan et al., 2015).

Even though pTDP-43 is associated with both ALS and FTD, genetic screening has been controversial, leaving the question open whether mutations in TDP-43 are causal to the neuropathology of ALS/FTD. Most of the genetic variability that is related to TDP-43, like mutations in the TARDBP gene that encodes for TDP-43, seems to be related to ALS cases (Sreedharan et al., 2008). To date there is hardly any evidence of mutations in TDP-43 associated with either FTD or ALS. Nevertheless, some genetic overlap between FTD and ALS can be found in other genes. For example, in both diseases mutations in the C9ORF2 gene and mutations in progranuline gene (PGRN) seem to play a role (Gao, Almeida, 2017; Ferrari et al., 2011).

Another resemblance between ALS and bvFTD was found in the previously mentioned areas affected by pTDP-43. Several of these cortical regions relate to each other via the Papez circuitry. The Papez circuit is a major loop of the limbic system and its main functions are the cortical control of emotion and the consolidation and retrieval of memory (Pessoa & Hof, 2015; Musio, 1997). The circuit is comprised of the hippocampal formation, mammillary bodies, fornix, anterior thalamic nucleus and the anterior cingulate (Fig. 1) (Shah, Jhawar & Goeal, 2012).

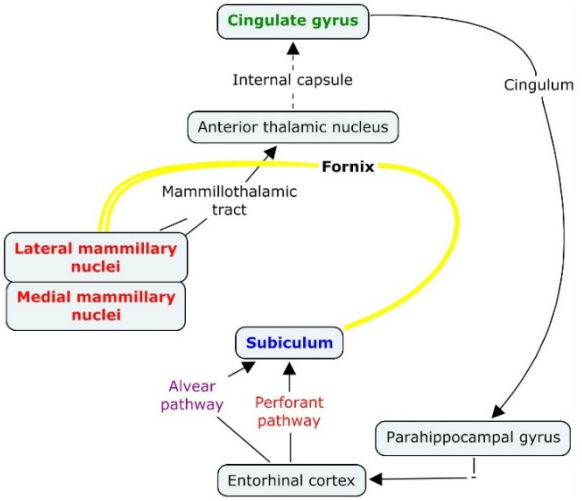


Figure 1. Schematic representation of the Papez circuit consisting of the entorhinal cortex, which projects to the subiculum via the perforant and alvear pathway. The subiculum has a projection to the mammillary nuclei via the fornix. From here the circuit continues to the cingulate gyrus via the anterior thalamic nucleus and the internal capsule. From the cingulate gyrus there is a projection to the parahippocampal gyrus via the cingulum ("figure Papez circuit," n.d.).

In vivo and post mortem MRI volumetric analyses demonstrate distinct regional grey and white matter atrophy of the Papez circuit in bvFTD (Hornberger et al., 2012; Irish et al., 2013). The main symptoms of bvFTD are functionally related to the Papez circuit since the main characteristics are progressive change in emotional regulation and changes in personality (Bott, Radke, Stephens & Kramer, 2014). Recent evidence also recognises episodic memory deficits in bvFTD, with the amount of atrophy in the Papez circuit determining the degree of episodic memory problems (Hornberger and Piguet, 2012; Hornberger et al., 2012).

Little is known about the involvement of the Papez circuitry in ALS. However, it was shown that the amount of TDP-43 pathology in ALS patients correlates with cognitive decline (Montreal Cognitive Assessment test scores), but not with the duration of the disease or its rate of progression (Brettschneider et al., 2013) brainstem motor nuclei of cranial nerves V, VII, and X-XII, and spinal cord α-motoneurons (stage 1; Cykowski et al., 2017). We examined the whole brains and spinal cords of 57 patients (35 men; 22 women; mean age 63.3 years; 15 patients with c9orf72-associated ALS [c9ALS]).

Most studies on neurodegeneration in ALS have focused on white matter areas related to motor function like the motor cortex, the corpus callosum and the corticospinal tract (Cirillo et al., 2012; Filippini et al., 2010; Horsfield & Jones, 2002; Lillo et al., 2012). To our knowledge, there are only three studies that investigated sub-cortical changes (Barbagallo et al., 2014; Thivard et al., 2007; Takeda, Uchihara, Mochizuki, Mizutani & Iwata, 2007). Two of these studies focused on grey matter integrity. An increase in mean diffusivity (MD) was found in ALS patients compared to controls in the amygdala, hippocampus and thalamus, which are part of the circuit of Papez (Barbagallo et al., 2014; Thivard et al., 2007).

The third study linked ALS to one of the projections of the circuit of Papez: the perforant path (Takeda, Uchihara, Mochizuki, Mizutani & Iwata, 2007). This path runs from the neurons in layer II and III of the entorhinal cortex (EC), through the pyramidal layer of the subiculum (Sub) to the molecular layer of the Sub, dentate gyrus (DG) and the CA3 region of the hippocampus (Fig. 2). The path is known to have an important function in memory and degeneration of this path is linked to the severe memory impairments in Alzheimer's disease (Hyman, Van Hoesen, Kromer & Damasio, 1986; Thal et al., 2000). The third ALS-related study however did not directly investigate the white matter of the perforant path, but rather evaluated the regions belonging to the path with a haematoxylin/eosin stain. Semi-quantitative analysis showed entorhinal cortex degeneration, subicular degeneration and laminar spongiosis of the DG in ALS patients. This damage, potentially caused by pTDP-43 inclusions, correlated with the degree of episodic memory deficits (Takeda, Uchihara, Mochizuki, Mizutani & Iwata, 2007). The previously mentioned white matter atrophy of the circuit of Papez in bvFTD also begins in the hippocampal area (Hornberger et al., 2012; Irish et al., 2013).

Since the perforant path seems to be affected in both ALS and dementia, we hypothesize that degeneration of the perforant path is one of the key projections to be affected in ALS, providing further evidence for ALS as a multisystem disease and providing a neural correlated for the cognitive impairments in ALS.

To verify our hypothesis, several imaging techniques were used to visualise the white matter tracts within and surrounding the hippocampus on post mortem brains of controls and of patients with ALS. First, diffusion magnetic resonance imaging (dMRI), a technique that provides 3D information about white matter tracts based on the diffusion of water within tissue (Bastiani & Roebroeck, 2015), was performed. By measuring the orientation dependence of water diffusion, an estimation of

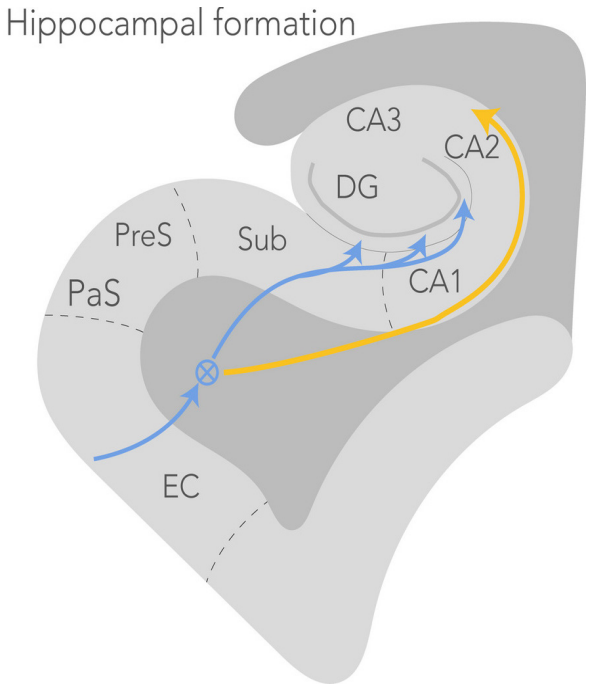


Figure 2. A schematic representation of the perforant path (blue) and the alvear path (yellow). The perforant path starts in the entorhinal cortex and projects through the subiculum via the molecular layer of the hippocampus to the dentate gyrus and the CA3 region of the hippocampus (circled-cross indicates out of screen – i.e. anterior-posterior – orientation). The alvear path projects via the CA1 region into the hippocampus. CA, cornu ammonis; DG, dentate gyrus; EC, entorhinal cortex; PaS, parasubiculum; PreS, presubiculum; Sub, subiculum.

the microstructure of the brain tissue was made (Jbabdi, Sotiropoulos, Haber, Van Essen & Behrens, 2015). dMRI has a resolution on the millimetre scale and because it is a non-invasive technique, it can be used for *in vivo* diagnostics and research (Bastiani & Roebroeck, 2015).

Second, polarised light imaging (PLI), a microscopy technique that quantifies the orientation of myelinated axons based on birefringence of myelin sheath in thin brain slices, was used. PLI is a relatively new technique and has, to our knowledge, never been used to compare properties of myelinated tracts between cases and controls. We correlated the results of PLI with the different dMRI metrics. To correlate differences in white matter with potential protein pathologies and to further complement the PLI and dMRI data, (immuno)histochemical stainings were carried out to visualise myelin, microglial activation, neurofilaments and pTDP-43 inclusions.

The aim of the research is to identify the potential changes in grey and white matter within the hippocampus in patients with ALS to further contribute to the hypothesis that ALS is a multisystem disorder and to get insight in the potential neural substrates of the cognitive deficits observed in ALS. Furthermore, our research elaborates on the hypothesis that ALS and bvFTD are part of the same spectrum of diseases.

Methods

Tissue

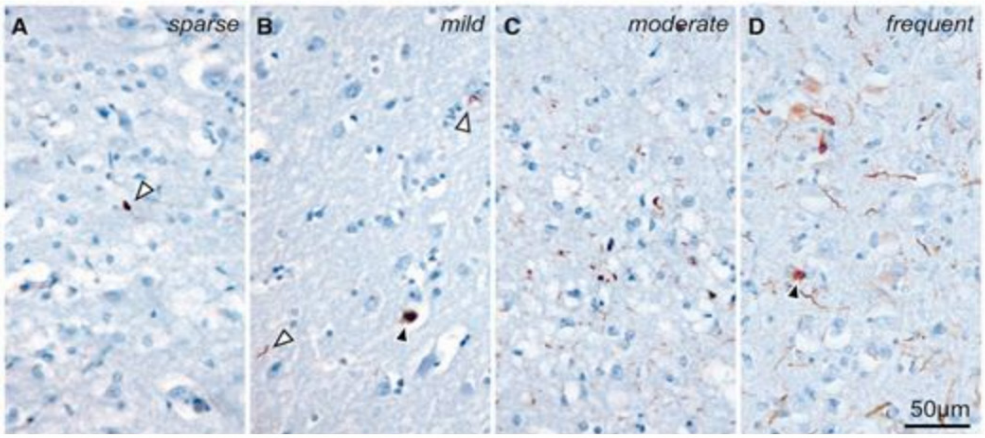
For this study, 13 post-mortem ALS human brains were acquired with permission from the Oxford Brain Bank at the John Radcliffe Hospital in Headington, United Kingdom. Five age matched post-mortem human control brains were acquired via the body donor program at the Department of Anatomy of the Radboud University Medical Centre, Nijmegen, The Netherlands. The human brains were removed from the skull and fixed in 2-4% formalin. Details on the history of each specimen are listed in Table 1. All brains were cut in coronal slabs of 5 mm thickness and the hippocampus was sampled in a slab comprising the lateral geniculate nucleus (LGN) as a key anatomical landmark. The hippocampal regions representing the perforant path were sampled for ALS cases (n=13, Oxford Brain Bank) and controls (n=5, Department of Anatomy, Nijmegen).

ALS classification was based on the neuropathological assessment of pTDP-43 proteinopathy in key gatekeeper regions previously defined by Tan and colleagues (Tan et al., 2015) encoded by TARDBP. pTDP-43 protein pathology was visually rated on a scale of 0 to 4 based on the rating scale of Tan and colleagues, with 0 indicating no protein pathology and 4 indicating severe pathology (Fig. 3).

Representative blocks of 13 other brain regions were sampled and stained following standard brain banking protocol for diagnostic purposes (Table 2) (Montine et al., 2012).

Magnetic resonance imaging acquisition

The imaging pipeline for all hippocampus blocks is shown in Figure 4. Out of the 18 blocks, 17 blocks were analysed with dMRI (Control n=5, ALS n=12). Prior to scanning all specimens were placed in phosphate buffered saline (PBS 0.1M pH=7.4) for at least 24 hours to reverse the decrease in T2-relaxation introduced by formalin (Shepherd, Thelwall, Stanisz & Blackcand, 2009) time-intensive MRI acquisitions without motion artifacts, such as those required



The severity of TDP-43 pathology in the anterior cingulate cortex rated on a five-point scale. 0 = no detectable pathology across the entire section (not shown), 1 = sparse (<5 inclusions across the entire section, **A**); 2 = mild (some pathology observed in most fields of view at × 100 magnification, **B**); 3 = moderate (**C**); and 4 = frequent (**D**). Solid arrowheads indicate neuronal inclusions. Open arrowheads show examples of neurites and glia.

Figure 3. Visual rating scale for pTDP-43 pathology developed by Tan et al. The anterior cingulate was stained for pTDP-43 and counterstained with haematoxylin. Pathology was rated on a scale from 0-4, with a 0 indicating no pathology, 1 indicating sparse pathology, 2 indicating mild pathology, 3 indicating moderate pathology and a 4 indicating severe pathology. Solid arrowheads indicate neuronal inclusions. Open arrowheads show examples of neurites and glia (Tan et al., 2015).

Table 1.
Post mortem specimen details from control (1-5) and ALS/ALS-FTD brains(6-18). Abbreviations: Post-mortem interval (PMI), fixation time (FT), cause of death (COD).

#	PMI (days)	FT (months)	Sex	COD	Age
1	0.5	12	M	Heart failure	67
2	0.5	14	M	Lungcancer	61
3	0.5	8	F	Old	82
4	0.5	8.5	M	Leukemia	73
5	1	15	F	Fluid in the lungs	71
6	4	47	F	Unknown	61
7	Unknown	Unknown	F	Unknown	77
8	3	42	M	Unknown	82
9	5	42	F	Unknown	53
10	4	39	M	Unknown	76
11	2	36	M	Unknown	78
12	1	36	F	Unknown	81
13	2	35	F	Unknown	81
14	2	26	F	Unknown	80
15	2	15	F	Infection (?), pneumonia. Difficulty with breathing.	67
16	2	14	F	Respiratory failure	74
17	4	14	M	Shortness of breath	55
18	2	13	F	Unknown	53

Table 2.
Stainings performed on 13 brain regions sampled for diagnostics, following standard brain banking protocol. Staining were performed for Alzheimer's disease (4G8 and AT8), Parkinson's disease (alpha-synuclein), ALS (TDP-43) and other general neuropathology (p62).

	HE	4G8	Alpha-synuclein	AT8	TDP-43	P62
Basal Ganglia	x	x	x			
Anterior hippo	x	x		x		
Amygdala	x	x	x	x	x	
Ant cingulate	x		x			
Mid front gyrus	x	x	x	x		x
Temp gyrus	x	x	x	x		x
Inferior parietal lobule	x	x	x	x		
Occipital gyus	x	x	x	x		
Midbrain	x	x	x			
Pons	x		x			
medulla	x		x			x
Cerebellum	x	x				x
M1/S1	x					x

for brain atlas projects, but the aldehyde fixatives used may significantly alter tissue MRI properties. To test this hypothesis, this study characterized the impact of common aldehyde fixatives on the MRI properties of a rat brain slice model. Rat cortical slices immersion-fixed in 4% formaldehyde demonstrated 21% and 81% reductions in tissue T1. For scanning, specimens were transferred to a 100 ml syringe filled with fluorinert® (Solvay Solexis Inc), a proton-free solution which is susceptibility matched to the tissue. If necessary, additional gauzes were place in the syringe to immobilize the tissue.

All scans were performed at room temperature on a 11.7T Bruker BioSpec Avance III preclinical MR system (Bruker BioSpin, Ettlingen, Germany) using a birdcage coil (Bruker Biospin). T1- and T2-weighted high-resolution structural images were obtained for anatomical reference.

T1-weighted structural images were acquired with a fast low angle shot gradient echo (FLASH) sequence at a resolution of 0.1x0.1x0.1 mm (TR=25 ms, TE= 3.4 ms, flip angle=10°). T2-weighted images were acquired with a multi-gradient echo sequence at a resolution of 0.1x0.1x0.1mm ((TR=35.813 ms, TE= 15 ms, flip angle=30°).

Diffusion weighted images were obtained using a spin-echo sequence with an EPI readout (TR=13.75s, TE= 30.066 ms, Δ= 12.5 ms, δ= 4.0). Two shells were acquired (b=2000,4000 s/mm^2), each employed at 64 gradient directions (128 in total) in addition to six images with negligible diffusion weighting (b=14s/mm^2). The specimens were covered in the coronal plane of the scanner with a field-of-view of 28.8 x 28.8 mm that was sampled with a 72 x 72 matrix. Together with a slice thickness of 0.4 mm this created isotropic voxels. The number of slices varied between 80 and 100 depending on the dimension of each specimen.

Probabilistic tractography

Processing of all MR images was performed with FSL (Woolrich et al., 2009) noisy images of the brain. This might be the inference of percent changes in blood flow in perfusion fMRI data, segmentation of subcortical structures from structural MRI, or inference of the probability of an anatomical connection between an area of cortex and a subthalamic nucleus using diffusion MRI. In this article we will describe how Bayesian techniques have made a significant impact in tackling problems such as these, particularly in regards to the analysis tools in the FMRIB Software Library (FSL). A diffusion tensor was estimated at each voxel in the dMRI data. Four different tensor metrics were derived from this diffusion tensor: the fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD) and the radial diffusivity (RD). The FA value is a metric describing the diffusion directionality of water ranging from 0 to 1, with a value of 0 indicating fully isotropic diffusion and 1 indicating fully anisotropic diffusion (Fig. 5). A reduction in FA may indicate white matter damage. The MD value is a measure for the total amount of diffusion, where an increase

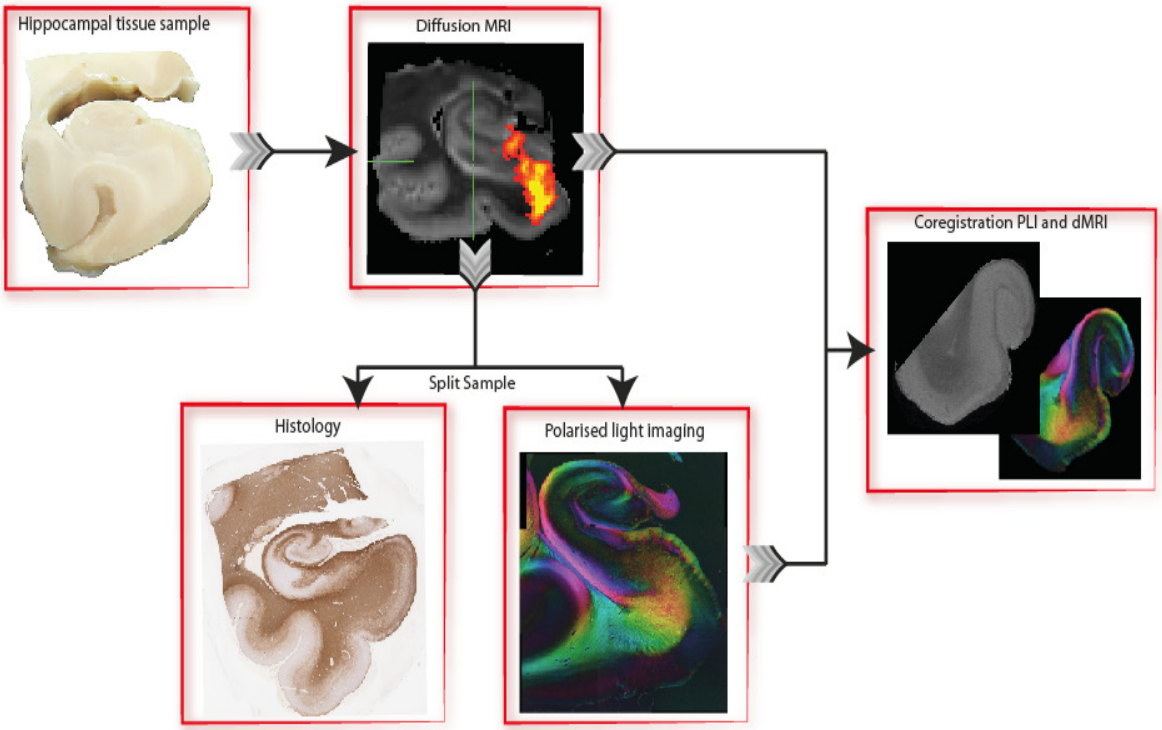


Figure 4. Processing and analysis pipeline of the specimens. All hippocampal blocks were scanned with diffusion MRI. After scanning, the blocks were bisected into two parts, one part was used for histological staining and the other part for polarized light imaging). In the end, the polarized light imaging data was co-registered with the diffusion MRI data.

in MD may also be possible marker for white matter damage. The AD is a direct indicator of the amount of the diffusion along the main diffusion axis, while the RD gives information about the diffusion along the axes parallel to the main diffusion direction. An increase in RD was previously shown to correlate with a reduction in myelin (Aung, Mar & Benzinger, 2013).

In order to extract the FA, MD, RD and AD metrics within the perforant path, probabilistic tractography was performed. To delineate the perforant path in the hippocampus specimen, tractography was run on the dMRI data. First, we estimated the fibre configuration in each voxel using the Bedpostx algorithm (Behrens et al., 2003) which was modified to handle multi-shell diffusion data (Jbabdi, Sotiropoulos, Savio, Graña & Behrens, 2012). We fitted up to three fibre orientations within each voxel depending on the level of support by the diffusion signal. Probabilistic tractography was then performed using the probtrackx2 algorithm (Behrens, Berg, Jbadbu, Rushworth & Woolrich, 2007). The (pre-)subiculum and the DG were manually segmented and served as seed and waypoints masks, respectively (Fig. 6). For each seedvoxel, we generated 5000 streamlines per voxel, only keeping those that terminate in the DG. The curvature threshold was set to 0.2 (corresponding to approximately 80°), with a step length of 0.2 and a minimal fibre volume threshold of 0.01. A loop check was performed such that pathways that looped back to themselves were excluded. After probabilistic tractography, a connection probability threshold was empirically determined and only the streamlines that followed the approximate course of the performant path were retained using exclusion masks that were determined for each slice individually. From all perforant path voxels, FA, MD, RD and AD values were extracted and averaged for each specimen separately.

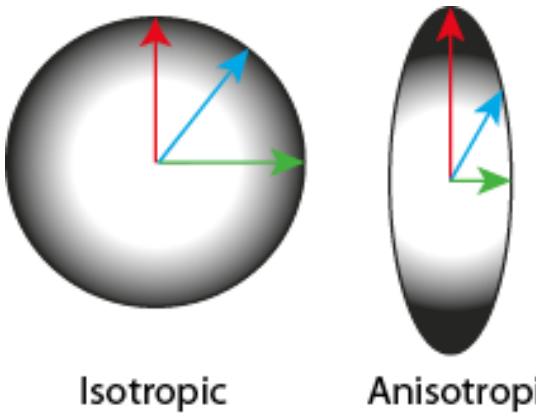


Figure 5. Isotropic diffusion (left) is associated with an FA value of 0. The more anisotropic (right) the diffusion, the higher FA value.

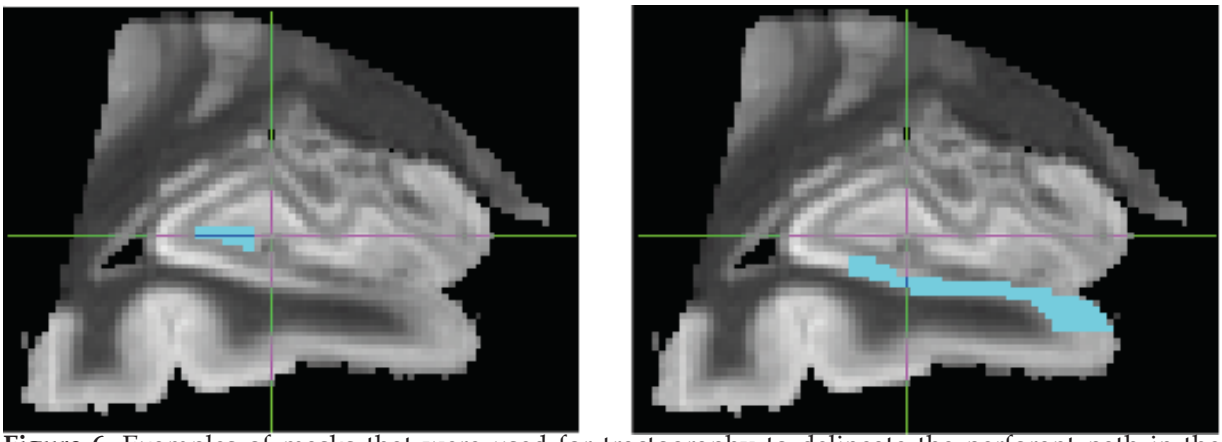


Figure 6. Examples of masks that were used for tractography to delineate the perforant path in the hippocampus. Left: The seed mask (start of tractography) for the presubiculum and subiculum. Right: The waypoint mask (end of tractography) for the dentate gyrus.

Polarized light imaging

PLI is a microscopy technique that is able to quantify the fibre orientation of myelinated axons using the birefringent property of myelin. When passing polarized light through brain tissue, the interaction of the light with the birefringent myelin results in a transmitted light intensity that depends on the angle between the myelin sheet and the polarizing filter on the microscope. The PLI raw signal consists of the transmitted light intensity at a single pixel over various polarizer orientations from which an in-plane fibre orientation can be estimated (Axer et al., 2011; Larsen, Griffin, Gräßel, Witte & Axer, 2007) the structural basis of the human connectome. In contrast to animal brains, where a multitude of tract tracing methods can be used, magnetic resonance (MR). After dMRI, all 18 hippocampal blocks were bisected along the coronal plane into two approximately equal parts. One part was embedded in paraffin for histology and the other part was frozen for PLI (Fig. 3). Frozen sections are preferred for PLI as the lipids in the myelin sheath remain preserved. After cryoprotection with 30% sucrose in PBS, all tissue blocks were sectioned at 50 µm thickness on a freezing microtome (Microm HM 440E Microtome). The sections representing the perforant path were mounted on glass slides coated with 0.5% gelatin and 0.05% potassium chrome(III) sulfatelycerin and were coverslipped with PVP, a water-soluble mounting medium.

PLI images were acquired on a Zeiss Axio Imager A2 microscope upgraded with a stationary polarizer, a quarter wave plate and a rotating polarizer. Light first passes through the stationary polarizer and a quarter wave plate positioned at a 45◦ angle relative to the stationary polarizer to create circularly polarized light. Once polarized, the light encounters the birefringent myelin that induces a phase shift. To capture the angular extent of this phase shift, images were acquired at 9 equiangular orientations of the rotating polarizer from 0° to 160° with a CCD camera. Together with a 1.25x magnifying objective, this yielded a spatial resolution of approximately 4 µm/pixel. Only the green channel was used for further processing as the quarter wave plate is designed for this wavelength. A set of background images was acquired for every rotation angle to correct for inhomogeneous background illumination. Background correction of the images was performed as described by Dammers and colleagues (2010). Three different parameters were derived from the raw PLI data by fitting the light intensity at each pixel to a sinusoid. The phase of the sinusoid gave the in-plane orientation, the phase shift induced to the light wave provided the retardance and the transmittance was calculated as the average amount of light passing through the tissue. The retardance and in-plane orientation maps were combined in the fibre orientation map, allowing for which visualization of the direction of myelinated fibres within the tissue (Axer et al., 2011b).

As the hippocampus sections were larger than the microscope's field-of-view (FOV), multiple FOVs were acquired to cover an entire specimen. We ensured that there was at least 10% overlap between neighbouring FOVs, such that these could be stitched together automatically. Landmarks in each FOV were described using Speed Up Robust Features (SURF) and matched together using custom written software in MATLAB (MATLAB and Statistics Toolbox Release 2017a, The MathWorks, Inc., Natick, Massachusetts, United States). A rigid transformation between the neighbouring FOVs

was then computed based on the distance between the matched features. After all transformations were estimated, the individual FOVs were stitched together to reconstruct the entire specimen.

Co-registration of PLI and MRI data

In order to make an anatomically correctly match between the dMRI and PLI data, non-linear registration of PLI-derived parametric images with corresponding MRI volumes was performed. The PLI and MRI images were registered in 2 dimensions after manually selecting a slice from the MRI volume that maximised the visual correspondence of observed anatomical contours between PLI and MRI images. The segmentation mask was subsequently applied to all parametric PLI images pertaining to the same specimen. Bi-directional non-linear image registration was carried out between segmented transmittance images and segmented 2D MR images so that aligned PLI-MRI image pairs were obtained at both at the resolution of MRI and that of PLI. The obtained registration parameters were subsequently used to register other parametric PLI images with the MRI volume (Fig. 7). The registration algorithm was implemented in Python 3.6 by Istvan N. Huszar, based on the Tensor Image Registration Library (Huszar, Miller, Pallebage-Gamarallage, Ansgore, & Mirfin, 2018).

The course of image registration comprised of three steps: resizing the input image, linear registration, and non-linear registration. The input image was resampled using bilinear interpolation to match the resolution of the target image. The isotropic scaling factor was inferred from the parameters of an initial rigid-body registration between the PLI-transmittance and the MR images using the linear registration model of the second step. The linear registration step utilised a 2D affine transformation model to maximise the normalised mutual information of the images using Powell's optimisation method. In the non-linear registration step, the images were re-represented using the Modality-Independent Neighbourhood Descriptor

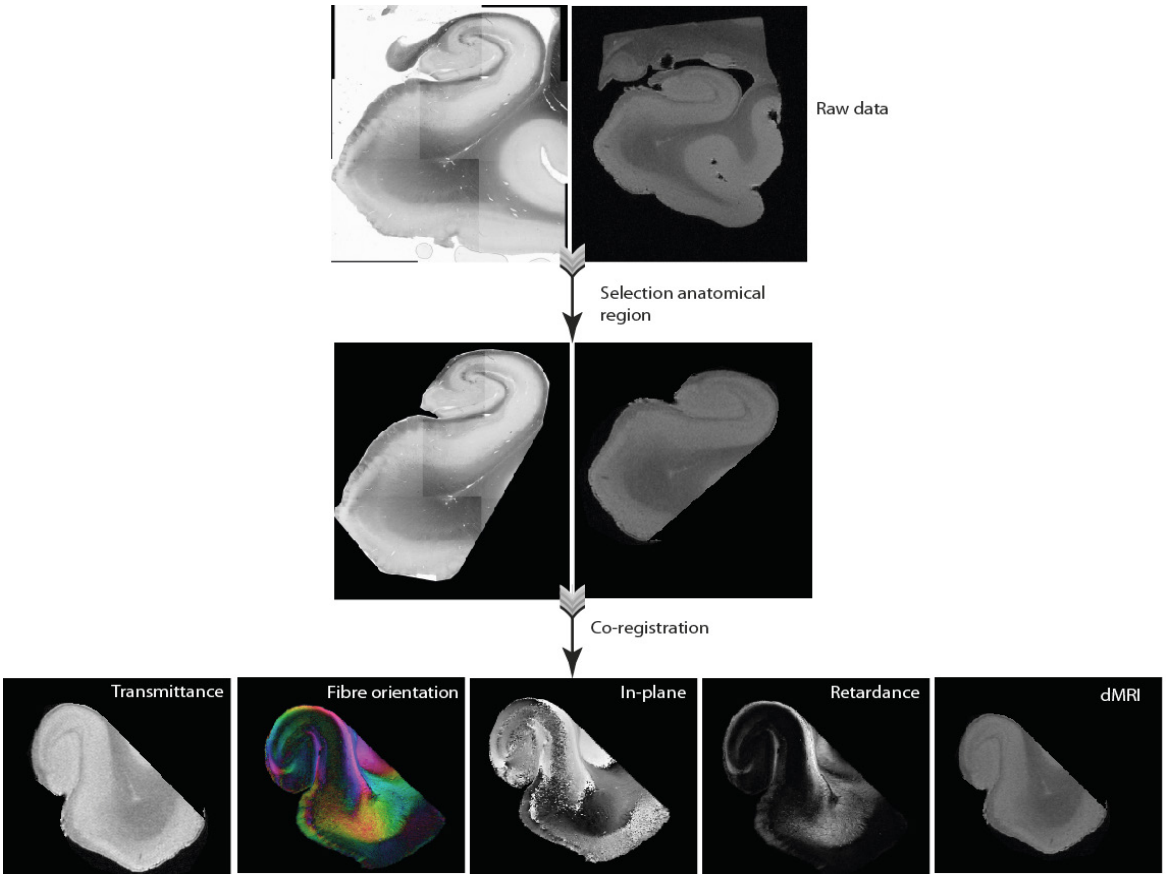


Figure 7. Example of co-registration of PLI with dMRI. On the top row, the transmittance map (left) and dMRI T1-weighted image (right) are depicted. A selection of the anatomical regions representing the perforant path was made to improve the co-registration result and the PLI selection was co-registered with the dMRI selection (middle row). The transformation that was applied to the transmittance map was also applied to the fibre orientation map, in-plane fibre orientation map and retardance map (bottom row).

(MIND) (Heinrich et al., 2012). This representation was used to calculate a cost function that was less sensitive to the differences in image contrast across the two modalities. The cost function was defined as the cumulative Euclidean distance between the image's MIND vectors, and it was minimised using Gauss-Newton optimisation for fast convergence (Heinrich et al., 2012). During the minimisation process, the input image was transformed voxel-wise, using a diffusion-regularised deformation field to ensure the smoothness of image deformation. A multi-resolution approach was adopted in both stages of the registration to avoid misalignment due to local minima in the cost functions. The image registration approach used in this study is a re-implemented and modified version of the algorithm that had been used previously by our research group to align histological images with MRI (Mollink et al., 2017). The modifications were necessary to obtain accurate spatial alignment between PLI and MRI.

Analysis of PLI data

Retardance.

In addition to fibre orientation, the retardance was derived from the PLI raw images. The retardance informs about how much the phase of the light wave is shifted due to interaction with myelin and is therefore used to evaluate the amount of myelin within a section. For one section per subject, the retardance values over the perforant path found with dMRI tractography were calculated by co-registering the PLI data with the dMRI data (Fig. 7). However, in between 8-12 PLI sections per subject were sectioned and processed. Due to time constraints, it was not possible to co-register all sections. Therefore, the other 7-11 sections were analysed by manually delineating the perforant path observed in diffusion MRI in the PLI transmittance map (supplementary Fig. 2). The retardance values under the dMRI path were extracted and averaged for all slices for each specimen to look for case-control differences.

Dispersion.

Another parameter that was extracted from the raw PLI data was the in-plane direction. The in-plane map describes the in-section direction of each fibre ranging from 0 to π. This map was used to estimate fibre dispersion. A higher degree of dispersion is associated with the beaded and fragmented fibre geometries that are associated with Wallerian axonal degeneration (Kerschensteiner, Schwab, Lichtman & Misgeld, 2005; Waller, 1850). Following a similar method to quantify the amount of dispersion from PLI fibre orientations as Mollink and colleagues (2017), a Von Mises distribution was fitted to fibre orientations in a local neighbourhood of 50x50 pixels. The Von Mises distribution is a circular normal distribution whose concentration is described by κ (supplementary Fig. 3). When all samples are oriented isotropic, $\kappa = 0$ and κ increases when the distribution becomes more centred around one angle. Dispersion is therefore calculated as:

$$Dispersion = 1/\kappa$$

The region of interest, namely the perforant path, was manually delineated in all polarized light imaging sections based on the tractography data for each subject separately and the dispersion values within the region of interest were calculated.

Immunohistochemistry

To investigate the pTDP-43 protein pathology, microglial activation, myelination and the amount of neurofilaments, antibodies against pTDP-43, cluster of differentiation 68 (CD68), proteolipid protein (PLP) and SMI-312 were used for immunohistochemistry, respectively. All tissue blocks were embedded in paraffin and cut in 6 µm thin sections for staining for CD68, PLP and SMI312. Sections for pTDP-43 had a thickness of 10 µm. For every specimen 5 sections were stained, one for each protein. The sections were then mounted on Superfrost Plus slides (ThermoFisher, Art. No. J1800BMNZ). The sections were deparaffinised in xylene and rehydrated through a series of 100%, 100%, 90% and 70% ethanol. The endogenous peroxidase was blocked by incubating the sections 3% H_2O_2 in PBS for 30 minutes. Subsequently, antigen binding sides for pTDP-43 and CD68 were retrieved by autoclaving in citrate buffer (pH6). Antigen retrieval for PLP and SMI-312 was achieved through heating the sections submerged in citrate buffer (pH=6) in a microwave. After antigen retrieval, the sections were washed with PBS. After washing, the primary antibody was added. All primary antibodies were diluted in TBS/T (pH=7.6). The primary antibodies against PLP (mouse-anti- PLP, 1:1000, Bio-Rad, MCA839G), CD68 (mouse-anti-CD68, 1:200, Dako Denmark, M0876) and SMI-312 (Mouse-anti-SMI-312, 1:2000, BioLegend, 837901) were incubated for one hour at room temperature. The primary antibody against pTDP-43 (Mouse-anti-pTDP-43, 1:40 000, CosmoBio, PS409/410) was incubated over night at 4 °C. After washing, the

secondary antibody containing HRP Rabbit/Mouse Serum (EnVision Detection Systems, Peroxidase/DAB, Rabbit/Mouse) was added and incubated for 1 hour. After washing, a DAB substrate buffer solution was added (EnVision Detection Systems, Peroxidase/DAB, Rabbit/Mouse) and incubated for 5 minutes. Sections were washed with PBS to stop the reaction. Finally, a counterstain for haematoxylin was performed to visualise the nuclei. After the counterstain sections were dehydrated and mounted with DPX.

Histological image analysis

Sections were scanned on the Aperio ScanScope AT Turbo device at a 20x magnification. All images were automatically segmented making use of great colour contrast between the DAB stain and the haematoxylin counterstain. A specific colour cluster for each structure of interest was empirically defined in HSV colour-space using MATLAB's interactive colour thresholder application (MATLAB and Statistics Toolbox Release 2017a, The MathWorks, Inc., Natick, Massachusetts, United States). The saturation and value thresholds were also optimised per stain so that the segmentations would represent the DAB-staining without much contribution of background staining. Pixels that fell within the colour segmentation were considered to be positive.

To quantify the amount of staining, a stained area fraction was calculated as the relative number of positive pixels within a local neighbourhood of 100x100 pixels. These area fraction maps were used to get an estimate of the histological structures, for example, myelin density, in a certain region of interest. The region of interest, namely, the perforant path, was manually drawn in all histological sections based on the tractography data, for each subject and stain separately (Fig. 12A).

Statistics

All statistical analyses were performed with IBM SPSS statistics for Windows, version 22 Armonk, NY: IBM Corp.

Diffusion MRI.

A students *t*-test was used to evaluate the differences in MD, FA, AD and RD values between groups. We tested one-sided since we expect the MD, AD and RD values to be higher and the FA values to be lower in ALS cases. Between the groups, there was a significant difference in fixation time ($p=0.003$). Since it is known that fixation time can cause a reduction in FA, MD and AD and an increase in RD (D'Arceuil & de Crespigny, 2007; Schmierer et al., 2008, 2007; Sun, Liang, Xie, Oyoyo & Lee, 2009) a chronic demyelinating disease. Conventional T2-weighted MRI (cMRI, a normalization step was performed before calculating the t-statistics. The normalisation was done by dividing the found FA, MD, AD and RD values in the perforant path by a reference values for each of these metrics in a control region within the same sample. The entorhinal cortex was used as a control region.

Polarized light imaging

A students T-test was used to evaluate the differences in retardance values within the perforant path between ALS cases and controls. We tested one-sided because we expected the retardance values to be lower in ALS cases compared to controls.

The relationship between the retardance and the normalised dMRI diffusion tensor metrics was assessed by calculating the Pearson correlation coefficient. The correlation was tested one sided, because we expected a negative correlation between the MD and retardance and a positive correlation between the FA and the retardance.

Histology

A students T-test was used to evaluate the differences in PLP, SMI-312, CD68 and pTDP-43 stained area fractions. Tissue fixation is known to reduce the availability of antigen binding sites and can therefore cause a decrease in the amount of staining (Werner, Chott, Fabiano & Battifora, 2000). Therefore, a normalisation was performed before calculating the T-test statistics. This normalization was only performed for the PLP and SMI-312 stain. The CD68 and TDP-43 stain were not normalised because inflammation and pTDP-43 pathology are expected to be widespread within the slice even if white matter degeneration is localised.

The normalisation was accomplished by dividing the stained area fraction in the perforant path by the stained area fraction in a control region within the same specimen. The EC was used as a control region. A Pearson correlation analysis was performed on the normalised dMRI and histology values.

Results

Magnetic Resonance Imaging

We aimed to identify the perforant path in all specimens using probabilistic tractography. After

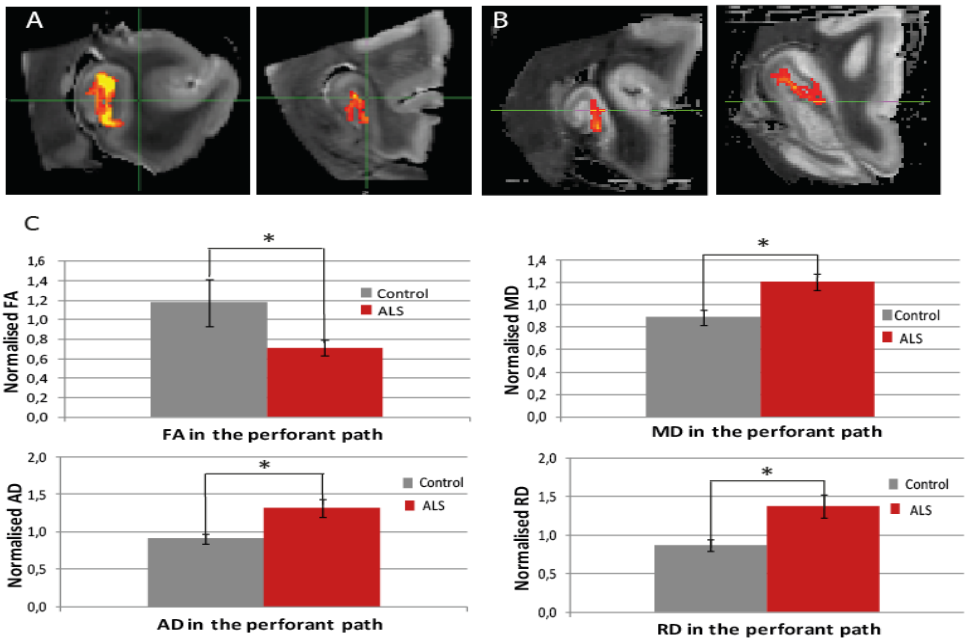


Figure 8. Examples of the tracts representing the perforant pathway running from the subiculum to the dentate gyrus for controls (A) and (B) for ALS cases. These tracts were used for computing the MD, RD, AD and FA values within the perforant path. (C) The FA value in the part of the perforant path that runs from the subiculum to the dentate gyrus is lower in ALS cases compared to controls (p = 0.014, T-test, one-sided). The MD value is significantly higher in the perforant path in the ALS specimen compared to controls (p = 0.011, T-test, one-sided) as well as the AD (p = 0.023, T-test, one-sided) and the RD (p = 0.028, T-test, one-sided) value.

running tractography analysis the identified tracts were inspected and exclusion masks were drawn in such a way that only the tracts representing the anatomical course of the perforant path were maintained for analysis (Fig. 8A).

To get an impression of white matter degeneration in the perforant path, the diffusion tensor metrics were compared between ALS cases and controls. Within the perforant path the FA values were significantly lower in ALS specimens compared to controls (p = 0.014), T-test one-sided). A reduced FA indicates that water diffuses more isotropically and could indicate myelin degeneration. This hypothesis is further strengthened by the increase that was found in MD in ALS specimens compared to controls (p = 0.011, T-test one-sided) (Fig. 8B and C). In addition, both RD (p = 0.028) and AD (p = 0.023) values were significantly higher in ALS cases compared to controls. An increase in RD value is caused by an increasing amount of diffusion along the radial diffusion direction, further implying loss of fibres.

Polarized light imaging

To get a detailed insight in the directionality and myelination of fibres within the tissue, PLI was performed. From the raw PLI data, the fibre orientation, transmittance, retardance and in-plane map were extracted. The transmittance map reflects the amount of light passing through the tissue. Grey matter lets through more light than white matter and therefore appears brighter in the transmittance map (Fig. 9A). The retardance map shows the degree to which the phase of the light is shifted due to interaction with the myelin and is therefore an indirect measure of myelin (Fig. 9B). The in-plane map gives the in-plane direction of the fibres with values ranging from 0 to π (Fig. 9C). The fibre orientation map is formed by combining the retardance and in-plane maps. It shows the direction of the myelinated fibres within the tissue, allowing for tracing the perforant path within each slice (Fig. 9D).

Retardance

To further investigate the amount of white matter in the perforant path in ALS, patients compared to controls and the PLI retardance maps were analysed. As previously mentioned, retardance measures how much a polarized light wave is shifted due to interaction with regularly arranged myelin sheets. The PLI retardance values within the perforant path were calculated and it was shown that the retardance values in ALS cases were lower, implicating a loss of myelin. However, the decrease was not significant (p

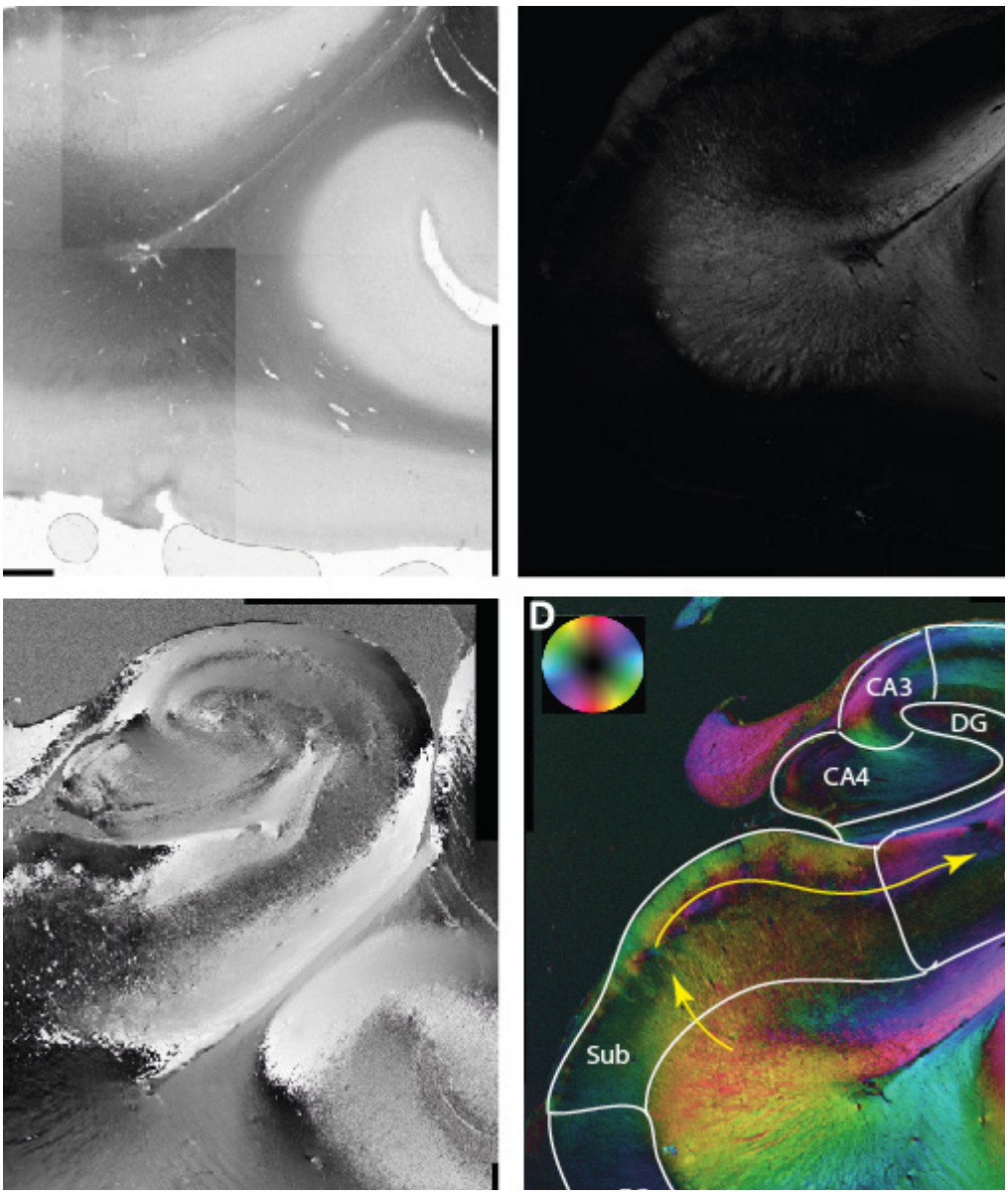


Figure 9. Example of the polarised light imaging data from a control case. (A) The transmittance map reflecting the amount of light passing through the tissue. (B) The retardance map showing the degree to which the light is shifted due to interaction with the myelin. (C) The in-plane map giving the in-plane direction of the fibres with values ranging from 0 to π. (D) The fibre orientation map showing the directionality of the myelinated fibres indicated with the colour wheel. This map is formed by combining the retardance and in-plane map. The anatomical regions within the hippocampus are delineated and the approximate direction of the perforant path is indicated with the yellow arrows.

= 0.16, one-sided; Fig. 10A).

Since PLI is a relatively new technique and to our knowledge not applied to pathological specimens before, we aimed to gain insight in the relationship between the dMRI metrics and the PLI retardance values. The PLI values were therefore correlated with the FA and MD. As expected, there was a positive correlation between the retardance and the FA ($p = 0.038$, one-sided). No correlation was found between the retardance and the MD ($p = 0.48$, one-sided; Fig. 10C). We, however, expected a negative correlation between the MD and retardance because an increase in MD is thought to reflect a decrease in white matter structure.

Dispersion

As a result of fibre degeneration, orientation dispersion may be a marker for the disturbed coherence of fibres. Indeed, the fibre dispersion was elevated in ALS cases compared to controls ($p = 0.044$), demonstrating an increase in the fanning of fibres within the perforant path (Fig. 11C). These results further implicate damage of the perforant path.

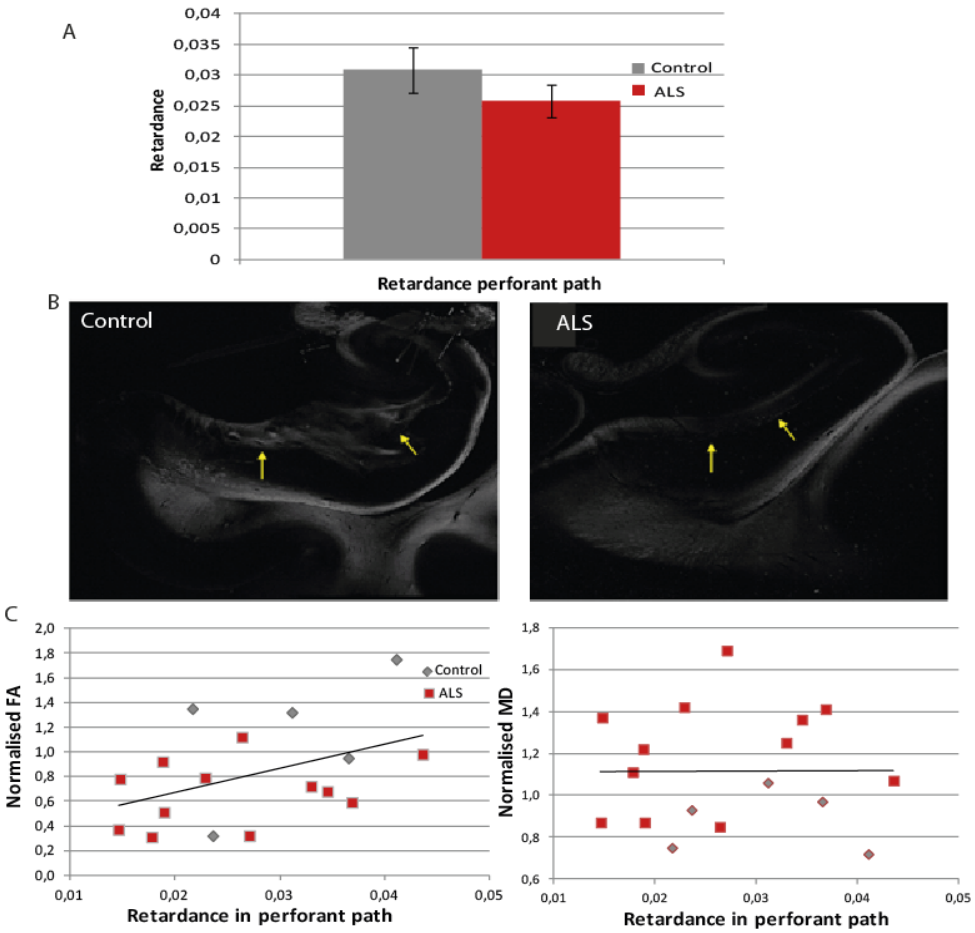


Figure 10. (A) The retardance value within the perforant path was found to be lower in ALS cases compared to controls, but was not significant (p = 0.16 one-sided). (B) Two examples of retardance maps from an ALS case (left) and control (right). The yellow arrows indicate the approximate location of the perforant path. (C) Correlation between the MD and FA values from the dMRI and retardance values from PLI for cases and controls combined. There was a significant correlation between the FA and the retardance (p = 0.038 one-sided), but not between the MD and the retardance (p = 0.48, one-sided).

Immunohistochemistry

While dMRI and PLI provide valuable information regarding white matter architecture in our specimens, protein pathological information within the tissue is lacking from these techniques. To gain more insight in the protein pathology, we performed immunohistochemical stains for pTDP-43, microglial activation (CD68), myelin (PLP) and neurofilaments (SMI-312).

The stained area fraction was calculated using an automatic segmentation for all stains. An example of each of these segmentations is shown in Figure 12. The segmentation for PLP and CD68 performed well, meaning that it picked up the true DAB-staining without including the background signal. For SMI-312, the segmentation proved to be more challenging due to a relatively low contrast between the background and true stain. To make sure it would not pick up any background signal, the segmentation settings were set to be more conservative. For this reason, the algorithm sometimes missed out on some of the true staining. To develop an automatic segmentation that correctly identified the pTDP-43 inclusions, it was necessary to set a minimum cluster size threshold of 15 pixels. After setting the size threshold, the amount of pTDP-43 inclusions was in general correctly quantified. However, the algorithm still picked up some degree of non-specific staining of pigments within the nucleus as can be observed in supplementary Figure 3, resulting in some false positive signal.

After determining the segmentation settings, the stained area fraction was computed for each stain in both cases and controls. A higher amount of myelin (PLP) was found in ALS cases compared to controls. However, this difference was not significant (p = 0.11). Furthermore, an increase in the amount of neurofilaments (SMI-312) was detected in ALS cases

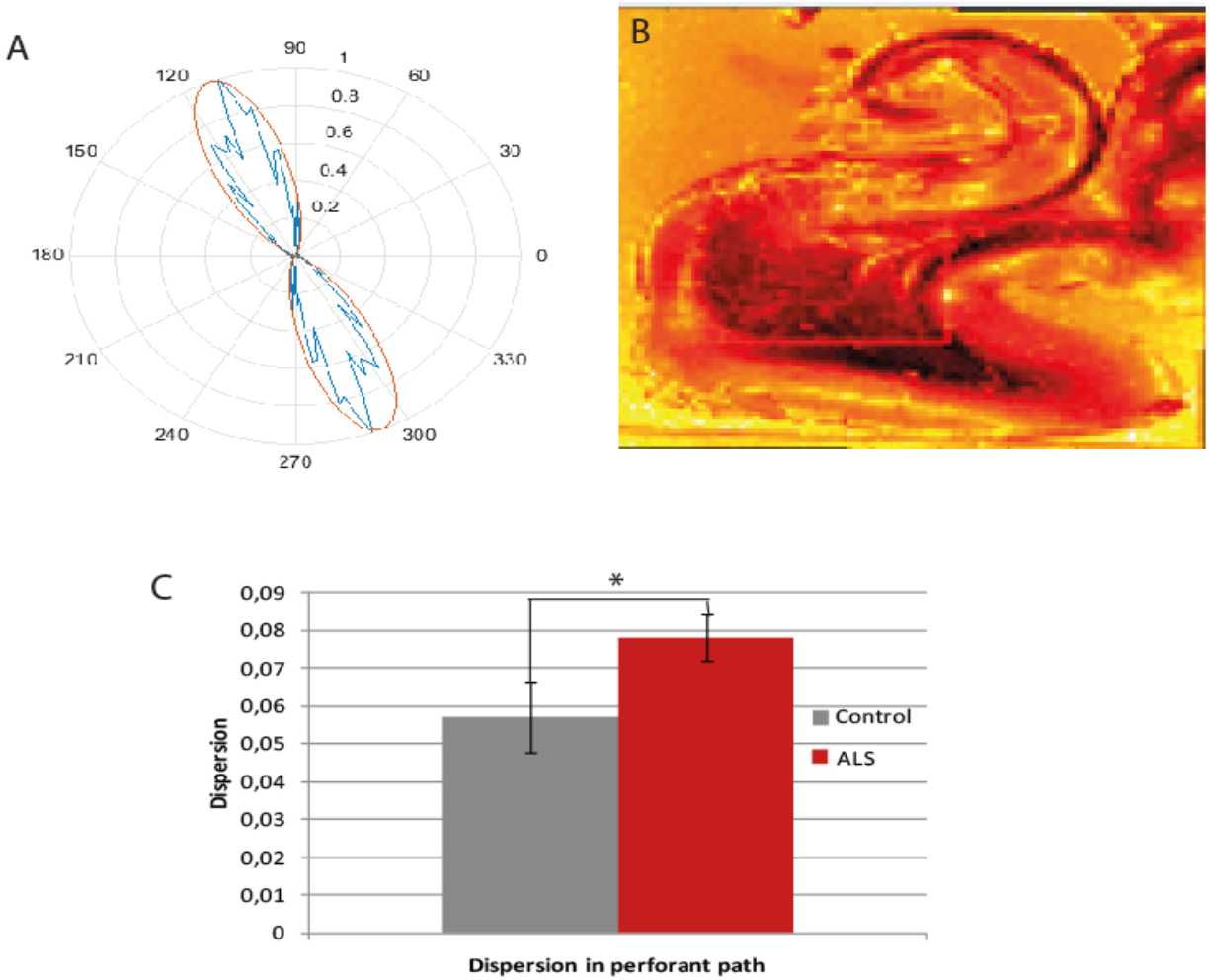


Figure 11. (A) A polar plot illustrating the raw PLI fibre orientation distribution (blue) from an exemplar neighbourhood of 50x50 pixels fitted with a von Mises distribution (red). (B) An example of the log(1/κ)-map. K is a measure for the broadness of the Von-Mises distribution, the higher the K value, the narrower the Von-Mises distribution. A narrow Von Mises distribution is associated with a low level of dispersion. Therefore, 1/κ is a measure for the amount of dispersion. Log (1/κ) was shown to increase visibility of the dispersion values. It can be observed that dispersion is lower in highly aligned white matter regions. (C) The amount of dispersion is higher in ALS cases compared to controls within the perforant path (p = 0.044).

(p = 0.03). These results contradict our previous findings with dMRI and PLI since they suggested a decrease in myelination and axonal integrity in ALS cases compared to controls. No changes were observed in the amount of CD68, indicating similar levels of inflammation in both groups. Finally, it was shown that two out of the 13 ALS cases exhibited pTDP-43 pathology within the hippocampus (Fig. 13B).

To investigate the relationships between PLI and SMI312 staining and the dMRI metrics, correlation analyses were performed. As expected, a positive correlation was present between the amount of myelin and the amount of neurofilaments (p = 0.025, R^2 = 0.499). An increase in MD and RD has previously been associated with demyelination. A negative correlation was therefore expected between PLP and MD/RD. However, contradictory to our hypothesis, a positive correlation was found between PLP and MD (p = 0.037, R^2 = 0.458), PLP and RD (p = 0.023, R^2 = 0.506) and PLP and AD (p = 0.021, R^2 = 0.514). The amount of SMI-312 stain did not correlate with any of the diffusion measures.

Discussion

Since ALS is mostly known as a motor disease, multiple studies investigating white matter alterations in ALS have focused on changes in the cortico-spinal tract and the corpus callosum (Cirillo et al., 2012; Filippini et al., 2010; Horsfield & Jones, 2002; Lillo et al., 2012). The clinical involvement of both compartments is characteristically variable and the site of onset debated. We sought to establish

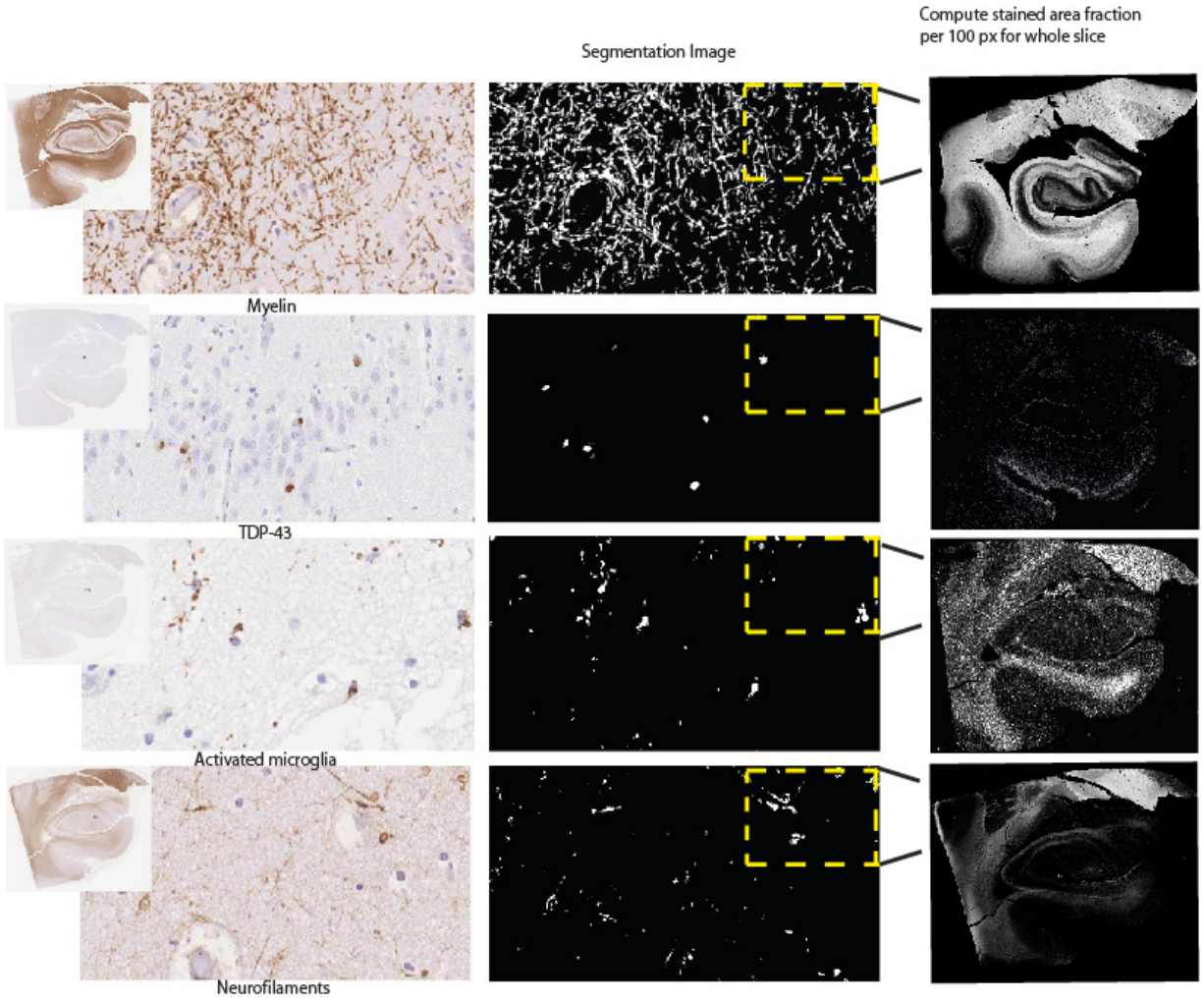


Figure 12. Overview of the histological data and analysis. The left column shows a snapshot of the original staining for myelin, TDP-43, activated microglia and neurofilaments. The middle column depicts the automatic segmentation of this snapshot where white indicates a pixel that is considered to be stained and black indicates a pixel that is considered to be background. After segmentation, the stained area fraction was calculated per 100 pixels resulting in a down-sampled image reflecting the amount of staining over the slice (yellow square). The right column gives an overview of the segmentation across the entire slice.

whether there is a consistent signature of cerebral white matter abnormalities in heterogeneous ALS cases. In this observational study, diffusion tensor imaging was applied in a whole-brain analysis of 24 heterogeneous patients with ALS and well-matched healthy controls. Tract-based spatial statistics were used, with optimized voxel-based morphometry of T1 images to determine any associated gray matter involvement. A consistent reduction in fractional anisotropy was demonstrated in the corpus callosum of the ALS group, extending rostrally and bilaterally to the region of the primary motor cortices, independent of the degree of clinical upper motor neuron involvement. Matched regional radial diffusivity increase supported the concept of anterograde degeneration of callosal fibers observed pathologically. Gray matter reductions were observed bilaterally in primary motor and supplementary motor regions, and also in the anterior cingulate and temporal lobe regions. A post hoc group comparison model incorporating significant values for fractional anisotropy, radial diffusivity, and gray matter was 92% sensitive, 88% specific, with an accuracy of 90%. Conclusion: Callosal involvement is a consistent feature of ALS, independent of clinical upper motor neuron involvement, and may reflect independent bilateral cortical involvement or interhemispheric spread of pathology. The predominantly rostral corticospinal tract involvement further supports the concept of independent cortical degeneration even in those patients with ALS with predominantly lower motor neuron involvement clinically. Over the past years, ALS has become more and more recognized as a multisystem neurodegenerative disease because areas of the brain outside the motor-areas also undergo

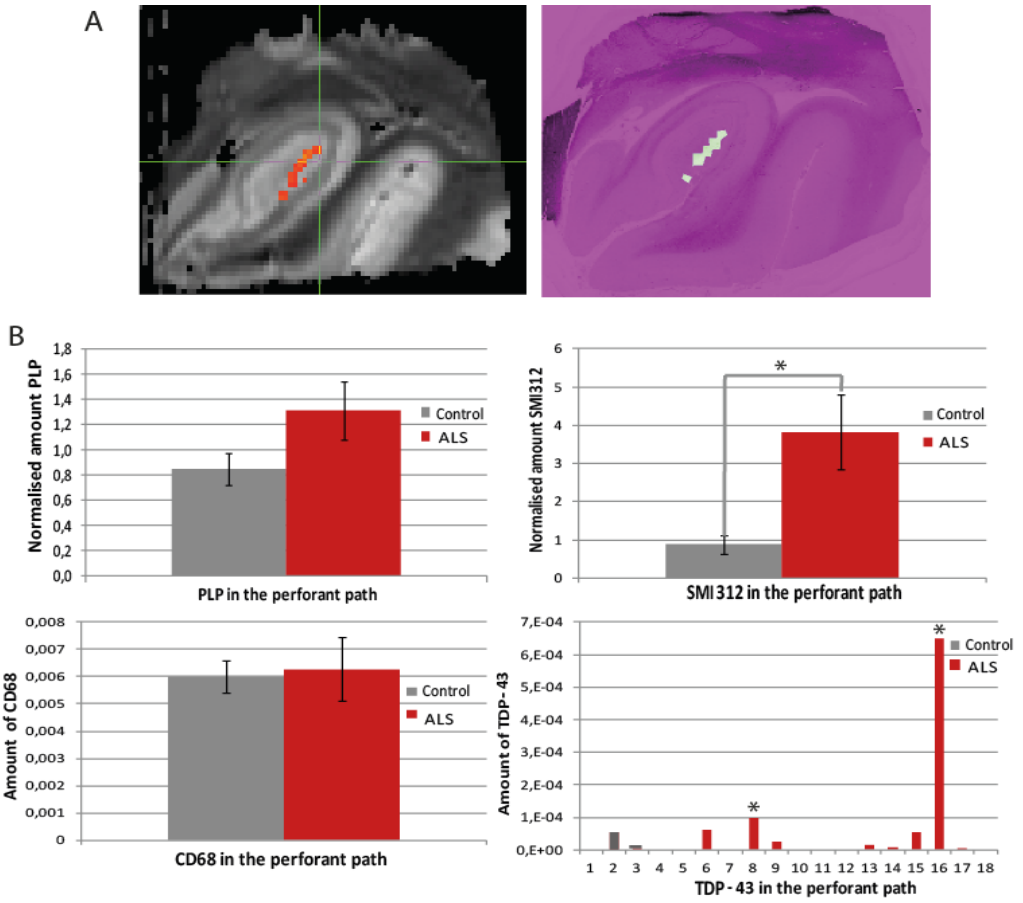


Figure 13. (A) Example of manual delineation of the perforant path (right) based on the dMRI tractography results (left). (B) The amount of PLP within the perforant path is higher in ALS cases compared to controls, but not significantly (p=0.11). Furthermore, an increase in the amount of neurofilaments is observed in ALS cases (p = 0.03). The amount of inflammation (CD68) is similar in both groups and two out of the 18 cases exhibit pTDP-43 pathology within the perforant path.

degeneration. Furthermore, around 50% of the ALS patients exhibit cognitive problems with 10-20% of the people meeting the criteria for bvFTD. Despite the relationship between ALS and memory deficits, detailed investigation of white matter tracts related to memory has been lacking to date. Here, we used dMRI, PLI and histology to investigate white matter changes in the perforant path, a projection of the circuit of Papez that has been implicated to be affected in ALS and bvFTD and was previously linked to cognitive deficits in Alzheimer's disease (Hyman et al., 1986; Thal et al., 2000).

Using dMRI, we demonstrated a reduction in the FA and an increase in the MD, AD and RD in the perforant path in ALS cases compared to controls (Fig. 8B and C), suggesting white matter damage. PLI retardance values were also lower in ALS specimens, however, a significant difference was not detected (Fig. 10A). The retardance did correlate with the FA, but not with the value MD within the perforant path (Fig. 10C). The amount of dispersion was significantly higher in ALS specimens, further indicating damaged axonal structure (Fig. 11C). Histochemical analysis showed no differences in the amount of microglial activation. However, they did show a (non-significant) increase in myelin and a significant increase in the amount of neurofilaments. There was a positive correlation between PLP and MD and RD and AD dMRI values. pTDP-43 pathology was found in two out of the 14 ALS cases within the hippocampus (Fig. 13B).

In order to extract the FA, MD, AD and RD values from dMRI within the perforant path, probabilistic tractography was performed. At first, tractography was performed by using the EC as a seed, keeping all paths terminating in the DG. With these settings we were unable to track the perforant pathway. Next, we tried to extract the perforant path by using the EC as a seed and the Sub as waypoint, again only maintaining the paths terminating at the DG. With these settings, it was also not possible to extract a path representing the anatomical course of the perforant path (Supplementary Fig. 1). This is in concordance with Augustinack and colleagues (2010), who also reported to be unable to track the whole perforant pathway from the EC to DG using tractography.

Third, we decided to divide the perforant path into two parts: One part from the EC to the Sub and the other part from the Sub to the DG. However, in the area from the EC to the Sub, two other fibre bundles, the alvear path and the collateral white matter, are crossing the perforant path (Supplementary Fig. 5). This makes it difficult to solely extract the perforant path for analysis in this region. Furthermore, the perforant path is originating from across the entire entorhinal cortex forming a coherent bundle only from the subiculum on. Since we are only interested in the dMRI, PLI and histology metrics within the perforant path, we decided to solely focus on the part of the path running from the Sub to the DG.

Part of the changes that were found in the dMRI metrics in ALS patients compared to controls have been linked to a loss of white matter integrity. Namely, a reduction in FA and an increased MD and RD values were repeatedly associated with white matter degeneration (Bozzali et al., 2002; Soares, Marques, Alves & Sousa, 2013). Interpretation of the increase in AD is less arbitrary as some studies found no changes in AD due to demyelination (Klawiter et al., 2011; Song et al., 2002), while other studies either find an increase in AD (Metwalli et al., 2010) or a decrease in AD due to white matter loss (Harsan et al., 2006; Tyszka et al., 2006). Possibly, an increase in AD is observed due to enlargement of the extracellular space as a result of axonal degeneration.

While the dMRI results suggest that white matter in the perforant path is damaged in ALS cases, changes in the dMRI metrics can also be driven by several other factors. Since dMRI metrics are calculated based on the diffusion of water, assuming that water diffuses along the line of an axon, they are indirect measures of the white matter geometry. For example, two kissing or crossing fibres could yield an FA value of 0 implying fully isotropic diffusion and no axons being present, not reflecting the actual white matter structure within that voxel (Jbabdi & Johansen-Berg, 2011). However, large differences between cases and controls in the amount of crossing fibres are not expected since we are investigating the same pathway in a similar area in both groups

Besides a change in white matter geometry, the diffusion metrics can also be affected by the way tissue is preserved after death. Several of the ALS cases had post–mortem intervals (PMI) of up to five days. Previous research has shown that long PMI can cause a reduction in FA and an increase in MD values (D'Arceuil & de Crespigny, 2007; Miller et al., 2011) and as a link between in vivo diffusion studies and \"gold standard\" histology/dissection. While there is a relatively mature literature on post mortem diffusion imaging of animals, human brains have proven more challenging due to their incompatibility with high-performance scanners. This study presents a method for post mortem diffusion imaging of whole, human brains using a clinical 3-Tesla scanner with a 3D segmented EPI spin-echo sequence. Results in eleven brains at 0.94. 0.94. 0.94. mm resolution are presented, and in a single brain at 0.73. 0.73. 0.73. mm resolution. Region-of-interest analysis of diffusion tensor parameters indicate that these properties are altered compared to in vivo (reduced diffusivity and anisotropy) but there are relatively few studies in human brains. While animal tissues are generally fixed pre-mortem or directly post-mortem, this is not possible for human tissue, therefore there is always some delay between death and tissue fixation. The elapsed time between death and tissue fixation, the post-mortem interval (PMI). In our data, a significant negative correlation between the PMI and the MD values ($p = 0.009$) was observed. No significant correlation was found between the FA ($p = 0.242$) or retardance ($p = 0.647$) and the PMI. Another factor that can influence the dMRI values is the fixation time. It has previously been shown that fixation causes a reduction in diffusivity within the tissue, leading to a drop in FA, MD, AD and RD (D'Arceuil & de Crespigny, 2007; Sun et al., 2009, 2005) but relatively few studies in human brains. While animal tissues are generally fixed pre-mortem or directly post-mortem, this is not possible for human tissue, therefore there is always some delay between death and tissue fixation. The elapsed time between death and tissue fixation, the post-mortem interval (PMI. Because the fixation time between groups differs significantly ($p = 0.003$), it is important to correct for these effects. Based on the assumption that any effects of PMI and fixation time are present in the whole hippocampal tissue block, an internal control region was used to normalise the FA, MD, AD and RD values within the perforant path.

First, the collateral white matter was chosen as a control region as this white matter region was present in all sampled hippocampal blocks. However, an (non-significant) increase in the MD value was observed within the collateral white matter ($p = 0.15$), while the expected effect of fixation time was in the opposite direction. In addition, the perforant path was partially crossing through the collateral white matter area (Supplementary Fig. 5), making the collateral white matter an unsuitable control region. Unfortunately, there was no other white matter region that was present in all sampled

hippocampal blocks and that was not part of the circuit of Papez. Therefore, the EC was used as a control region. The perforant path originates here but very few fibres of perforant path run through this area. After normalisation, a significant decrease in the FA value ($p = 0.014$) and a significant increase in the MD value ($p = 0.011$), AD value ($p = 0.021$) and RD value ($p = 0.028$) was found in ALS cases compared to controls (Fig. 8C). For future research, specimens have to be matched on fixation time, PMI, age, gender and region to ensure that any observed differences are driven by case-control differences.

The dMRI results implicate white matter loss in the perforant path in ALS but can, as previously mentioned, also be caused by a difference in fibre configuration between the groups. We performed PLI and histology to gain further insight in the nature of the changes found with dMRI.

The PLI retardance maps were used to compare the degree of myelination between patients and controls. This is, to our knowledge, the first time that PLI was used to make a case-control comparison. A decrease in the retardance was observed in ALS cases compared to controls within the perforant path. This decrease was, however, not significant ($p = 0.16$, one-sided, Fig. 10A).

In addition to the birefringent properties of myelinated axons, the retardance is also dependent on the inclination angle of fibres. A reduction in retardance is for example observed when equally myelinated fibres are propagating less radial to the imaging plane (Axer et al., 2011)the structural basis of the human connectome. In contrast to animal brains, where a multitude of tract tracing methods can be used, magnetic resonance (MR). A small tilt in the plane could have been induced during cutting the sections or during sampling of the hippocampal blocks from the brains, leading to a variation in the retardance values.

We were unable to perform a fixation time correction on the retardance values because the exact effects of fixation time on retardance are unknown. Retardance measures the birefringence induced by the regular arrangement of myelin proteins and lipids. During formalin fixation, proteins are cross-linked keeping protein structures relatively intact (Thavarajah, Mudumbaimannar, Rao, Ranganathan & Elizabeth, 2012). It has been shown that there is a loss of lipids in the central nervous system during fixation, possibly resulting in lower retardance values when fixation time is increasing (Hopwood, 1969). However, no significant negative correlation between fixation time and retardance values was found ($R^2 = 0.306$, $p = 0.109$, one-sided). Another inadequacy is that the effect of aggregations of proteins like pTDP-43 on the retardance signal have never been investigated.

To gain further insight in the relationship between retardance and dMRI metrics, a Pearson correlation analyses was performed. There was a positive correlation between the FA and retardance, meaning that a higher retardance value was associated with a more isotropic water diffusion direction and supposedly a more intact white matter architecture. Surprisingly, no negative correlation between the MD and retardance was observed. It was expected that when the MD value rises due to loss of axonal integrity, the retardance values would go down. Potentially, the decrease in MD value is more a result of loss of coherence in fibre structure than of demyelination and severe degeneration of axons.

To further investigate the degree of myelination within the perforant path, immunohistochemical stainings for proteolipid protein (PLP) were performed. Myelin consists of a high proportion of lipids (75-80%) and consequently a lower portion of proteins (15-30%). PLP and myelin basis protein (MBP) together make up about 60-80% of the proteins present in myelin (Brady & Siegel, 2012). Since PLP is one of the main components of myelin, a reduction in PLP is expected when myelin is degrading (Garbern et al., 2002). Instead, our results indicate a (non-significant) increase in PLP in the subiculum in ALS cases compared to controls ($p = 0.11$, one-sided). This result is contradicting the previously discussed dMRI and PLI results suggesting a reduction in white matter. Possibly, the change in white matter structure is driven by a reduction in MBP or lipids, unrelated to the amount of PLP. Another reason for the discrepancy between the dMRI and PLI and the PLP stain could be that axonal degeneration is occurring in the absence of demyelination (Garbern et al., 2002). As a result, axons would proceed to disentangle, increasing the amount of myelin-stained surface. The spreading of axonal fibres is known to be the begin stage of Wallarian degeneration (Waller, 1850). The significant increase in the spreading of fibres (dispersion, $p = 0.044$) and the positive correlation between the MD, RD and AD values and the PLP stain further substantiates this hypothesis.

An increase in the amount of SMI-312 stain was observed, again contradicting our hypothesis and the dMRI and PLI results. However, several previous papers report that during axonal degeneration in ALS swollen axon segments occur that are rich of the neurofilament proteins that SMI-312 stains for (Delisle & Carpenter, 1984; Lingor et al.,

2012) inflammatory and degenerative neurological diseases. Increasing evidence suggests that axonal degeneration occurs early in the course of these diseases and therefore represents a promising target for future therapeutic strategies. We review the evidence for axonal destruction from pathological findings and animal models with particular emphasis on neurodegenerative and neurotraumatic disorders. We discuss the basic morphological and temporal modalities of axonal degeneration (acute, chronic and focal axonal degeneration and Wallerian degeneration. The increase in SMI-312 stain could thus be caused by accumulation of neurofilaments in these axonal swellings. Another possible rationale behind the increase in neurofilaments is that accumulation of phosphorylated neurofilaments is often observed in the amyloid beta plaques associated with aging and Alzheimer's disease (Dale & Garcia, 2012; Dickson, King, McCormack & Vickers, 1999). The degree of amyloid beta like pathology should be taken into account as soon as the diagnostic stains are fully evaluated.

pTDP-43 pathology was detected in only two out of the 13 ALS cases within the hippocampal area. In one case, mild pathology was observed; in the other case, the pathology was classified as frequent (Fig. 3). Tan and colleagues (2015) developed a staging system for pTDP-43 pathology in ALS and FTD and it is shown that the hippocampal area is one of the last areas to be affected by pTDP-43 pathology in ALS. It is thus not surprising that not all specimens exhibit hippocampal pathology. In bvFTD however, the hippocampus is one of the first areas to be enforced by pTDP-43.

The histology data showed no differences in the amount of activated microglia, a measure for inflammation, in ALS cases compared to controls. This contradicts earlier findings of enhanced microglial activation in white matter lesions in ALS in the corpus callosum and the spinal cord (Henkel et al., 2004; Sugiyama et al., 2013). However, the average age of the control specimens was 70.8 years. It is known that the amount of activated microglia increases upon aging, potentially making it harder to detect significant differences in microglial activation between groups (Sparkman & Johnson, 2008). In our data, we did not find a correlation between activated microglia and age. Until now, only one section per subject was stained and analysed. Increasing sample size could result in a significant difference in CD68 staining between the groups.

The fact that only one section per subject was stained for each histochemical stain also resulted in a regional difference between the control and ALS cases. On average, the control sections were more representative of the posterior part of the hippocampus while the ALS sections represented a more anterior part (supplementary Fig. 3). Unfortunately, these regional differences were not only present in the single histology slices but also in the sampled hippocampal blocks. The control blocks were sampled in Nijmegen, The Netherlands, while the ALS blocks were sampled in Oxford, United Kingdom. Despite the fact that both groups used the LGN as an anatomical reference point for sampling the blocks, the control blocks were sampled more posterior than the ALS blocks. It could thus be that the observed dMRI, PLI and histology differences were driven by regional differences rather than by case-control variation. It is, however, hard to correct for this.

To allow further research of white matter changes in the entire circuit of Papez, other regions of the circuit were also sampled from all specimens. Unfortunately, we were not able to use ALS/FTD or pure bvFTD specimens for this research. It would be interesting to see whether the observed white matter changes correlate with clinically observed cognitive decline. Since the magnitude of all dMRI metrics is going down as a result of fixation, it is also interesting to perform *in vivo* dMRI of the perforant path in controls, ALS/ALS-FTD and bvFTD patients. This is challenging because the perforant path is small and therefore hard to detect with tractography analysis. Because we used post-mortem tissue, we were able to work with long scanning times (3-4 hours) and an ultra-high field MRI scanner (11.7T) resulting in very high resolution dMRI. Routine clinical MRI scanners have a field strength of 1.5-3T and scanning time would be maximally an hour, drastically reducing the achievable resolution. Another interesting target of future research is the effect of fixation time and protein aggregation on retardance measures. Since PLI is a new technique, this remains largely unknown so far.

Our current results linked the perforant path, a memory related white matter tract, to ALS. This research therefore contributes to the hypothesis of ALS being a multisystem neurodegenerative disease. Till now, the main focus in ALS research has been on motor neuron degeneration and degeneration of motor related areas and pathways. Instead, our research has focussed on a potential neural substrate of the cognitive problems that are observed in 50% of the ALS patients. We aimed to shed more light on the cause of these memory related problems and thereby hope to raise awareness about this under exposed aspect of the disease. Hopefully, our results

will contribute to faster recognition and alleviation of cognitive disabilities in ALS patients in the clinic.

Furthermore, our research substantiates the hypothesis that ALS and bvFTD are a spectrum of diseases rather than two separate disorders, by linking a pathway that is known to be involved in bvFTD to ALS. Future studies should therefore focus on investigating the perforant path in bvFTD and ALS/bvFTD patients to see if damage to the perforant path directly correlates with the degree of cognitive deficits in these diseases.

Conclusion

This work investigated white matter changes in the perforant path in ALS patients using dMRI, PLI and histological analyses. A reduction in the FA and an increase in the MD, AD and RD observed in ALS cases relative to controls, implicating white matter degeneration of the perforant path. This observation was supported by a (non-significant) decrease in the PLI retardance values and increased dispersion in the perforant path. However, histology results showed an increase in PLP and neurofilaments, potentially due to disentangling of axonal fibres and neurofilamental aggregation in axonal swellings, respectively.

We conclude that the axonal integrity of the perforant path is damaged in ALS cases compared to controls. Our worked hereby provides further proof of ALS being a multisystem disease and therefore contributes to the understanding of ALS. Since the perforant path is also implicated in bvFTD, we provide further evidence of ALS and bvFTD being a spectrum of diseases. Furthermore, we show that white matter degeneration of the perforant path likely occurs even before clinical diagnosis of cognitive symptoms. Finally, our results provide the perforant path as a potential neural correlate of the cognitive symptoms in ALS and further research should be done to substantiate this hypothesis.

Acknowledgements

We would like to acknowledge "Stichting Alzheimer Nederland" for their funding support. Furthermore, we would like to acknowledge P.J.W.C. Dederen for instructing in the cutting of the sections for PLI. Furthermore, we'd like to thank Dr. M. Wiessman for his excellent support with the MRI protocols and C. Grabitz and J. de Ruyter van Steveninck for their lab work. Finally, we acknowledge the Oxford brain bank as well as the body Donor program from the department of Anatomy, Radboud UMC Nijmegen for providing the post-mortem specimens.

References

Augustinack, J.C., Helmer, K., Huber, K.E., Kakunoori, S., Zöllei, L., & Fischl, B. (2010). Direct visualization of the perforant pathway in the human brain with ex vivo diffusion tensor imaging. *Frontiers in Human Neuroscience. 4*, 42.

Aung, W.Y., Mar, S., & Benzinger, T.L. (2013). Diffusion tensor MRI as a biomarker in axonal and myelin damage. *Imaging Med.*

Axer, M., Amunts, K., Grässel, D., Palm, C., Dammers, J., Axer, H., Pietrzyk, U., & Zilles, K. (2011). A novel approach to the human connectome: Ultra-high resolution mapping of fiber tracts in the brain. *Neuroimage. 54*, 1091–1101.

Barbagallo, G., Nicoletti, G., Cherubini, A., Trotta, M., Tallarico, T., Chiriaco, C., Nisticò, R., Salvino, D., Bono, F., Valentino, P., & Quattrone, A. (2014). Diffusion tensor MRI changes in gray structures of the frontal-subcortical circuits in amyotrophic lateral sclerosis. *Neurological Science. 35*, 911–918.

Bastiani, M., & Roebroeck, A. (2015). Unraveling the multiscale structural organization and connectivity of the human brain: the role of diffusion MRI. *Frontiers in Neuroanatomy. 9,* 77.

Behrens, T.E.J., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., & Woolrich, M.W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage 34,* 144–155.

Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., & Smith, S.M. (2003). Characterization and Propagation of Uncertainty in Diffusion-Weighted MR Imaging. *Magn. Reson. Med. 50*, 1077–1088.

Bonafede, R., & Mariotti, R. (2017). ALS Pathogenesis and Therapeutic Approaches: The Role of Mesenchymal Stem Cells and Extracellular Vesicles. *Frontiersin Cellular Neuroscience. 11*.

Bott, N.T., Radke, A., Stephens, M.L., & Kramer, J.H. (2014). Frontotemporal dementia: diagnosis, deficits and management. *Neurodegener. Dis. Manag. 4,* 439–54.

Bozzali, M., Falini, A., Franceschi, M., Cercignani, M., Zuffi, M., Scotti, G., Comi, G., & Filippi, M. (2002). White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. *J. Neurol. Neurosurg. Psychiatry* 72, 742–746.

Brady, S.T., & Siegel, G.J. (2012). Characteristic composition of myelin. Basic Neurochem. *Princ. Mol. Cell. Med. Neurobiol. 1096*.

Brettschneider, J., Del Tredici, K., Irwin, D.J., Grossman, M., Robinson, J.L., Toledo, J.B., Fang, L., Van Deerlin, V.M., Ludolph, A.C., Lee, V.M.Y., Braak, H., & Trojanowski, J.Q. (2014). Sequential distribution of pTDP-43 pathology in behavioral variant

frontotemporal dementia (bvFTD). *Acta Neuropathol. 127*, 423–439.

Brettschneider, J., Del Tredici, K., Toledo, J.B., Robinson, J.L., Irwin, D.J., Grossman, M., Suh, E., Van Deerlin, V.M., Wood, E.M., Baek, Y., Kwong, L., Lee, E.B., Elman, L., McCluskey, L., Fang, L., Feldengut, S., Ludolph, A.C., Lee, V.M.Y., Braak, H. & Trojanowski, J.Q., (2013). Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Ann. Neurol. 74,* 20–38.

Cirillo, M., Esposito, F., Tedeschi, G., Caiazzo, G., Sagnelli, A., Piccirillo, G., Conforti, R., Tortora, F., Monsurrò, M.R., Cirillo, S. & Trojsi, F., (2012). Widespread microstructural white matter involvement in amyotrophic lateral sclerosis: A whole-brain DTI study. *Am. J. Neuroradiol. 33,* 1102–1108. 8

Cykowski, M.D., Powell, S.Z., Peterson, L.E., Appel, J.W., Rivera, A.L., Takei, H., Chang, E. & Appel, S.H., (2017). Clinical Significance of TDP-43 Neuropathology in Amyotrophic Lateral Sclerosis. *J. Neuropathol. Exp. Neurol. 76,* 402–413.

D'Arceuil, H. & de Crespigny, A., (2007). The effects of brain tissue decomposition on diffusion tensor imaging and tractography. *Neuroimage, 36*, 64–68.

Dale, J.M., & Garcia, M.L. (2012). Neurofilament Phosphorylation during Development and Disease: Which Came First, the Phosphorylation or the Accumulation? *J. Amino Acids.*

Delisle, M.B., & Carpenter, S. (1984). Neurofibrillary axonal swellings and amyotrophic lateral sclerosis. *J. Neurol. Sci.*

Dickson, T.C., King, C.E., McCormack, G.H., & Vickers, J.C. (1999). Neurochemical diversity of dystrophic neurites in the early and late stages of Alzheimer's disease. *Exp. Neurol.*

Fen-Biao Gao& Sandra Almeida, R.L. (2017). Dysregulated molecular pathways in amyotrophic lateral sclerosis–frontotemporal dementia spectrum disorder. *Embo J.*

Ferrari, R., Kapogiannis, D., D. Huey, E., & Momeni, P.,(2011). F*TD and ALS: A Tale of Two Diseases. Curr. Alzheimer Res. 8*, 273–294.

figure Papez circuit [WWW Document], n.d.

Filippini, N., Douaud, G., MacKay, C.E., Knight, S., Talbot, K., & Turner, M.R. (2010). Corpus callosum involvement is a consistent feature of amyotrophic lateral sclerosis. *Neurology, 75,* 1645–1652.

Garbern, J.Y., Yool, D.A., Moore, G.J., Wilds, I.B., Faulk, M.W., Klugmann, M., Nave, K.A., Sistermans, E.A., Van Der Knaap, M.S., Bird, T.D., Shy, M.E., Kamholz, J.A., & Griffiths, I.R. (2002). Patients lacking the major CNS myelin protein, proteolipid protein 1, develop length-dependent axonal degeneration in the absence of demyelination and inflammation. *Brain, 125*, 551–561.

Ghosh, S., Lippa, C.F. (2015). Clinical Subtypes of Frontotemporal Dementia. *Am. J. Alzheimers. Dis. Other Demen. 30,* 653–661.

Harsan, L.A., Poulet, P., Guignard, B., Steibel, J., Parizel, N., De Sousa, P.L., Boehm, N., Grucker, D., & Ghandour, M.S. (2006). Brain dysmyelination and recovery assessment by noninvasive in vivo diffusion tensor magnetic resonance imaging. *J. Neurosci. Res. 83,* 392–402.

Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F. V., Brady, S.M., &Schnabel, J.A. (2012). MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Med. Image Anal.*

Henkel, J.S., Engelhardt, J.I., Sikls, L., Simpson, E.P., Kim, S.H., Pan, T., Goodman, J.C., Siddique, T., Beers, D.R., & Appel, S.H. (2004). Presence of Dendritic Cells, MCP-1, and Activated Microglia/Macrophages in Amyotrophic Lateral Sclerosis Spinal Cord Tissue. *Ann. Neurol. 55*, 221–235.

Hopwood, D., (1969). Fixatives and fixation: a review. *Histochem. J.*

Hornberger, M., & Piguet, O. (2012). Episodic memory in frontotemporal dementia: a critical review. *Brain 135,* 678–92.

Hornberger, M., Wong, S., Tan, R., Irish, M., Piguet, O., Kril, J., Hodges, J.R., & Halliday, G. (2012). In vivo and post-mortem memory circuit integrity in frontotemporal dementia and Alzheimer's disease. *Brain 135*, 3015–3025.

Horsfield, M.A., & Jones, D.K. (2002). Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases - A review. *NMR Biomed.*

Hu, W.T., & Grossman, M. (2009). TDP-43 and frontotemporal dementia. *Curr. Neurol. Neurosci. Rep.*

Huszar IN, Miller KL, Pallebage-Gamarallage M, Ansorge O, & Mirfin C, H.M. (2018). Pipeline for registering histological sections to MRI volumes; *abstract ISMRM conference Paris., in: Pipeline for Registering Histological Sections to MRI Volumes.*

Hyman, B.T., Van Hoesen, G.W., Kromer, L.J., & Damasio, A.R. (1986). Perforant pathway changes and the memory impairment of Alzheimer's disease. *Ann. Neurol.*

Irish, M., Devenney, E., Wong, S., Dobson-Stone, C., Kwok, J.B., Piguet, O., Hodges, J.R., & Hornberger, M. (2013). Neural substrates of episodic memory dysfunction in behavioural variant frontotemporal dementia with and without C9ORF72 expansions. *NeuroImage Clin.*

Jbabdi, S., & Johansen-Berg, H. (2011). Tractography: Where Do We Go from Here? *Brain Connect. 1,* 169–183.

Jbabdi, S., Sotiropoulos, S.N., Haber, S.N., Van Essen, D.C., & Behrens, T.E. (2015). Measuring macroscopic brain connections in vivo. *Nature Neuroscience. 18*, 1546–1555.

Jbabdi, S., Sotiropoulos, S.N., Savio, A.M., Graña, M., & Behrens, T.E.J. (2012). Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. *Magn. Reson. Med. 68,* 1846–1855.

Kerschensteiner, M., Schwab, M.E., Lichtman, J.W., & Misgeld, T. (2005). In vivo imaging of axonal

degeneration and regeneration in the injured spinal cord. *Nat. Med.*

Klawiter, E.C., Schmidt, R.E., Trinkaus, K., Liang, H.F., Budde, M.D., Naismith, R.T., Song, S.K., Cross, A.H., & Benzinger, T.L. (2011). Radial diffusivity predicts demyelination in ex vivo multiple sclerosis spinal cords. *Neuroimage, 55,* 1454–1460.

Larsen, L., Griffin, L.D., Gräßel, D., Witte, O.W., & Axer, H. (2007). Polarized light imaging of white matter architecture. *Microsc. Res. Tech. 70*, 851–863.

Lillo, P., Mioshi, E., Burrell, J.R., Kiernan, M.C., Hodges, J.R., Hornberger, & M. (2012). Grey and White Matter Changes across the Amyotrophic Lateral Sclerosis-Frontotemporal Dementia Continuum. *PLoS One, 7.*

Lingor, P., Koch, J.C., Tönges, L., & Bähr, M. (2012). Axonal degeneration as a therapeutic target in the CNS. *Cell Tissue Res.*

Logroscino, G., Traynor, B.J., Hardiman, O., Chio, A., Mitchell, D., Swingler, R.J., Millul, A., Benn, E., & Beghi, E. (2010). Incidence of amyotrophic lateral sclerosis in Europe. J. Neurol. Neurosurg. *Psychiatry, 81*, 385–390.

Mackenzie, I.R., & Rademakers, R. (2008). The role of transactive response DNA-binding protein-43 in amyotrophic lateral sclerosis and frontotemporal dementia. *Current Opinions in Neurology. 21,* 693–700.

Metwalli, N.S., Benatar, M., Nair, G., Usher, S., Hu, X., & Carew, J.D. (2010). Utility of axial and radial diffusivity from diffusion tensor MRI as markers of neurodegeneration in amyotrophic lateral sclerosis. *Brain Res. 1348,* 156–164.

Miller, K.L., Stagg, C.J., Douaud, G., Jbabdi, S., Smith, S.M., Behrens, T.E.J., Jenkinson, M., Chance, S.A., Esiri, M.M., Voets, N.L., Jenkinson, N., Aziz, T.Z., Turner, M.R., Johansen-Berg, H., & McNab, J.A. (2011). Diffusion imaging of whole, post-mortem human brains on a clinical MRI scanner. *Neuroimage, 57,* 167–181.

Mollink, J., Kleinnijenhuis, M., Cappellen van Walsum, A.M. van, Sotiropoulos, S.N., Cottaar, M., Mirfin, C., Heinrich, M.P., Jenkinson, M., Pallebage-Gamarallage, M., Ansorge, O., Jbabdi, S., & Miller, K.L. (2017). Evaluating fibre orientation dispersion in white matter: Comparison of diffusion MRI, histology and polarized light imaging. *Neuroimage.*

Montine, T.J., Phelps, C.H., Beach, T.G., Bigio, E.H., Cairns, N.J., Dickson, D.W., Duyckaerts, C., Frosch, M.P., Masliah, E., Mirra, S.S., Nelson, P.T., Schneider, J.A., Thal, D.R., Trojanowski, J.Q., Vinters, H. V., & Hyman, B.T. (2012). National institute on aging-Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease: A practical approach. *Acta Neuropathol. 123,* 1–11.

Pessoa, L., & Hof, P.R. (2015). From Paul Broca's great limbic lobe to the limbic system. J. *Comp. Neurol.*

Phukan, J., Elamin, M., Bede, P., Jordan, N., Gallagher, L., Byrne, S., Lynch, C., Pender, N., & Hardiman, O.(2012). The syndrome of cognitive impairment in amyotrophic lateral sclerosis: a population-based study. J. Neurol. Neurosurg. *Psychiatry, 83*, 102–8.

Schmierer, K., Wheeler-Kingshott, C.A.M., Boulby, P.A., Scaravilli, F., Altmann, D.R., Barker, G.J., Tofts, P.S., & Miller, D.H. (2007). Diffusion tensor imaging of post mortem multiple sclerosis brain. *Neuroimage, 35,* 467–477.

Schmierer, K., Wheeler-Kingshott, C.A.M., Tozer, D.J., Boulby, P.A., Parkes, H.G., Yousry, T.A., Scaravilli, F., Barker, G.J., Tofts, P.S., & Miller, D.H. (2008). Quantitative magnetic resonance of postmortem multiple sclerosis brain before and after fixation. 268–277.

Scotter, E.L., Chen, H.J., & Shaw, C.E. (2015). TDP-43 Proteinopathy and ALS: Insights into Disease Mechanisms and Therapeutic Targets. *Neurotherapeutics.*

Shah, A., Jhawar, S.S., & Goel, A. (2012). Analysis of the anatomy of the Papez circuit and adjoining limbic system by fiber dissection techniques. *Journal of Clininacal Neuroscience 19,* 289–298.

Shepherd, T.M., Thelwall, P.E., Stanisz, G.J., & Blackband, S.J.(2009). Aldehyde fixative solutions alter the water relaxation and diffusion properties of nervous tissue. *Magn. Reson. Med. 62*, 26–34.

Soares, J.M., Marques, P., Alves, V., & Sousa, N., (2013). A hitchhiker's guide to diffusion tensor imaging. *Frontiers in Neurosciense.*

Song, S.K., Sun, S.W., Ramsbottom, M.J., Chang, C., Russell, J., & Cross, A.H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage, 17*, 1429–1436.

Sparkman, N.L., & Johnson, R.W. (2008). Neuroinflammation associated with aging sensitizes the brain to the effects of infection or stress. *Neuroimmunomodulation.*

Sreedharan, J., Blair, I.P., Tripathi, V.B., Hu, X., Vance, C., Rogelj, B., Ackerley, S., Durnall, J.C., Williams, K.L., Buratti, E., Baralle, F., de Belleroche, J., Mitchell, J.D., Leigh, P.N., Al-Chalabi, A., Miller, C.C., Nicholson, G., & Shaw, C.E.(2008). TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science, 319,* 1668–72.

Sugiyama, M., Takao, M., Hatsuta, H., Funabe, S., Ito, S., Obi, T., Tanaka, F., Kuroiwa, Y., & Murayama, S. (2013). Increased number of astrocytes and macrophages/microglial cells in the corpus callosum in amyotrophic lateral sclerosis. *Neuropathology, 33*, 591–599.

Sun, S.W., Liang, H.F., Xie, M., Oyoyo, U., & Lee, A. (2009). Fixation, not death, reduces sensitivity of DTI in detecting optic nerve damage. *Neuroimage, 44*, 611–619.

Sun, S.W., Neil, J.J., Liang, H.F., He, Y.Y., Schmidt, R.E., Hsu, C.Y., & Song, S.K. (2005). Formalin fixation alters water diffusion coefficient magnitude but not anisotropy in infarcted brain. *Magn. Reson. Med., 53*, 1447–1451.

Taddei-Ferretti C. Musio, (1997). Neuronal Bases and Psychological Aspects of Consciousness.

Takeda, T., Uchihara, T., Mochizuki, Y., Mizutani, T., & Iwata, M. (2007). Memory deficits in amyotrophic lateral sclerosis patients with dementia and degeneration of the perforant pathway A clinicopathological study. *Journal of Neurological Science, 260,* 225–230.

Takeda, T., Uchihara, T., Mochizuki, Y., Mizutani, T., & Iwata, M. (2007). Memory deficits in amyotrophic lateral sclerosis patients with dementia and degeneration of the perforant pathway. A clinicopathological study. *Journal of Neurological Science. 260,* 225–230.

Tan, R.H., Ke, Y.D., Ittner, L.M., & Halliday, G.M. (2017). ALS/FTLD: experimental models and reality. *Acta Neuropathol.*

Tan, R.H., Kril, J.J., Fatima, M., McGeachie, A., McCann, H., Shepherd, C., Forrest, S.L., Affleck, A., Kwok, J.B.J., Hodges, J.R., Kiernan, M.C., & Halliday, G.M. (2015). TDP-43 proteinopathies: Pathological identification of brain regions differentiating clinical phenotypes. *Brain, 138*, 3110–3122.

Thal, D.R., Holzer, M., Rüb, U., Waldmann, G., Günzel, S., Zedlick, D., & Schober, R. (2000). Alzheimer-related τ-pathology in the perforant path target zone and in the hippocampal stratum oriens and radiatum correlates with onset and degree of dementia. *Experimental Neurology.*

Thavarajah, R., Mudimbaimannar, V., Rao, U., Ranganathan, K., & Elizabeth, J. (2012). Chemical and physical basics of routine formaldehyde fixation. *J. Oral Maxillofac. Pathol. 16,* 400.

Thivard, L., Pradat, P.-F., Lehéricy, S., Lacomblez, L., Dormont, D., Chiras, J., Benali, H., & Meininger, V. (2007). Diffusion tensor imaging and voxel based morphometry study in amyotrophic lateral sclerosis: relationships with motor disability. *J. Neurol. Neurosurg. Psychiatry, 78,* 889–92.

Tyszka, J.M., Readhead, C., Bearer, E.L., Pautler, R.G., & Jacobs, R.E. (2006). Statistical diffusion tensor histology reveals regional dysmyelination effects in the shiverer mouse mutant. *Neuroimage, 29,* 1058–1065.

Waller, A. (1850). Experiments on the Section of the Glossopharyngeal and Hypoglossal Nerves of the Frog, and Observations of the Alterations Produced Thereby in the Structure of Their Primitive Fibres. *Philos. Trans. R. Soc. London.*

Warren, J.D., Rohrer, J.D., Rossor, & M.N. (2013). Clinical review. Frontotemporal dementia. *BMJ, 347*, f4827.

Werner, M., Chott, A., Fabiano, A., & Battifora, H. (2000). Effect of formalin tissue fixation and processing on immunohistochemistry. *Am. J. Surg. Pathol. 24,* 1016–1019.

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., & Smith, S.M. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage.*

The Role of Motor Representation in Infants' Sensitivity to Emotional Information in Action Kinematics

Lisanne Schröer[1]
In collaboration with: Deniz Çetin[2], Stefania Vacaru[1] and dr. Johanna van Schaik[3]
Supervised by: Prof. dr. Sabine Hunnius[1]

[1]Donders Institute of Brain, Cognition and Behaviour, Radboud University Nijmegen
[2]Institute of Psychology, Leiden University
[3]Institute of Education and Child Studies, Leiden University

Emotional information can be conveyed by deviations in action kinematics (Montepare et al., 1999; Pollick et al., 2001). By 11- to 12-months of age, infants showed sensitivity to the emotional valence of action kinematics (Addabbo et al., in preparation). Whilst the underlying mechanisms of this sensitivity remain unclear, it is widely accepted that our motor system represents observed actions of others. A recent study provided the first evidence that perceived emotional states of others are dependent on our own movement kinematics (Edey et al., 2017). This suggests that infants might become sensitive to emotional information conveyed in kinematics once they have a sufficiently detailed motor representation allowing them to detect deviations in another person's movement kinematics.

This study aimed to understand how young infants become sensitive to emotional information conveyed in kinematics. Firstly, this study examined whether it could replicate the results of Addabbo and colleagues (in preparation) in a large sample. Secondly, this study investigated whether infants who have a more detailed motor representation, indicated by less kinematic variability in their movement, were more sensitive to deviations in kinematics conveying emotional information. Action kinematics of 12- to 13-month-old infants were investigated in two transport tasks using motion capture. Infants' sensitivity to kinematics of angry and happy transport actions was investigated using facial electromyography (EMG), following Addabbo et al. (in preparation). Forty-six infants with sufficient EMG data were included in the analysis to examine whether infants were sensitive to emotional information conveyed in kinematics. Twenty-four infants with sufficient data for both tasks were included in the analysis to investigate whether infants with a more detailed motor representation were more sensitive to emotional information conveyed in kinematics.

The EMG data did not provide evidence that infants this age are already sensitive to emotional information conveyed in action kinematics. The combined data of both tasks indicated a significant correlation between the measurement of motor representation and infants' sensitivity to happy kinematics. However, in contrast to our predictions, infants with higher variability, hypothesized as a less detailed motor representation, showed more zygomaticus muscle activation in response to happy stimulus videos.

This unexpected finding that more variable infants were more sensitive to emotional information (i.e., more zygomaticus compared to corrugator activation to happy stimuli) might be due to expressive infants that were more active and happy overall over both sessions, resulting in more variability in their movement and more zygomaticus activation in the EMG session. There was no evidence for a relationship between the measurement of motor representation and the sensitivity to emotional information conveyed in kinematics. However, it might be that our motor task did not capture the detailedness of the infants' motor representations as assumed. Future research should design an age-appropriate task in order to measure the detailedness of motor representation.

Keywords: Emotion Perception, Emotional Information in Kinematics, Action Kinematics, Motor Representation, Infants

Corresponding author: Lisanne Schröer; E-mail: l.schroer@bbk.ac.uk

Emotion recognition

Developing the ability to recognize emotions in other persons is essential for infants and young children for successful social interaction (Harms, Martin, & Wallace, 2010). Research has shown that infants become sensitive to emotional expressions during the first year of life. At 3 months of age, infants can already discriminate between facial expressions of surprise and happiness (Young-Browne, Rosenfeld, & Horowitz, 1977), as well as anger and happiness (Barrera & Maurer, 1981) and by 7 months of age, infants can discriminate between facial expressions of fear and happiness (Kotsoni, de Haan, & Johnson, 2001).

Importantly, facial expressions are not the only source of information that can be used for recognizing how another person feels. Information about another person's emotions can also be retrieved from vocal expressions, body posture and body motion patterns (Dael, Mortillaro, & Scherer, 2011; Heberlein & Atkinson, 2009). In particular, body motion patterns are an important source of emotional information for adults. For example, adults were able to identify emotions in body movements and gestures in actors with blurred faces (Montepare, Koff, Zaitchik, & Albert, 1999). In addition, people were capable of recognizing and identifying emotions from gait information (Montepare, Goldstein, & Clausen, 1987). Furthermore, adults could recognize emotions in point-light displays based on motion cues in arm movements (Pollick, Paterson, Bruderlin, & Sanford, 2001), walking movement (Nackaerts et al., 2012) and dance movement (Dittrich, Troscianko, Lea, & Morgan, 1996; Walk & Homan, 1984). Stern (2010) named these emotional actions vitality forms describing the 'how' of an action. In vitality forms, emotions can be detected on the basis of movement dynamics, time profile, force, space or direction (Di Cesare et al., 2014). These emotional actions (e.g., Montepare et al., 1999; Pollick et al., 2001) appeared to deviate in terms of their kinematics compared to normal actions that do not convey emotional information. For example, participants rated angry body movements and gestures as jerkier than happy or sad body movements and gestures, while happy body movements and gestures were rated as smoother (Montepare et al., 1999). An analysis of movement kinematics provided evidence that positive affect in movement kinematics is related to longer duration, slower velocity, slower acceleration and less jerk in arm movements performing drinking and knocking actions (Pollick et al., 2001). In sum, emotional information in action appears to be conveyed by deviations in kinematics.

To date, several researchers have stressed that emotional information in body movement and posture might be even a more important source for emotion recognition than facial expressions (Aviezer, Trope, & Todorov, 2012; de Gelder, 2006). When the emotional cues from the body and the face of an image were mismatched, judgement of the facial expression is limited and is biased into the direction of the emotion expressed by the body (Meeren, van Heijnsbergen, & de Gelder, 2005). This provides evidence that emotional cues from the body bias discrimination of emotional facial expressions in favour of body cues in adults (Aviezer et al., 2012; de Gelder, 2006; Meeren et al., 2005) and in infants (Rajhans, Jessen, Missana, & Grossman, 2016).

However, to date, there has been little research done into how recognition of emotional information in body movement develops in infancy. Using facial electromyography (EMG), a recent study provided the first evidence that infants of 11-to 12-month-old are already sensitive to emotional information conveyed in kinematics of movements. Happy expressions in the kinematics of an action induced a greater response in the zygomaticus major ('smiling muscle'), while angry expressions in the kinematics induced a greater response in the corrugator supercilii ('frowning muscle'), providing evidence that infants of this age show sensitivity to the emotional valence of action kinematics (Addabbo, Meyer, Vacaru, & Hunnius, in preparation).

Action experience

Whilst the underlying mechanisms of this

sensitivity remain unclear, it is widely accepted that our motor system represents actions we observe in others (e.g., Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Hari et al., 1998; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Rizzolatti & Luppino, 2001). Such motor representations of actions become more detailed as experience with that certain action increases. Observed actions that are part of our motor repertoire activated the observer's motor system. However, actions that are outside of our motor repertoire led to little activation in the motor areas (Buccino et al., 2004). Expert dancers trained in either classical ballet or capoeira showed more motor activation when observing movements they had been trained to perform compared to movements they had not (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2004; Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006). This motor activation is thought to reflect an internal motor representation that is activated during action observation (Buccino et al., 2004; Calvo-Merino et al., 2004; Calvo-Merino et al., 2006; Hunnius & Bekkering, 2014). Such internal motor representations become more detailed with experience, likewise in infancy. Infants who received motor training with an action with a sound effect showed more motor activation to the sound associated with the learned action compared to a familiar sound not associated with the movement that the infant produced (Gerson, Bekkering, & Hunnius, 2015).

Infants and adults use their own motor representation of actions to predict and interpret actions of others (e.g., Blakemore & Decety, 2001; Gallese & Goldman, 1998; Hunnius & Bekkering, 2014; Sommerville & Woodward, 2005; Southgate, Johnson, El Karoui, & Csibra, 2010; Wilson & Knoblich, 2005). More motor experience with a certain action, and respectively an improved motor representation, enhanced predicting and interpreting of actions in expert sport players (Abernethy, Zawi, & Jackson, 2008; Aglioti, Cesari, Romani, & Urgesi, 2008; Brault, Bideau, Kulpa, Craig, 2012; Diersch, Cross, Stadler, Schütz-Bosbach, & Rieger, 2012; Jackson, Warren, & Abernethy, 2006; Sebanz & Shiffrar, 2009). For example, expert basketball players with high levels of motor experience in a certain action showed an enhancement of perception in discriminating (Sebanz & Shiffrar, 2009) and predicting that certain action (Aglioti et al., 2008).

Hunnius and Bekkering (2014) suggest that action experience is essential for the infants' developing action understanding to form associations between motor representations and the sensory consequences of these actions. For example, the extent of an infant's motor experience with crawling or walking, therefore the detailedness of the infant's motor representation, determined the accuracy of the infant's prediction of another person's crawling or walking action (Stapel, Hunnius, Meyer, & Bekkering, 2016). Movement experience of an infant with a certain action, and therefore a more detailed motor representation of that certain action, improved predicting and interpreting another person's action (Cannon, Woodward, Gredebäck, von Hofsten, & Turek, 2012; Gerson & Woodward, 2014; Sommerville, Woodward, & Needham, 2005; Stapel et al., 2016).

Now, there is the first evidence that such motor representations play a role in deciphering emotions in another person's actions (Edey, Yon, Cook, Dumontheil & Press, 2017). Edey and colleagues (2017) hypothesized that we use our own motor representations of action kinematics to make judgments about the emotional states of others. In their experiment, participants had to judge the emotion of point-light-display walkers on a 10 point scale from not at all happy/angry/sad to very happy/angry/sad. A person who walks with high velocity is generally rated as angry, while a person who walks with low velocity is rated as sad. In addition, the own walking speed of the participants was assessed. There was a relationship between the participants' own walking speed and their judgments about the emotion of the point-light-display walker. Faster walkers rated high velocity point-light-display walkers as less intensely angry, while low velocity point-light-display walkers were rated as more intensely sad. This evidence that perceived emotion in kinematics is dependent on participants' own movement characteristics suggests that we use motor representations of our own movement

kinematics to make judgements about emotional states of others (Edey et al., 2017).

Current study

Our own motor experience is essential to improve the motor representation of that specific action. In addition, there is the first evidence that we use the deviations from these motor representations as an indication of emotional states of others (Edey et al., 2017). This would imply that infants become sensitive to emotional information conveyed by kinematics, once they have a sufficiently detailed motor representation that allows them to detect deviations in another person's movement kinematics. Infants are still very variable in the movements they execute, but with practice, movements and their respective motor representations improve (Calvo-Merino et al., 2004; Calvo-Merino et al., 2006; Fetters & Todd, 1987; Gerson et al., 2015; Hunnius & Bekkering, 2014; Konczak, Borutta, Topka, & Dichgans, 1995; Mathew & Cook, 1990; von Hofsten, 1991).

This study aimed to understand how young infants become sensitive to emotional information conveyed in kinematics. We hypothesized that infants need a detailed motor representation in order to become sensitive to emotional information conveyed in action kinematics. Developing a detailed motor representation in infancy is dependent on motor development. Here, we examined whether infants who have a better detailed motor representation of their kinematics were indeed more sensitive to deviations in kinematics conveying emotional information.

In a two-part study, we investigated whether infants' motor representation plays a role in deciphering emotions in another person's actions. In the first session, infant's own motor representation was investigated with the infant moving an object measured by motion capture. Actions conveying emotional information seem to be different in terms of their kinematics compared to normal actions (Montepare et al., 1999; Pollick et al., 2001). Therefore, we decided to measure variability (Cook, Blakemore & Press, 2013) over the infant's own movement kinematics as a measurement of detailedness of motor representation. We hypothesized that infants require a sufficiently detailed motor representation in terms of their kinematics in order to identify deviations in kinematics in observed movement conveying emotional information. In the second session, infants' sensitivity to kinematics of angry and happy actions was investigated by measuring facial muscle activity in response to emotional videos with angry and happy action kinematics in transport movements using facial EMG.

We hypothesized that infants would show activation in the zygomaticus muscle and deactivation in the corrugator muscle for actions with happy kinematics, while they would show activation in the corrugator muscle and deactivation in the zygomaticus muscle for actions with angry kinematics, replicating the findings of Addabbo and colleagues (in preparation). Secondly, we hypothesized that infants with a more detailed motor representation of their own kinematics would display a greater sensitivity to emotional information conveyed in kinematics of observed actions, following the reasoning based on Edey and colleagues (2017).

Methods

Participants

Total sample. A total of eighty-three 12- to 13- months-old infants were tested in this study (Table1). In Figure 1, a flowchart illustrates the participant inclusion and exclusion in the different analyses. Families were recruited from the Baby & Child Research Center database in Nijmegen, a medium sized city in the Netherlands, and its surroundings. Participation in the research was voluntary and parents of the infants were called to see if they wanted to participate with their infant. Parents were informed beforehand about the test procedure and all parents gave written informed consent. Families were given a thank-you gift for participation. All procedures were approved by the local ethics committee.

Sample with sufficient EMG data. Forty-six infants (Table 1) were included in the analysis to test whether infants were sensitive to emotional information conveyed in kinematics. From this sample, twenty-four infants (Table 1) had sufficient data for both tasks to test whether infants with a more detailed motor representation were more sensitive to emotional information conveyed in kinematics (see below). An additional thirty-seven infants (Table 1) were tested, but data was excluded from both analyses, because they did not want to wear the EMG electrodes (N = 9), were sick at the second session (N = 3), there was a technical error in recording the video during the emotional sensitivity task (N = 3), there was a technical error in recording

the EMG data (N = 5), they were too fussy (N = 6), they did not watch enough trials (N = 10) or were chewing during the experiment (N = 1). Similar drop-out rates have been reported before in facial EMG research with infants and children (Geangu, Quadrelli, Conte, Croci, & Turati, 2016; Isomura & Nakano, 2016; Vacaru, van Schaik, & Hunnius, under review).

Sample with sufficient data for both tasks. Twenty-two infants, who were included in the first analysis, were excluded from the analysis to investigate whether infants with a more detailed motor representation were more sensitive to emotional information conveyed in kinematics, because they had no trials for both motion tasks (N= 13), they had too little trials (less than 3) for the motion task (N = 7), they did not want to wear the motion capture markers (N = 1), or they had no data for the hand marker (N = 1). This left twenty-four infants to be included in the analysis.

Stimuli and procedure

This study consisted of two parts. Infants' motor representation was measured in the first session using motion capture. In the second session, infants' sensitivity to kinematics of angry and happy actions was investigated by measuring facial muscle activity in response to emotional videos with angry and happy kinematics in transport movement using facial EMG. The movement session always occurred first, with the second session measuring emotional sensitivity following preferably within 1 to 10 days. This was done in order to avoid potential biases of the emotional video stimuli on the movement of the infants.

Task on own movement kinematics. Each infant was seated at a table in a baby chair. The parent was seated next to the infant and the experimenter was seated in front of the infant. In order to track the movement of the infant's hand, reflective markers were placed on the infant's preferred hand and corresponding wrist. Parents were asked beforehand about the preferred hand of their infant. If the parent was unable to indicate a preferred hand, the infant was assumed to be right-handed. The experimenter used doubled-sided tape to place one marker on the knuckle of the infant's middle finger (3rd metacarpal) and one marker on the

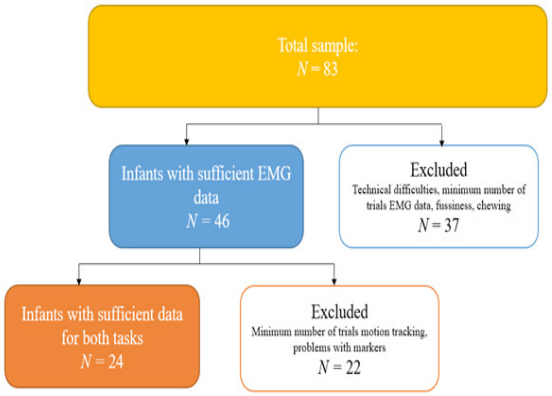


Figure 1. Flowchart of infants included in the different analyses. NB: For testing hypothesis **1**: only EMG data needed and for testing hypothesis **2**: both EMG and motion capture data was required.

Table 1.

Sample size, age in days and months and the mean days between two sessions of the samples.

	Total sample	**Sample with sufficient EMG data**	**Sample with sufficient data for both tasks**
Sample size	83 (46 females)	46 (26 females)	24 (9 females)
Age in days at the emotional sensitivity task	389.76 (*SD* = 11.41; range: 370 - 419)	391.00 (*SD* = 10.71; range: 375 - 418)	391.50 (*SD* =11.46; range: 375 - 418)
Age in months at the emotional sensitivity task	12.8 (*SD* = 0.38; range: 12.1 – 13.8)	12.8 (*SD* = 0.35; range: 12.3 – 13.7)	12.9 (*SD* = 0.38; range 12.3 – 13.7)
Mean days between sessions	7.81 (*SD* = 5.05)	8.30 (*SD* = 5.90)	7.17 (*SD* = 5.19)

corresponding location on the infant's wrist. Movements were recorded at 100 Hz using a 3D optical motion capture system (Qualisys AB, Götenborg, Sweden) with eight infrared cameras positioned around the whole table. The session was filmed using the Qualisys system camera at 13 Hz to record the start and the end of each transport movement. Parents were instructed to encourage their infant to play with the balls only using their preferred hand containing the markers.

We used two different tasks to measure the infant's transport movement. The two tasks, one easier and one more difficult, were picked based on the pilot results. These pilot results indicated that the more difficult task was the most appropriate for the age group, therefore it was decided to carry out this task first. The first task (more difficult task) was to transfer a ball from a red block onto a track (see Fig.2A). When the ball was placed on the track, it would roll down the track. The experimenter encouraged the infant verbally and non-verbally to repeat this action at least 10 times. The distance between the red block and the track was 28 cm. Next, the experimenter switched the setup to the second task. The second task (easier task) was to transfer a ball from a red block into a bowl (see Fig. 2B). Again, the infant was encouraged to repeat this action at least 10 times. The distance between the red block and the bowl was 28 cm.

Task on emotional sensitivity. Infants watched stimulus videos portraying actions with happy or angry kinematics. Stimulus videos featured an adult transporting the object (a green donut, a coloured ball, a red bar, a coloured donut or a purple ball) into a tray either from left to right or right to left displaying an angry action or a happy action. The two emotional actions were identical but performed expressing a different emotion (angry or happy). Analysis of the stimuli videos showed that actions with happy kinematics were associated with slower velocity, slower acceleration and less jerk in movement compared to actions with angry kinematics. These stimuli videos were used in previous research (Addabbo et al., in preparation). Four different actors featured in the videos, resulting in a total of 64 different stimuli videos: 32 angry and 32 happy. Only the torso of the actor was visible during the actions, whereas their face was out of view. Videos were presented in a pseudo-randomized way. The pseudo-randomized sequences of videos were created in the program "Mix" (van Casteren & Davis, 2006). The whole experiment consisted of 3 blocks with in total 256 trials: 128 angry and 128 happy trials.

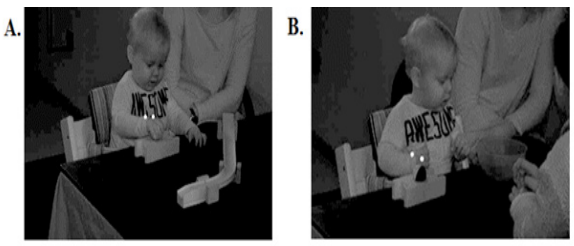


Figure 2. A. The experimental set-up for the transporting task onto the track. **B**. The experimental set-up for the transporting task into the bowl.

Each trial (see Fig. 3) started with a fixation cross paired with a beep sound to attract the infant's attention to the centre of the screen. The fixation cross was displayed with a varying time between 600 to 1000 milliseconds. This fixation cross was used as baseline. Next, a video depicting either happy or angry emotion conveyed in the kinematics was played with a length of 2800 milliseconds. After each trial, there was a 500 milliseconds inter-trial interval (grey screen) before the beep sound was played again. The fixation cross was displayed 45 to 50 milliseconds later. There was the following constraint: no emotion could occur more than two times successively.

EMG procedure and recordings.

Infants were seated on the lap of the parent, and the parent was asked to hold the hands of their child. First, the infant's face was cleaned with baby skin cleanser and scrubbed lightly with Nuprep Skin Prep Gel to ensure good quality signal recordings from the EMG electrodes. Infants were entertained with nursery rhymes movies or bubbles during the preparation. Conductive OneStep clear gel was placed on the electrodes to improve their impedances. It was aimed to keep impedance below 10 kΩ (following Vacaru et al., under review and Geangu et al., 2016).

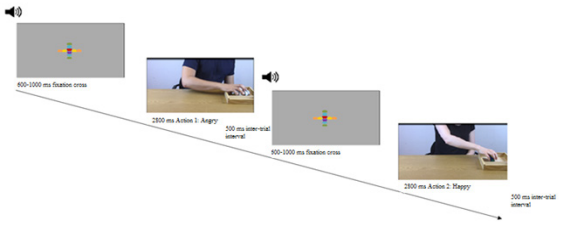


Figure 3. Example frames from the experiment illustrating two successive trials. Trials began with a static fixation cross paired with a beeping sound, followed by an action video, displaying either happy or angry emotion. Next, there was an inter-trial interval of 500 milliseconds before the next beep was played. Markers were time-locked to the onset of each action video.

Electromyography (EMG) was recorded for the zygomaticus major and the corrugator supercilii on the left side of the face with 4 Neuroline EMG Ag/AgCl electrodes in a bipolar configuration with 10 mm inter-electrode distance (see Fig. 4) (Cacioppo, Petty, Losch, & Kim, 1986; Cacioppo, Tassinary, & Berntson, 2007). In previous studies, the zygomaticus major was found to be a reliable indication for differential facial expression for happiness, while the corrugator supercilii was considered a reliable indication for differential facial expression for anger (Addabbo et al., in preparation; Cacioppo et al., 1986; Ekman & Friesen, 1976). Two additional Neuroline EMG Ag/AgCl electrodes were used for the reference and the ground. The reference electrode was positioned on the left mastoid and the ground was positioned just below the hairline in the middle of the forehead (see Fig. 4). The EMG signal was amplified using a Brain Products Amplifier, recorded continuously at a sample rate of 2500 Hz and band-pass filtered (0.016 – 120 Hz) with Brain Vision Recorder (Brain Products GmbH, Munich, Germany).

Infants were shown the stimulus videos on a 17 inch monitor (1280 x 1024 pixels) at a distance of approximately 50 cm from the infant and parent. Infants were monitored by the experimenter in order to determine when they had lost interest. Parents were instructed not to interact with their infant during the videos, except pointing to the screen to reorient the infant's attention to the screen. The experimental session was video-recorded for offline movement and attention coding. The session was ended when the infant became fussy or inattentive. Parents were debriefed after the sessions about the aim of the experiment.

Data-analysis

Motion capture data-analysis. The start and the end of each transport movement for both tasks were determined offline. Trials in which the infant transported the ball with two hands or the non-preferred hand, trials with a different starting point other than the red block, or with a different endpoint other than the track or bowl, or with a pause in the transporting movement were excluded. Furthermore, attempts in which the ball did not reach the end goal position (track or bowl) were excluded. In addition, trials with any missing marker locations for the hand marker were excluded. Lastly, trials in which the parent 'helped' the infant in their movement were excluded.

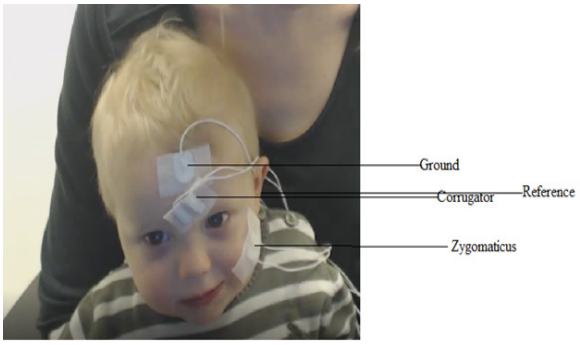


Figure 4. Positions of the EMG electrodes. All electrodes were placed on the left side of the face.

The mean standard error of the mean absolute jerk, acceleration and velocity that were calculated from the movement data can be seen as a measurement of variability in movement kinematics over trials (following Cook et al., 2013). We assumed here that infants with a low variability over trials, meaning better movement control, have a more detailed motor representation.

All pre-processing steps and filters were based on previous research (Cook et al., 2013). Velocity of the movement for each trial was calculated as the square root of the sum of the squared differentials of the x, y, and z vectors of the hand marker[1]. The velocity vectors were low-pass filtered using a Butterworth 1st order filter with a low-pass of 10 Hz, and 10 data points were trimmed from the end of each velocity vector to remove possible artefacts associated with the filter. Acceleration and jerk for each trial were calculated as the first and the second order differentials of these filtered velocity vectors. The distance for each trial was estimated by multiplying the mean velocity of that trial by the number of frames (duration) of that trial. Mean standard error of the mean (SEM) absolute jerk (in mm/frames^3), acceleration (in mm/frames^2), and velocity (in mm/frames) over the first 3 trials were calculated for each participant in analogy to a previous study (Cook et al., 2013). In addition, the mean distance (in mm) and the mean duration (in frames) over the first 3 trials were calculated for each participant (Cook et al., 2013).

However, due to the low number of infants with enough trials (N = 11 with EMG data and 3 or more trials) for the first task (train track task), and no significant correlations of these mean SEM values with the mean SEM values of the bowl task (see Table 2), it was decided not to analyse the data for the first task. The non-significant correlations could indicate that the first task (train track task)

1 The hand marker represented the movement kinematics of the infants better than the wrist marker. Therefore, it was decided to use the hand marker in the analysis.

Table 2.

The correlation of the two tasks in their mean standard error of the mean of the different measurements.

***N* = 12**	**Correlation between values of task 1 and task 2**
Mean SEM of absolute velocity	r = .362 (p = .248)
Mean SEM of absolute acceleration	r = .240 (p = .453)
Mean SEM of absolute jerk	r = .102 (p = .752)
Mean distance	r = .345 (p = .272)
Mean duration	r = .534 (p = .074)

Note: Task 1 is the train track task; task 2 is the bowl task. The correlation is a Pearson correlation, two-tailed.

was not measuring the same concept (in this case detailedness of motor representation) as the second task (bowl task), or that there was too little variance in the first task.

Infants had to have a minimum of 3 trials in the bowl task in order to be included in the final statistical analysis. This inclusion criterion was based on the mode of the trials of all the participants. Infants had a mean number of 10.25 trials (SD = 8.00, range: 3-30). There was one left-handed infant included in the final sample. Mean SEM values were always calculated from the hand marker of the preference hand of the first 3 trials of each infant.[2] Given that the velocity, acceleration and jerk scores are related directly through a mathematical operation, a single composite score was calculated to summarize these three variables.

A factor analysis on velocity, acceleration and jerk scores was performed using the regression method calculating the resulting kinematic score. This was done following previous research (Cook et al., 2013).

To investigate whether there is a relationship between the infant's detailedness of motor representation and the infant's sensitivity to emotional information, we aimed to execute a correlation analysis between the kinematic score and the sensitivity scores of the conditions (angry and happy) in the EMG data (see paragraph *EMG data reduction and analysis*). However, tests[3] demonstrated that the kinematic score variable violated the assumptions of most commonly used Pearson correlation analyses.

Therefore it was decided to conduct a non-parametric test, the Spearman's correlation, to determine the relationship since it does not require these assumptions. All statistical analyses were conducted in SPSS statistical software version 25.0 (IBM Corporation, Armonk, New York).

EMG data reduction and analysis. Videos of the EMG session were coded offline whether the infant was paying attention to the video stimuli or not using ELAN annotation software (Max-Planck Institute for Psycholinguistics, Nijmegen, the Netherlands). Trials in which the infant did not pay attention to the video stimuli were excluded. Two coders coded the first 10 infants, with a good agreement (*Cohen's kappa* = .75).

The EMG data was pre-processed using Brain Vision Analyzer 2.1 (Brain Products GmbH, Munich, Germany). The pre-processing steps were based on previous research (Vacaru et al., under review). The EMG signal was filtered offline using a band rejection filter of 50 Hz with bandwidth of 0.2 Hz and order 4. In addition, an infinite impulse response (IIR) zero phase shift Butterworth filter with a low cut-off filter of 20 Hz and a high cut-off filter of 500 Hz and order 8 was applied on the data. The different scores of the bipolar electrodes of each muscle (zygomaticus and corrugator) were calculated. Next, the data was segmented into trials based on stimulus onset. After excluding the trials based on attention coding (when the infant was not paying attention to the video stimuli), trials with

2 In addition, movement units of each trial were calculated in the MATLAB toolbox TimeStudio (Nyström, Falck-Ytter, Gredebäck, 2016). Movement units consist of an acceleration and deceleration phase, and are often used as measurements of movement control (e.g., Konczak & Dichgans, 1997; von Hofsten, 1991). Filtering values were based on earlier research (Grönqvist, Brodd, & von Hofsten, 2011) and the TimeStudio motion-tracking analysis manual (Gottwald & Ekberg, in preparation). However, there was very little variance (M = 1.083, SD = 0.177, range: 1.000 – 1.667) in movement units per trial in our sample based on these filtering values. We therefore decided not to look further in movement units.

3 The kinematic score deviated significantly from a normal distribution in a Shapiro-Wilk test (W = .663, p <.001), and by visual analysis. In addition, the skewness and kurtosis values were not within acceptable range based on Field (2009).

signal noise or motion artefacts contaminated the signal were also discarded based on visual inspection of the data. Lastly, the data was rectified: all values were made absolute values because of our interest in the absolute amplitude of the signal.

Mean activation values were calculated for the baseline (500 ms pre-stimulus onset until stimulus onset) and for the trial interval (700-2800 ms post-stimulus onset). The choice of this trial interval was based on previous research (Addabbo et al., in preparation) and confirmed through visual inspection of the data (see Fig. 5). Activation immediately after the stimulus onset is often seen as a startle response with no difference in activation between the zygomaticus and corrugator muscle activation, and is usually discarded (Addabbo et al., in preparation; Geangu et al., 2016; Isomura & Nakano, 2016). Baseline correction was calculated as the percentage change in activation during stimulus presentation compared to baseline activation during the fixation cross.

Lastly, sensitivity scores were calculated in order to investigate the relationship between the infant's sensitivity to emotional information conveyed in kinematics and the detailedness of their motor representation. The difference between the zygomaticus and the corrugator mean activation in response to happy stimuli was calculated as the happy sensitivity score. The difference between the corrugator and the zygomaticus mean activation in response to angry stimuli was calculated as the angry sensitivity score. Positive values indicated more activation in the corresponding muscle compared to the non-corresponding muscle. This was done in analogy to previous research (Vacaru et al., under review).

Infants had to reach a minimum of 3 trials per condition in order to be included in the final statistical analyses. Forty-six infants were included in the final analyses. The mean number of trials in the happy condition was 12.61 (*SD* = 9.507; range: 3-42) and the mean number of trials in the angry condition was 13.74 (*SD* = 9.733; range: 3-41).

Results

Are infants sensitive to emotional information conveyed in kinematics?

Figure 5 shows the descriptive results of the zygomaticus and the corrugator muscle activation in response to both emotional stimuli over time. First, we tested whether infants were sensitive to emotional information conveyed in kinematics. We hypothesized that infants would show activation in the zygomaticus and deactivation in the corrugator while observing happy actions, and that they would show activation in the corrugator and deactivation in the zygomaticus while observing angry actions, replicating the findings of Addabbo and colleagues (in preparation). We conducted a 2 (Emotion: happy, angry) x 2 (Muscle: zygomaticus, corrugator) repeated measures ANOVA with percentage change score (in the time-window 700-2800 milliseconds; see Fig. 5) from baseline in EMG activation as dependent variable.

The repeated measures ANOVA did not yield a significant interaction between emotion and muscle ($F(1,45) = 0.011$, $p = .916$). This meant that

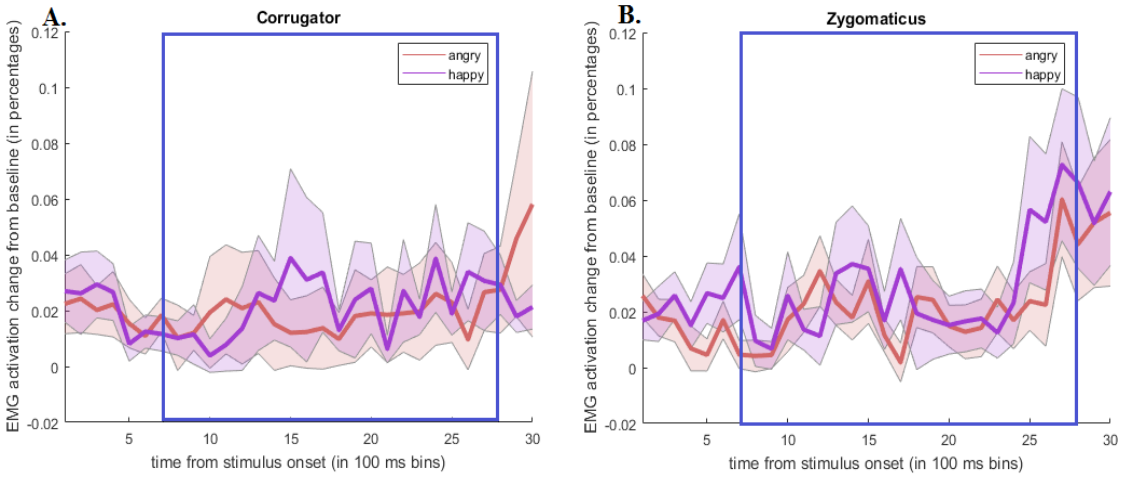


Figure 5. Percentage change EMG activation from baseline (shaded areas indicate the standard error of the mean) for corrugator (**A**) and zygomaticus (**B**) muscle when observing action with happy (purple line) and angry (red line) emotional information conveyed in kinematics. The stimulus presentation lasted 2800 milliseconds. Our time window of interest is indicated by the dark blue square.

Table 3.

Means (and standard deviations) of percentage change in EMG activation from the zygomaticus and corrugator muscle to both emotional expressions in time-window 700-2800 ms.

***N* = 46**	**Corrugator *M* (*SD*)**	**Zygomaticus *M* (*SD*)**
Angry	0.9762 (3.8380)	1.9460 (4.9344)
Happy	1.1921 (2.6116)	2.2694 (6.3648)

the difference in activation from baseline in the zygomaticus muscle and the corrugator muscle was not significantly different in the happy and the angry condition (see Table 3 and Fig. 6). The absence of the interaction effect between emotion and muscle does not correspond to our predictions, namely, that infants would show activation in the zygomaticus and deactivation in the corrugator during the observation of happy kinematics and would show activation in the corrugator and deactivation in the zygomaticus during the observation of angry kinematics. There were no additional significant effects in the repeated measures ANOVA.

Does motor representation play a role in deciphering emotion in another person's action in infancy?

Secondly, we tested whether infants who had a more detailed motor representation were more sensitive to emotional information conveyed in kinematics of observed action as hypothesized. We conducted a non-parametric Spearman's correlation with the kinematic score and the sensitivity score of both emotions in the EMG data. The kinematic score did not correlate significantly with the angry sensitivity score (r_s = -0.126, N = 24, p = .279, one-tailed), but there was a significant correlation between the happy sensitivity score and the kinematic score (r_s = .368, N = 24, p = .038, one-tailed; see Table 4 and Fig. 7).[4] In addition, there was a significant correlation between the angry sensitivity score and the happy sensitivity score (r_s = -.619 N = 24, p = .001, one-tailed).

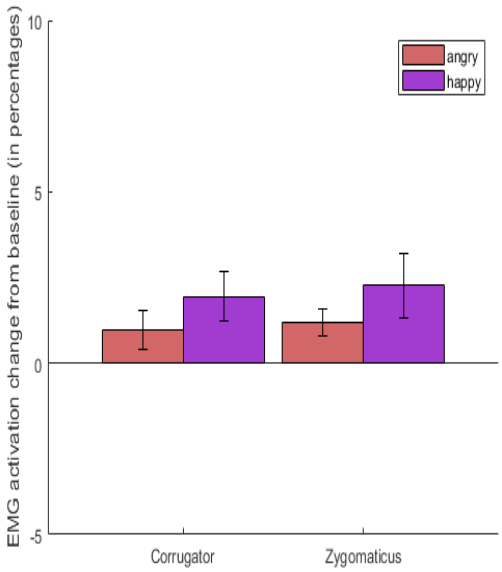


Figure 6. Percentage change EMG activation (error bars represent standard errors) in the time interval 700-2800 ms after stimulus onset compared to baseline in both emotional conditions (angry and happy) in the two muscles (corrugator and zygomaticus).

There was a weak positive relationship between the happy sensitivity score and the kinematic score (r_s = .368). A higher score on the kinematic score, indicated a higher variability in the movements was associated with a higher score on the happy sensitivity score, indicated that these infants showed more activation for the zygomaticus muscle compared to the corrugator muscle in the happy condition. This was in contrast with our hypothesis stating that infants that were more variable in movement (a higher score on kinematic score) were less sensitive to emotional information conveyed in kinematics (had a lower score on the happy score). In addition, there was a moderate to strong negative relationship between the happy and angry sensitivity score (r_s = -.619). A higher sensitivity happy score was associated with a lower angry sensitivity score. Infants with higher activation for the zygomaticus muscle compared to the corrugator muscle in the happy condition, showed lower activation for the corrugator muscle compared to the zygomaticus muscle in the angry condition. This indicates a tendency of infants to activate a similar muscle to a certain extent to both

4 The same analysis was executed excluding the four infants that had more than 10 days in between sessions. However, excluding these infants did not change the results (angry score: r_s = .065, N = 20, p = .393; happy score: r_s = .314, N = 20, p = .089).

video stimuli compared to baseline. This is as well in contrast with our hypothesis, as we expected infants to be sensitive to emotional information conveyed in kinematics, which would result in high angry and happy sensitivity scores, and a strong positive correlation between these scores.

In addition, a Pearson correlation between the sensitivity score of both emotional conditions and the kinematic score was conducted excluding the kinematic score, the happy sensitivity score and the angry sensitivity score outliers[5] in order to control for the potential effects of the outliers on the correlation. Without these outliers, the assumptions of the Pearson correlation were not violated. Outliers can have adverse effects on correlations (Osborne & Overbay, 2004).

5 Outliers with a score above or below 2 standard errors of the mean were excluded. When excluding outliers with a score above or below 3 SD of the mean, the kinematic score still deviated significantly from a normal distribution in a Shapiro-Wilk test ($W = .660$, $p < .001$) and the kurtosis and skewness values were not within acceptable range (based on Field, 2009). It was therefore decided to exclude outliers with a score above or below 2 standard errors of the mean.

The kinematic score correlated significantly with the happy sensitivity score ($r = .434$, $N = 18$, $p = .036$, one-tailed) and correlated marginally significantly with the angry sensitivity score ($r = -.373$, $N = 18$, $p = .064$, one-tailed). Both correlations were weak. In addition, there was a significant negative strong correlation between the angry sensitivity score and the happy sensitivity score ($r = -.744$, $N = 18$, $p < .001$, one-tailed) similar to the previous analysis (see Table 5 and Fig. 7).

A higher score on the kinematic score was associated with higher scores on the happy sensitivity score, and lower scores on the angry sensitivity score. Infants with high variability had higher activation in the zygomaticus muscle compared to the corrugator muscle to observed happy stimuli, and lower activation in the corrugator muscle compared to the zygomaticus muscle to observed angry stimuli. This indicated a tendency for infants with high variability to display more zygomaticus activation to the observed stimuli compared to baseline. This was in contrast with our hypothesis that infants with less variability in movement (low score on kinematic score) were expected to be more sensitive to emotional information conveyed in kinematics.

Table 4.

The correlations between the EMG measurements and the kinematic score

***N* = 24**	**EMG measurements**		**Motion capture**
	Angry sensitivity Score	Happy sensitivity Score	Kinematic Score
Angry sensitivity Score	-		
Happy sensitivity Score	r_s =-.619 (p = .001)*	-	
Kinematic Score	r_s =-.126 (p = .279)	r_s =.368 (p = .038)*	-

Note: The correlation was a Spearman's correlation, one-tailed. * means a significant correlation.

Table 5.

The correlations between the EMG measurements and the kinematic score excluding the outliers

***N* = 18**	**EMG measurements**		**Motion capture**
	Angry sensitivity Score	Happy sensitivity Score	Kinematic Score
Angry sensitivity Score	-		
Happy sensitivity Score	r =-.744 ($p < .001$)*	-	
Kinematic Score	r =-.373 (p = .064)	r =.434 (p = .036)*	-

Note: The correlation was a Pearson's correlation, one-tailed. * means a significant correlation

Discussion

The current study examined whether infants need a detailed motor representation in order to become sensitive to emotional information conveyed in kinematics. Our first hypothesis was that infants were sensitive to emotional information conveyed in kinematics, showing a differential facial response to both emotional stimuli (replicating results of Addabbo et al., in preparation). Our second hypothesis was that infants with a more detailed motor representation, and therefore less variable movement, would be more sensitive to emotional information conveyed in kinematics.

Are infants sensitive to emotional information conveyed in action kinematics?

In contrast to our hypothesis, the findings of Addabbo and colleagues (in preparation) were not replicated in this current study. Our results yielded no indication that infants were sensitive to emotional information conveyed in kinematics as found by Addabbo and colleagues (in preparation). There was no evidence of differential facial response to the two emotional stimuli, meaning that we found no evidence that infants in this age group (12- to 13-months-old) are sensitive to emotional information conveyed in movement kinematics, and there was no evidence that infants were able to differentiate between angry and happy emotional information in

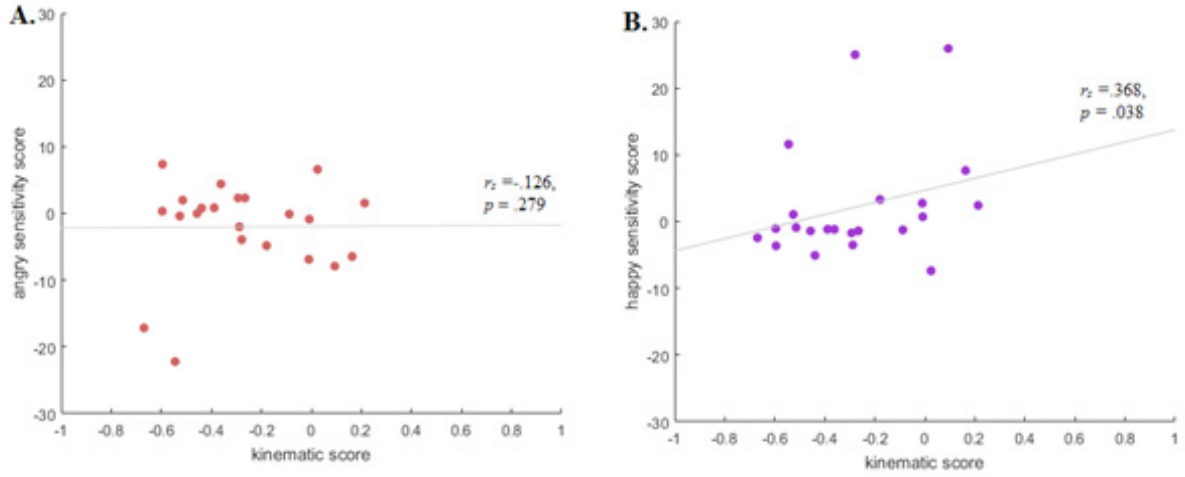


Figure 7. A. Scatterplot depicting the relationship between the kinematic score and the angry sensitivity score. **B**. Scatterplot depicting the relationship between the kinematic score and the happy sensitivity score.

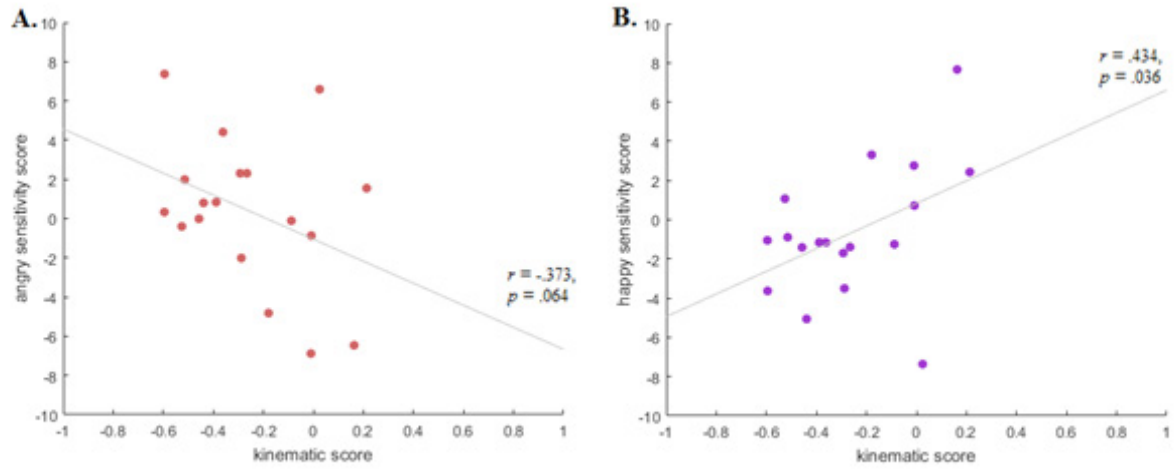


Figure 8. A. Scatterplot depicting the relationship between the kinematic score with the angry sensitivity score excluding the outliers. B. Scatterplot depicting the relationship between the kinematic score with the happy sensitivity score excluding the outliers.

actions. However, we found a negative correlation between the two (angry and happy) sensitivity scores. Infants with more activation in zygomaticus muscle compared to corrugator muscle in the happy condition showed less activation in corrugator muscle compared to zygomaticus muscle in the angry condition. This might indicate a tendency of infants to display a similar response (for example, activate the zygomaticus muscle more) in response to both emotional stimuli. One possible explanation is that some infants found the baseline (the fixation cross) boring and activated their zygomaticus muscle more when the stimuli movies started playing. In sum, our study yielded no evidence that infants are yet sensitive to emotional information conveyed in kinematics.

In the literature, there is some discussion on the muscle reaction to angry stimuli in infants and children. In adults, there is activation in the corrugator muscle in response to angry faces (Dimberg, 1982). In childhood, however, Geangu and colleagues (2016) reported a frontalis muscle (lifts the brows in fear) activation in response to angry faces in 3-year-old children, while Deschamps and colleagues (2012) found an activation in the corrugator muscle to angry dynamic faces in children aged 6 to 7 years. In infancy, 4-months-old infants were found not to show selective facial reactions to any facial expressions (Kaiser, Crespo-Llado, Turati, & Geangu, 2017). 7-months-old infants showed an increased zygomaticus activation to dynamic happy facial expressions, while they did not show a differential response to angry faces (Kaiser et al., 2017), nor was there evidence of a differential corrugator muscle activity (Datyner, Henry, & Richmond, 2016). We based our decision to measure the corrugator muscle response as an indication for sensitivity to angry stimuli on the previous findings of Addabbo and colleagues (in preparation), however, it seems that in infancy and childhood the corrugator muscle activation might not necessarily be the best measurement of differential facial response to angry stimuli. This might explain the lack of differential facial response found for the angry stimuli videos. However, this cannot explain the lack of differential facial response found for the happy stimuli videos.

Isomura and Nakano (2016) suggest a system that elicits facial muscle in response to emotional stimuli that matures over infancy, first only triggered by bimodal emotional information. Later in infancy, when the system is matured, it is as well triggered by unimodal emotional information. It might be that our stimuli did not trigger this not fully matured system to elicit facial muscle in response to emotional stimuli. Isomura and Nakano (2016) found that 4- to 5-month-old infants only show an increased corrugator response to combined audio-visual cries and an increased zygomaticus response to combined audio-visual laughter. These responses were absent for both the visual and auditory unimodal emotion stimuli. They suggested that in infancy a system starts to mature in order to elicit facial muscles responses to emotional stimuli, but that system has not matured yet fully, and motor responses are only generated when multimodal information is present. Future research could look into whether infants are able to display a differential facial response to bimodal stimuli with emotional information conveyed in kinematics, for example combined with vocal or auditory emotional information.

In addition, it might be that infants in our age group are able to understand emotional information in facial expressions, however, the sensitivity to emotions in bodily expressions and movement develops later in toddlerhood or even childhood, explaining the lack of differential facial response to the different emotions conveyed in kinematics. Geangu and colleagues (2016) found that 3-years-old showed the expected increased response in the zygomaticus muscle to happy facial stimuli, however, they did not show the expected increased response in the zygomaticus muscle to bodily expressions. They explained this finding by suggesting that children fail to associate the emotional body posture of an observed person with the causing emotional state. This ability might develop at a different rate than interpreting emotional facial information.

Furthermore, it might be that infants and young children are not sensitive to actions that are not directed at them personally. In daily life, infants experience many action directed at the infant, such as being picked up or being fed. This may explain the differences between this current study and the studies reporting sensitivity to emotional facial expressions in which stimuli are generally directed at the infant. However, it is generally accepted in the literature that young infants can learn models of actions acquired through observational statistical learning, meaning that they observe someone performing actions not necessarily directed at them (e.g., Hunnius & Bekkering, 2014; Monroy, Gerson, & Hunnius, 2017; Monroy, Meyer, Schröer, Gerson, & Hunnius, 2019). Whether infants can learn emotions in actions from observing actions performed that are not directed at them personally remains unknown. More research with into different age groups (for example, toddlers and children) is necessary in order to be able to understand the development of sensitivity to emotional information

in movement kinematics.

Lastly, there is a possible methodological limitation of this study. It might be that our current time window is cut off too soon for the EMG response, resulting in the effect bleeding over in the next trial's baseline during the fixation cross. However, our choice of the time window was based on Addabbo and colleagues (in preparation), who did find an effect in this time window. In addition, previous studies have that facial muscle generally begin to show a differential activation response to facial emotional expressions starting around 500 ms after stimulus onset, and usually decreasing the response 2000 ms after stimulus onset (Beall, Moody, McIntosh, Hepburn, & Reed, 2008; Geangu et al., 2016; Oberman, Winkielman, Ramachandran, 2009).

Does the infant's motor representation play a role in deciphering emotion in another person's action?

Our second hypothesis was that infants with a more detailed motor representation, indicated by less variability in their movements, would be more sensitive to the emotional information conveyed in the kinematics of happy and angry observed actions. However, we only found a significant correlation between the movement variability measure and the sensitivity score to happy stimuli in the EMG data. The relationship between the movement variability score and happy score indicated that infants with higher variability in their movement showed a higher zygomaticus muscle activation compared to corrugator muscle activation to happy stimuli, which is in contrast with our hypothesis. When controlling for outliers, there was also a marginally significant correlation between the movement variability measure and the sensitivity score to angry stimuli. This relationship indicated that infants with higher variability in their movement showed lower corrugator muscle activation compared to zygomaticus muscle activation to angry stimuli. These two relationships between the kinematic score and the emotional sensitivity scores indicated that infants with more variability tended to activate their zygomaticus muscle more during the EMG session. One explanation might be that infants with more expressive and extraverted temperament were overall more active and happy in both sessions, resulting in more variability in their movement as well as more smiling during the EMG session.

These results might indicate that the kinematic score was not capturing the detailedness of motor representations well in our sample. We assumed that infants with more variable movement would have a less detailed motor representation. The mean standard deviation of the mean absolute jerk, acceleration and velocity are described as measurement of movement control (Cook et al., 2013). Movement control and planning itself uses motor representations (Kawato, 1999). Often, motor activation in the motor system during the observation of actions is taken as an indication of motor representation (Buccino et al., 2004; Calvo-Merino et al., 2004; Calvo-Merino et al., 2006). Another possibility of a measurement of motor representations might be motor activation in observation of the neutral action, for example, measured by electroencephalogram (EEG) in infancy and childhood (e.g., Gerson et al., 2015). More motor activation would be an indication of more motor experience, and a respectively improved motor representation (Calvo-Merino et al., 2004; Calvo-Merino et al., 2006). However, with the current methods, it is still impossible to measure motor representations directly in infancy.

Unfortunately, our movement task was too difficult for many infants. This resulted in several infants who did not have any trials for the bowl task (N = 13) or too few trials (N = 7), who we were not able to include in our analysis. These infants generally showed less competent movement control, consisting of pausing in movements, putting the ball first in the mouth before moving to the bowl, or using a different start- or endpoint. It might be that our task therefore measured in a lesser extent the detailedness of the infants' motor representation. Future research should design a more appropriate motor task for this age. Infants in this age group have very limited motor abilities and often still mouth stuff. A good movement task aiming to get an idea of the detailedness of the infant's motor representation should be able to tolerate for pauses in movement and be more flexible in order to include infants with 'less competent' movement.

Based on piloting, we used two tasks to measure infant's variability in movement. However, the first task (train task) was very hard to execute by the infants. Infants required fine control in order to place the ball on the train track to make it roll. The bowl task was more flexible with an endpoint that was not a small precise location. There was more variance in the standard error of the mean velocity, acceleration and jerk in the bowl task compared to the train task. In addition, very few infants in this age group were able to do the train task (N = 11). It might be that the few infants that were able to

do the train track task were very competent in their movement, resulting in low variability, and therefore low variance within the included group. However, most of the infants in our sample were not able to do this task. An easier task would therefore be more appropriate for this age group.

Conclusion

This study investigated how young infants become sensitive to emotional information conveyed in kinematics. We had two hypotheses in this study: first, we hypothesized that infants of this age (12- to 13-months) are sensitive to emotional information conveyed in kinematics (replicating Addabbo et al., in preparation) and second, we hypothesized that infants with a more detailed motor representation would be more sensitive to emotional information conveyed in kinematics, based on the idea of Edey and colleagues (2017) that we use our own motor representation to make judgments about the emotional states of others.

In our results, we were unable to replicate the results of Addabbo and colleagues (in preparation). This study did find a relationship between the angry and the happy sensitive score, indicating that infants showed a similar activation pattern in response to both stimuli (such as smiling in a certain extent to both angry and happy action kinematics). Our research did not provide evidence that infants of 12- to 13-months-old are sensitive to emotional information conveyed in action kinematics.

In our second analysis, we found a correlation between the infants' variability in movement and their sensitivity to emotional information conveyed in kinematics when observing happy actions (and only when controlling for outliers, a marginal significant correlation when observing angry actions). Infants with a higher variability score tended to activate their zygomaticus muscle more in response to the happy and angry stimuli. It might be that more expressive, extraverted infants are overall more active and comfortable during the experiment, showing more variability in the motion task and smiled more during the emotional sensitivity task. It might be that our kinematic score does not provide us with a measurement of the detailedness of the infant's motor representation to the full extent. In addition, our motor task also excluded infants with less competent movement kinematics. Future research should use a more appropriate movement task to investigate whether infants need a detailed motor representation to become sensitive to emotional information conveyed in kinematics.

In sum, this study did not provide evidence that infants of 12- to 13-months-old are yet sensitive to emotional information conveyed in action kinematics, and nor did it provide evidence that this sensitivity is related to the infant's own motor representation.

Acknowledgments

We would like to thank Angela Khadar and Rosa Pranger for their help with recruitment of the families. We would also like to thank dr. Margaret Addabbo for providing us with the emotional sensitivity stimuli. Furthermore, we would like to thank Erik van den Berge, dr. Marta Bakker and dr. Janny Stapel for their help with the motion capture set-up. In addition, we would like to thank Erik Verhaar with his help with analysing the video stimuli and the movement data. Lastly, we would like to thank all families that participated in this study.

References

Addabbo, M., Meyer, M., Vacaru, S., & Hunnius, S. (in preparation). Binding emotion and action in infancy.

Abernethy, B., Zawi, K., & Jackson, R.C. (2008). Expertise and attunement to kinematic constraints. *Perception, 37,* 931-948.

Aglioti, S.M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*(9), 1109-1116.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions discriminate between intense positive and negative emotions. *Science, 338,* 1225-1229.

Barrera, M.E., & Maurer, D. (1981). The perception of facial expression by the three-month-old. *Child Development, 52*(1), 203-206.

Beall, P.M., Moody, E.J., McIntosh, Hepburn, S.L., & Reed, C.L. (2008). Rapid facial reactions to emotional facial expressions in typically developing children and children with autism spectrum disorder. *Journal of Experimental Child Psychology, 101*(3), 206-223.

Blakemore, S.J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience, 2*(8), 561-567.

Brault, S., Bideau, B., Kulpa, R., & Craig, C.M. (2012). Detecting deception in movement: The case of the side-step in rugby. *PLOS ONE, 7*(6), e37494.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi. F., … & Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. *Journal of Cognitive Neuroscience, 16*(1), 114-126.

Cacioppo, J.T., Petty, R.E., Losch, M.E., & Kim, H.S. (1986). Electromyographic activity over facial muscle

regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology, 50(2)*, 260-268.

Cacioppo, J.T., Tassinary, L.G., & Berntson, G. (Eds.) (2007). *Handbook of Psychophysiology.* Cambridge University Press

Calvo-Merino, B., Glaser, D.E., Grèzes, J., Passingham, R.E., & Haggard, P. (2004). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*(8), 1243-1249.

Calvo-Merino, B., Grèzes, J., Glaser, D.E., Passingham, R.E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology, 16*, 1905-1910.

Cannon, E.N., Woodward, A.L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-months-old infants' attention to others' actions. *Developmental Science, 15*(1), 35-42.

Cook, J.L., Blakemore, S.J., & Press, C. (2013). Atypical basic movement kinematics in autism spectrum conditions. *Brain, 136*(9), 2816-2824.

Dael, N., Mortillaro, M., & Scherer, K.R. (2011). Emotion expression in body action and posture. *Emotion, 12*(5), 1085-1101.

Datyner, A., Henry, J.D., & Richmond, J.L. (2016). Rapid facial reactions in response to happy expressions in 7-month-old infants. *Developmental Psychobiology, 59,* 1046-1050.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience, 7*(3), 242-249.

Deschamps, P.K., Schutte, I., Kenemans, J.L., Matthys, W., & Schutter, D.J. (2012). Electromyographic responses to emotional facial expressions in 6-7 years olds: A feasibility study. *International Journal of Psychophysiology, 85*(2), 195-199.

Di Cesare, G., Di Dio, C., Rochat, M.J., Sinigaglia, C., Bruschweiler-Stern, N., Stern, D.N., Rizzolatti, G. (2014). The neural correlates of 'vitality form' recognition: An fMRI study. *Social Cognitive and Affective Neuroscience, 9*(7), 951-960.

Diersch, N., Cross, E.S., Stadler, W., Schütz-Bosbach, S., & Rieger, M. (2012). Representing others' actions: The role of expertise in the aging mind. *Psychological research, 76,* 525-541.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology, 19*(6), 643-647.

Dittrich, W.H., Troscianko, T., Lea, S.E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25*(6), 727-738.

Edey, R., Yon, D., Cook, J., Dumontheil, I., & Press, C. (2017). Our own action kinematics predict the perceived affective states of others. *Journal of Experimental Psychology – Human Perception and Performance, 43*(7), 1263-1268.

Ekman, P., & Friesen, W.V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behaviour, 1*(1),56-75.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*(6), 2608-2611.

Fetters, L., & Todd, J. (1987). Quantitative assessment of infant reaching movements. *Journal of motor behavior, 19*(2), 147-166.

Field, A. (2009). *Discovering statistics using SPSS*. London: SAGE.

Gallese, V., & Goldman, A (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493-501.

Geangu, E., Quadrelli, E., Conte, S., Croci, E., & Turati, C. (2016). Three-year-olds' rapid facial electromyographic responses to emotional facial expressions and body postures. *Journal of Experimental Child Psychology, 144,* 1-14.

Gerson, S.A., Bekkering, H., & Hunnius, S. (2015). Short-term motor training but not observational training alters neurocognitive mechanisms of action processing in infancy. *Journal of Cognitive Neuroscience, 27*(6), 1207-1214.

Gerson, S.A., & Woodward, A. L. (2014). Learning from their own actions: The unique effect of producing actions on infants' action understanding. *Child Development, 85*(1), 264-277.

Gottwald, J.M., & Ekberg, T.L. (in preparation). TimeStudio process manual – Analysis of Motion-tracking data.

Grönqvist, H., Brodd, K.S., & von Hofsten, C. (2011). Reaching strategies of very preterm infants at 8 months corrected age. *Experimental Brain Research, 209*(2), 225-233.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences of the United States of America, 95*(25), 15061-15065.

Harms, M.B., Martin, A., & Wallace, G.L. (2010). Facial emotion recognition in Autism Spectrum Disorders: A review of behavioural and neuroimaging studies. *Neuropsychogical reviews, 20,* 290-322.

Heberlein, A.S., & Atkinson, A.P. (2009). Neuroscientific evidence for simulation and shared substrates in emotion recognition: Beyond faces. *Emotion Review, 1,* 162-177.

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of Royal Society B, 369:* 20130490.

Isomura, T., & Nakano, T. (2016). Automatic facial mimicry in response to dynamic emotional stimuli in five-month-old infants. *Proceedings of the Royal Society B, 283*(1844), 20161948.

Jackson, R.C., Warren, S., & Abernethy, B. (2006). Anticipation skills and susceptibility to deceptive movement. *Acta Psychologica, 123*, 355-371.

Kaiser, J., Crespo-Llado, M.M., Turati, C., & Geangu,

E. (2017). The development of spontaneous facial responses to others' emotions in infancy: An EMG study. *Scientific reports, 7,* 17500.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology, 9*(6), 718-727.

Konczak, J., Borutta, M., Topka, H., & Dichgans, J. (1995). The development of goal-directed reaching in infants: Hand trajectory formation and joint-torque control. *Experimental Brain Research, 106*(1), 156-168.

Konczak, J., & Dichgans, J. (1997). The development toward stereotypic arm kinematics during reaching in the first 3 years of life. *Experimental Brain Research, 117*(2), 346-354.

Kotsoni, E., de Haan, M., & Johnson, M.H. (2001). Categorical perception of facial expression by 7-month-old infants. *Perception, 30,* 1115-1125.

Mathew, A., & Cook, M. (1990). The control of reaching movements by young infants. *Child Development, 61*(4), 1238-1257.

Meeren, H.K., van Heijnsbergen, C.C., & de Gelder, B. (2005). Rapid perceptual integration of facial expressions and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America, 102*(45), 16518-16523.

Monroy, C.D., Gerson, S.A., & Hunnius, S. (2017). Toddlers' action prediction: Statistical learning of continuous action sequences. *Journal of Experimental Child Psychology, 157*, 14-28.

Monroy, C.D., Meyer, M., Schröer, L., Gerson, S.A., & Hunnius (2019). *NeuroImage, 185*, 947-954.

Montepare, J., Goldstein, S.B., & Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behaviour, 11*(1), 33-42.

Montepare, J., Koff, E., Zaitchik, D. & Albert, M. (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behaviour, 23*(2), 133-152.

Nackaerts, E., Wageman, J., Helsen, W., Swinnen, S.P., Wenderoth, N., Alaerts, K. (2012). Recognizing biological motion and emotion from point-light displays in Autism Spectrum Disorders. *PLOS One, 7*(9): e44473.

Nyström, P., Falck-Ytter, T., & Gredebäck, G. (2016). The TimeStudio Project: An open source scientific workflow system for behavioural and brain sciences. *Behavioural Research Methods, 48*(2), 542-552.

Pollick, F.E., Paterson, H.M., Bruderlin, A., & Sanford, A.J. (2001). Perceiving affect from arm movement. *Cognition, 82*(2), B51-B61.

Oberman, L.M., Winkielman, P., & Ramachandran, V.S. (2009). Slow echo: Facial EMG evidence for delay of spontaneous, but not voluntary emotional mimicry in children with autism spectrum disorders. *Developmental Science, 12*(4), 510-520.

Osborne, J.W., & Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). *Practical Assessment, Research, and Evaluation, 9*(6), 1-12.

Rajhans, P., Jessen, S., Missana, M., Grossman, T. (2016). Putting the face in context: Body expressions impact facial emotion processing in human infants. *Developmental Cognitive Neuroscience, 19*, 115-121.

Rizzolatti, G., Fadiga, L.,Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131-141.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron, 31*, 889-901.

Sebanz, N., & Shiffrar, M. (2009). Detecting deception in a bluffing body: The role of expertise. *Psychonomic Bulletin & Review, 16*(1), 170-175.

Sommerville, J.A., & Woodward, A.L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition, 95,* 1-30.

Sommerville, J.A., Woodward, A.L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*, B1-B11.

Southgate, V., Johnson, M.H., El Karoui, I., & Csibra, G. (2010). Motor system activation reveals infant's on-line prediction of others' goals. *Psychological Science, 21*(3), 355-359.

Stapel, J.C., Hunnius, S., Meyer, M., & Bekkering, H. (2016). Motor system contribution to action prediction: Temporal accuracy depends on motor experience. *Cognition, 148*, 71-78.

Stern, D.N. (2010). *Forms of Vitality Exploring Dynamic Experience in Psychology, Arts, Psychotherapy and Development.* Oxford: Oxford University Press.

Vacaru, V.S., van Schaik, J.E., & Hunnius, S. (under review). Do young children mimic selectively? An investigation into three-year-olds' facial mimicry and its modulation by affiliation tendencies.

van Casteren, M., & Davis, M.H. (2006). Mix, a program for pseudorandomization. *Behavioural Research Methods, 38*(4), 584-589.

von Hofsten, C. (1991). Structuring of early reaching movements: A longitudinal study. *Journal of Motor behavior, 23*(4), 280-292.

Walk, R.D., & Homan, C.P. (1984). Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society, 22*(5), 437-440.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Cognition, 95*, 1-30.

Young-Browne, G., Rosenfeld, H.M., & Horowitz, F.D. (1977). Infant discrimination of facial expressions. *Child Development, 48*(2), 555-562.

Bridging the Gap between Deep Learning and Neuroscience: An Investigation of the Biological Plausibility of Deep Neural Networks for the Task of Visual Object Recognition

Fiammetta Strazzera Perniciani [1]
Supervisors: Leonardo Franco[2], Paul Tiesinga [1]

[1]Radbouboud University Nijmegen, Donders Institute for Brain Cognition and Behaviour, The Netherlands
[2]Universidad de Málaga, Spain

In recent years, Deep Learning has achieved superhuman abilities in many tasks such as visual object recognition. Nevertheless, the brain outperforms Deep Networks in its ability to generalize to distorted images. Yet, the exact mechanisms used to achieve this invariance are still not completely understood. The interplay between neuroscience and Deep Learning could both advance the knowledge on the processes that occur in the brain and help the development of more efficient artificial networks. The aim of the present paper is to study the link between the brain and artificial neural models by comparing the behavior of a Convolutional Neural Network to our knowledge of the processing of visual information in the human cortex. The network's recognition ability under invariance conditions was tested when presenting input images that were different from the images employed for the training of the network. The test images were modified either with geometric deformations, by varying the rotation, position and size of the objects within the image, or by compromising the extent of visual information transmitted from the input when changing the quality, contrast and amount of noise. The results are compared to neural data obtained from behavioral and neuroimaging studies in which the subject's response time, accuracy and neural activations were recorded following the presentation of images with the various types of deformations. Furthermore, the fundamental characteristics of the architecture of the network and the backpropagation algorithm used for the training process are discussed in comparison to the structure of the visual stream and to the synaptic update processes that are thought to be employed by the brain for learning. Our investigation highlights that a great issue with current Deep Neural Networks is the limited performance under image distortions as compared to humans' invariant recognition ability. Furthermore, the present study underlines the differences in the implementation of the learning algorithm in computational models and in the brain as a starting point to improve Deep Learning towards more efficient and more biologically plausible networks.

Keywords: Convolutional Neural Network, Deep Learning, neuroscience, visual object recognition

Corresponding author: Fiammetta Strazzera Perniciani; E-mail: fiammetta.strazzera@gmail.com

Building a representation of visual information is one of the most crucial functions of the visual system. Recognizing or classifying objects is a particularly complex task, since an object can appear in the visual field over various viewing conditions: this is referred to as the invariance problem. Transformations that preserve the identity of an object include changes in position, size, illumination and rotation of the object along with its background. By comparing the behaviour of object recognition algorithms under these invariance conditions to our knowledge of the processing of visual information in the human cortex, we aim at studying the link between the brain and artificial neural networks.

Given its great success over the last years, in the present study, the task of object recognition will be tackled using Deep Learning. Deep Learning, a research area in the field of Machine Learning, is the latest development of artificial neural network models that comprise several hidden layers. It has become the new gold standard among different applications in artificial intelligence. This is supported by its superhuman abilities in several tasks such as pattern recognition, game playing, medical diagnosis and social network filtering. This new technology is inspired by several features of the mammalian brain, without being constrained by any biological limitations.

Initially influenced by neuroscience, Deep Learning algorithms have strongly developed over the past years, making it possible to train artificial neural networks with several layers to complete various tasks efficiently. Nevertheless, those algorithms have now little explicit resemblance to the processes occurring in the mammalian brain. Yet, we claim that the interplay between neuroscience and Deep Learning can advance the study of learning processes in the brain. Neuroscience can help Machine Learning to develop the best strategies, optimizing functions and architectures. Moreover, it can formulate constraints on the implementation of learning algorithms so to properly match the real neural processes. Likewise, Deep Learning provides a tool with which hypotheses from neuroscience can be tested empirically.

The pioneer studies on visual processing in the brain were carried out by Hubel and Wiesel and represent the starting point for the development of Deep Learning algorithms (Hubel &Wiesel, 1959; Hubel & Wiesel, 1962). By performing various experiments investigating the visual system in the cat's brain, they showed that some regions in the visual cortex are sensitive to specific areas of the visual field, called receptive fields, or specific orientations or shapes. Specifically, the authors identified two types of cells in the brain: simple cells and complex cells. The former are neurons in the cortex that respond exclusively to one position or orientation while being silent to stimuli outside their focal area. The latter are units which fire in the presence of specific movements of the object in the visual field. Object recognition might be performed in the brain by integrating information of both types of cells (Schiller, Finlay & Volman,1976). In line with this work, several neural network architectures were proposed.

For instance, Convolutional Neural Networks, that perform nearly as robustly as our brains under several transformations of the objects in the visual field. In these networks each region has its visual receptive field and responds to specific features (Lecun, Bottou, Bengio & Haffner, 1998; Lecun, Haffner, Bottou & Bengio, 1999; Krizhevs, Suskever & Hinton, 2012; Szegedy et al., 2015). Additionally, other networks were designed in which layers that resemble the functioning of simple and complex cells are alternated (Serr, Oliva & Poggio, 2007; Riesenhuber & Poggio, 1999). Amongst the most popular Deep Neural Networks are LeNet-5 (Lecun et al., 1998), a five layers neural network usually applied to the task of recognizing handwritten numbers, HMAX (Serre et al., 2007), a biologically inspired hierarchical neural network, AlexNet (Krizhevsky et al., 2012), an extension of LeNet, GoogLeNet (Szegedy et al., 2015), a 22 layers deep network, and the VGG-16 and Very Deep networks comprising 16 and 19 layers respectively (Simonyan and Zisserman, 2014). For a complete overview of the fundamental innovations and techniques that led to the great performance of neural networks please refer to Nielsen (2018), and Yamins and DiCarlo (2016).

Image classification occurs instantaneously in the brain, whereas it is a challenging task for an artificial neural network. Building an artificial neural network that performs object recognition as accurately and efficiently as our own visual system might be achieved by mapping the spatial organization of the brain areas and portions of the cortex involved in this process. The ability to recognize objects relies on largely feedforward computations that flow throughout the visual ventral stream of the mammalian brain. The transmission of visual information starts in the retina, continues in the lateral geniculate nucleus of the thalamus (LGN) and then through the primary visual cortex V1, secondary visual cortex V2, visual cortex V4 to the inferior temporal cortex (IT) (Trappenberg, 2002). Each cortical area responds to

specific features of an image and unravels different types of information (Blumberg & Kreiman, 2010). It is likely that the IT is the portion of the visual stream that is mainly responsible for object recognition (DiCarlo, Zoccolan & Rust, 2012).

The algorithm used in the brain to solve object recognition is still not completely understood. Empirical findings in neuroscience, concerning the organization and structure of the visual ventral stream, can help to define the hypothesis space and orient the implementation of a possible algorithm. For instance, clues can be taken by studying the activity of neurons in the ventral visual stream, their firing rate, their sparseness and their tolerance, that is the ability to preserve preference for a limited range of object variables. Given its success in the last decades, several attempts have been made to integrate Deep Learning results and neuroscience data (Kheradpishe, Ghodrati, Ganjtabesh & Masquelier, 2016b; Baldi & Sadowski, 2014; Baldi & Sadowski, 2016; Dodge & Karam, 2016; Dodge & Karam, 2017; Geirhos et al., 2017). Several studies compared neural data obtained using functional magnetic resonance imaging (fMRI), electroencephalography (EEG) or magnetoencephalography (MEG) to activation of units in artificial neural networks (Kheradpisheh, Ghodrati, Ganjtabesh & Masquelier, 2016a; Güçlü & van Gerven, 2015).

It was shown that, as a neural network is trained to recognize objects, a hierarchical structure, in which increasingly complex features are processed, naturally emerges along its layers (Cichy, Khosla, Pantazis, Torralba & Oliva, 2016; Güçlü & van Gerven, 2015). This increasing complexity is comparable to the processing of visual information in the brain. Specifically, the last layer of a neural network is particularly predictive of IT neurons' responses and the previous layer is predictive of the responses of neurons in the V4 cortex (Yamins et al., 2014; Cadieu et al., 2014). In contrast, the biological plausibility of the training procedures applied in Deep Learning is still questioned. As a matter of fact, it is unlikely for neurons to perform backpropagation, the most common algorithm used to train neural networks (Rumelhart, Hinton & Williams, 1988). Nevertheless, it could be possible for the brain to approximate this training algorithm. Additionally, the implementation of its optimization and activation functions is largely consistent with the observations and hypotheses regarding the functioning of our brain (Marblestone, Wayne & Kording, 2016).

The aim of the present project is to carry out a detailed analysis of the aspects involved in the functioning of Deep Learning algorithms for object recognition. Specifically, we aim at analyzing whether these aspects have a neural correlate in the mammalian brain and can represent effective simplifications of the processes occurring in biological systems or whether they are completely artificial tools. Firstly, the paper will analyze the behavior of an artificial neural network when modifying the characteristics of the representation in the input images and compare it to neural data. This consists in studying the accuracy in recognizing objects and the activation of neural units when varying the rotation, position or size of the objects as well as changing the quality, the contrast, or adding noise to the input images. Secondly, the characteristics of artificial networks will be discussed in terms of their biological plausibility based on neuroscientific data. These characteristics include the architecture and connectivity, the neural activation functions, the training process, the use of the backpropagation algorithm and the dropout scheme to prevent overfitting.

Methods

The task of image classification consists in taking an input image and giving the class that it belongs to among a fixed set of categories representing different objects or scenes. In order for the network to learn the correct classification, error signals are used to update the parameters of the network proportionally to the derivative of the classification error. In the present paper, image classification will be investigated using Deep Neural Networks (DNNs), that are characterized by several hidden layers (Goodfellow, Bengio & Courville, 2016). Our DNN was implemented in Keras (Chollet et al., 2015), an efficient and flexible application program interface (API). As a Deep Learning framework, we used Google's Tensorflow (Abadi et al., 2015), an open source library written in Python and used frequently in Machine Learning (Rampasek & Goldenberg, 2016).

Three fundamental factors shape DNNs and determine the correlation between representations in DNNs and cortical visual representations: the architecture, the task and the training procedure (Cichy et al., 2016).

Deep Neural Network Architecture

It was proven that a network with a single hidden layer, given it has enough units, can approximate any function and operation of a Deep Network (Cybenko, 1989). It is true, however, that the

number of units needed in order to learn decreases exponentially with the depth of the network (Cohen, Sharir & Shashua, 2016; Liang & Srikant, 2016). Moreover, DNNs can represent a large number of possible configurations in the input space with very rich descriptions and are crucial in order to solve the complex problems required for artificial intelligence (Hastad, 1986; Bengio & Delalleau, 2011). With distributed representations, Deep Networks have the advantage of learning the input with a number of parameters that scales linearly and not exponentially with the dimensionality of the feature space, as opposed to non-parametric approaches (Hinton, 2014). Nevertheless, training neural networks with many layers is computationally expensive and frequently has the disadvantage of overfitting the data (Hastad & Goldmann, 1991; Bengio & Lecun, 2007). Overfitting occurs as the network has more parameters than training data and overlearns the input images, losing the ability to generalize. Therefore, in order to avoid these problems, we used a deep architecture with only five layers, resembling the networks proposed by Yamins et al. (2014) and Serre et al. (2007).

The type of neural network we chose for image classification is Convolutional. Convolutional Neural Networks (CNNs) are a type of feed-forward neural network, that is, a network in which the information flows in forward direction from one input layer to one output layer and which have no cycles. CNNs consist of a series of convolutional layers, followed by fully connected (dense) layers, in which the units are connected to all units in the previous layer with a linear operation and by an output layer in which each unit represents a different target class.

Convolutional layers

Convolutional layers take the input and convolve it with a weight matrix, called filter or kernel. Convolution consists in sliding the filter through the input units and, for each slide, multiplying it element-wise to the corresponding portion of the input, adding up the result to form one unit of the output. The size of the kernel is smaller than the size of the input, therefore, each unit has a local receptive field (Lawrence, Giles & Tsoi, 1997). Each convolutional layer can have more than one filter, leading each to a different output, called a feature map. With the use of convolution, for each feature map the same weight matrix is shared throughout all the input units. Therefore, the weights of a convolutional layer are denominated shared weights. Thus, the same feature of the input object, such as orientation or shape, is detected in a feature map regardless of its position in the visual field: in this sense each filter learns to recognize a specific characteristic of the input.

Max-pooling layers

Each convolutional layer is typically followed by a max-pooling or downsampling, operation, which reduces the size of each feature map by extracting subregions of the input layer with the maximum value (Zhou & Chellappa, 1988). Specifically, the max-pooling operation divides its input into disjoint regions of a given size and takes the maximum over all the values in each region. Therefore, only the locations that show the maximum correlation with each feature are kept, creating a new, smaller layer, whereas the other values in the region are discarded. Max-pooling preserves features specificity and helps increase robustness to clutter by discarding objects that cause low responses. This reduces the number of parameters of the network and thus the computational cost and processing time (Boureau, Ponce & Lecun, 2010). By reducing the number of parameters of the network, max-pooling additionally helps to prevent overfitting.

Regularization techniques

In order to reduce overfitting, regularization techniques were applied to the architecture of network, such as dropout and L2-regularization. Dropout is a technique that allows the network to avoid learning the training data too specifically and being unable to classify new images (Hahnloser, Bengio, Frasconi & Schmidhuber, 2000). With dropout, noise is injected into the network in order to increase robustness over variations of the input images (Baldi & Sadowski, 2014). Dropout consists in randomly dropping units during the training of artificial neural networks, preventing each unit to rely excessively on the output of a specific input neuron (Hinton, Krizhevsky, Suskever & Salakhutdinov, 2012). When applying a dropout of probability (or level) p to a layer in the network, in each training iteration every unit in the network layer is deleted with probability p. The remaining weights are then trained according to the chosen training algorithm (Srivastava, Hinton, Krizhevsky, Sutskever & Salakhutdinov, 2014). In order to choose the level of dropout in each layer, a grid search was done by varying the dropout level in the convolutional layers and in the fully connected layer in the set of ten uniformly distributed values between 0 and 0.9. Dropout indeed increased the performance

of the network, confirming the findings of Paine, Khorrami, Han and Huang (2014). The best performing model which was selected for the experimental manipulations had a dropout of 0.1 after the convolutional layers and of 0.5 following the dense layer and reached 88.38% validation accuracy.

In addition to dropout, L2-regularization was included in the learning algorithm, as illustrated in the training procedure section below. Contrary to the findings by Loshchilov and Hutter (2017), this technique was found to be more effective than weight decay as a method to penalize excessively high connections between neurons and therefore reduce overfitting (M Zur, Jiang, Pesce & Drukker, 2009).

Activation functions

Activation functions are applied to the output of each layer, adding a non-linearity that is necessary in order for the network to perform complex tasks (Jarrett, Kavukcuoglu, Ranzato & LeCun, 2009). The sigmoidal activation function is commonly used in Deep Learning, since it introduces non-linearities in the model. However, a known issue with this function is the vanishing gradient problem (Hochreiter, 1991; Hochreiter et al., 2001). The sigmoid approaches a constant value when moving away from the y-axis and consequently its derivative assumes infinitely small values. Therefore, the error signals needed for learning tend to vanish. In order to avoid this problem, we utilized the rectifier function following all the layers, with the exception of the softmax function that was applied to the output layer, in accordance with the work of Güçlü & van Gerven (2015).

The rectifier applies the function R(zi) = max(0,zi) to the output zi of neuron i in the layer, eliminating the neurons with negative outputs, thus giving rise to a sparse representation. Biologically inspired (Hahnloser, Seung & Slotine, 2003), it is the most frequently employed function in Deep Neural Networks (Ramachandran, Zoph and Le, 2017) because of its efficiency (Glorot, Bordes & Bengio, 2011; Nair & Hinton, 2010). A unit to which this activation function is applied is called Rectified Linear Unit (ReLU) (Nair & Hinton, 2010). The softmax activation function is defined as $S(z)_i = \frac{e^{z_i}}{\sum_j e^{z_j}}$, where z_i represents the output of the neuron i of a given layer. This function is usually used in the final layer of a network used for classification (Bishop, 2006) due to its normalizing effect on its output, preventing it from becoming too large.

Batch normalization

The values of the input pixels as well as the activations of the units in each layer can have very distinct values throughout the layer, differing by several orders of magnitude. Those values can be adjusted by normalizing the training data and the activations of the layers, a technique called batch normalization. Constraining the units to have the same mean and variance reduces the covariance shift, that measures the amount of variation between activations in one layer (Ioffe & Szegedy, 2015). Batch normalization limits the amount to which updating the parameters in the earlier layers can affect the distribution of values of the following layers. This stabilizes the network, that becomes robust to changes in the input distribution. Therefore, each layer learns more independently, and this speeds up learning and gives the network the ability to generalize (Ma & Klabjan, 2017). Additionally, since the mean and variance for the normalization

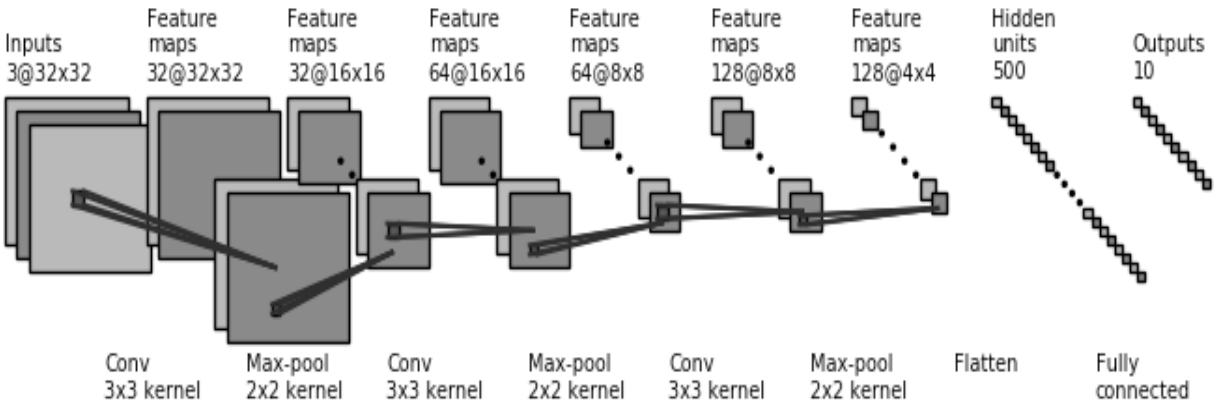


Figure 1. Neural network architecture. The first three layers consist of a convolution, with 3x3 kernels and 32, 64 and 128 filters respectively, and a max-pooling operation of size 2x2. Each plane is a feature map. The last convolutional module is followed by a fully connected layer and the output layer. Source: own elaboration.

are computed on batches of data rather than on the whole dataset, this adds noise to learning and therefore has a slight regularization effect that helps preventing overfitting. Batch normalization was added after each activation function, as it was shown to perform better if added after rather than before the layer of non-linearity (Mishkin & Matas, 2015).

As illustrated in Figure 1, the first 3 layers of the network are convolutional, followed by one fully connected layer and the output. Each convolutional layer consists of the convolution, a ReLu activation function, batch normalization, a max-pooling operation of size 2x2 and a dropout layer of probability 0.1. The weight matrices used for convolution have a 3x3 kernel and the filters are 32, 64 and 128 respectively. The fully connected layer is followed by a ReLu activation function and a dropout of probability 0.5. The softmax activation function is applied to the output layer.

Task

Training on real world objects is critical for the correspondence between layers of the CNNs and cortical visual pathways, as shown by Cichy et al. (2016). The neural network described in the present paper was trained to recognize images in the CIFAR-10 dataset (Krizhevsky, 2009), which consists of 60000 RBG images of size 32x32 representing items from 10 categories. As shown in Figure 2, the categories are: airplane, automobile, bird, cat, deer, dog, frog, horse, ship and truck.

Training procedure

Figure 2. Example of images from the CIFAR-10 dataset. For each of the ten categories, 10 random images are shown. Source: https://www.cs.toronto.edu/~kriz/cifar.html

The images in the dataset are divided into training set, validation set and test set, in a 10:1:1 ratio. The first set is used to train the network and the remaining two for testing. The training set contains examples of inputs with their associated outputs. Learning is supervised in the sense that the network learns by comparing its prediction to a given target output (supervision). Neural network learning aims at reducing the prediction error, that is, the difference of activation between the actual and the desired output. This is achieved by propagating information from the output layer back to the input layer and updating the layer weights accordingly.

Training the neural network consists in finding the parameters θ of a neural network that significantly reduce a cost function that measures to which degree the predicted output differs from the target output. This cost function includes the loss of the network, that is a measure of the classification error over all the training set, as well as additional regularization terms. The backpropagation algorithm (backwards propagation of error) (Werbos & J. Paul John, 1974; Rumelhart et al., 1988) finds a local optimum of the function that the network is trying to learn by updating its parameters and going towards the direction of lower error.

The per-example loss function is given by:

$$L(x, \hat{y}, \theta) = -\hat{y} log\big(y(x, \theta)\big) \quad (1)$$

where $y(x,\theta)$ is the predicted output vector when the input is x, representing the probabilities of x being in each of the classes, and $\hat{y}$ is the target output vector. In order to penalize network weights with high magnitudes, the regularization term

$$\frac{\lambda}{2}\|\theta\|^2 \quad (2)$$

is added to the loss function, where λ is a given penalization factor. Thus, when a batch B of N example images $x_1, \ldots, x_N$ with target outputs $\hat{y}_1, \ldots, \hat{y}_N$ is presented to the network, the total cost is:

$$J(B, \theta) = \frac{-1}{N}\sum_{n=1}^{N} \hat{y}_n \, log\big(y(x_n, \theta)\big) + \frac{\lambda}{2}\|\theta\|^2 \quad (3)$$

In our network, we choose N = 64 and λ = 10−6. In order to minimize the cost function J, an adaptive learning rate optimization algorithm, Adam, whose name derives from adaptive moment estimation, was used (Kingma & Ba, 2015). Adam chooses a separate learning rate for each parameter of the objective function, speeding learning when a different learning rate is needed for each parameter, and uses momentum: for each timestep, a fraction of

the previous update is added to the current update, moving faster to the direction of the minimum and decreasing the oscillations around it. The use of adaptive learning rates combined with momentum makes the algorithm efficient and fairly robust to the choice of hyperparameters (Reddi, Kale & Kumar, 2018).

Data augmentation

When trained with "small" datasets such as CIFAR-10, which have less images than the total number of parameters of the network, often the models tend to overfit the data (Perez & Wang, 2017). In addition to adding dropout to the architecture of the network and weight regularization to the learning algorithm, another technique used to prevent overfitting is data augmentation (Simard et al., 2003; C. Wong et al., 2016; Cagli et al., 2017). Data augmentation has been proven particularly effective for image classification (Perez & Wang, 2017). This strategy consists of increasing the amount of training samples by applying a transformation, such as reflection, rotation, shear and shift, to the training images. For every epoch a new transformation is applied to every input image. Thus, distinct images are presented to the network each time. An example of such a transformation is illustrated in Fig.3, where an original image from the CIFAR-10 dataset is rotated by 15◦.

In this study, the training dataset was augmented by rotating the initial samples of a random angle between −15◦ and 15◦, translating them horizontally and vertically by 10% of their total width and height, and reflecting them across the vertical axis. Data augmentation significantly improved the performance of the neural network, confirming the results presented by Paine et al. (2014).

Indeed, the final accuracy of the network increased by around 7.3% when applying data augmentation compared to training with the original dataset, as illustrated in Figure 4.

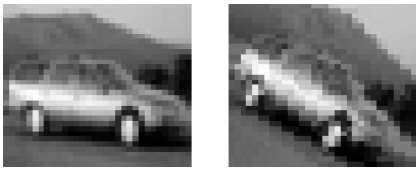


Figure 3. Example of the data augmentation process applied to one original image from the CIFAR-10 dataset. Left: original sample image. Right: new image created from the original through a 15◦ rotation.

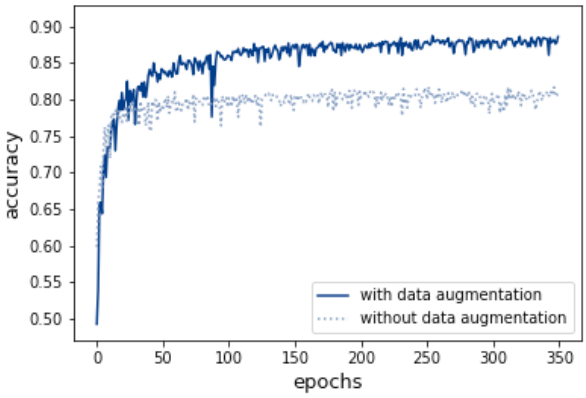


Figure 4. Performance of the Convolutional Neural Network model when training for 350 epochs. The black and gray lines show the validation accuracy over epochs when training with and without data augmentation respectively.

Experiments

Experiments regarding the performance of humans in tasks in which images are modified with various transformations trace back to Koffka (1935) and Walsh and Kulikowski (1988). The visual system is particularly robust to deformations of the objects in the visual field (Rolls, 1992; Rolls & Deco, 2002) and recent computational models have shown similar behaviors (Huiping, Bingfang & Jinlong, 2003; Dodge & Karam, 2016; Kheradpisheh et al., 2016b), although in limited extent (Ghodrati, Farzmadhi, Rajaei, Ebrahimpour & Khaligh-Razavi, 2014; Pinto, Barhomi, Cox & DiCarlo 2011; Pinto, Cox & DiCarlo, 2008). The results section discusses whether the same holds in the chosen network, illustrating the results of a series of experiments that test the behavior of the model in order to compare it to behavioral and neurological data of the human visual stream.

The first set of experiments tested the view invariance of the network to geometric transformations of the input images. Motivated by the considerable translation invariance found in the inferior temporal visual cortex (Rolls & Deco, 2002), in Experiment 1, we translated the input images vertically and horizontally. The use of convolution strides in the first layers of the network suggests that it would show robustness to translation of the objects in the input images. In order to compare the experiment to previous studies that use different datasets, we first reduced the size of the images so that the objects would be mostly contained in the DNN's visual field when their position was varied and pasted them on backgrounds created with an inpainting technique (Telea, 2004). Additionally,

we aimed to investigate the rotation invariance of the model. Since the network was trained with data augmentation techniques, we expected it to show view independence when the rotation angle was within the range of ±15◦ used in the training phase. Moreover, according to Roll's hypothesis (Rolls, 1992), invariant representations can be created by associating different learned views and, therefore, training on every rotation is not necessary in order to build view invariant representations (Booth & Rolls, 1998). In Experiment 2, we tested this hypothesis by studying whether this rotation invariance would be present also when the rotation was greater than ±15◦. Finally, in Experiment 3, we studied the performance of the network when varying the scale of the objects contained in each test image. This was achieved by reducing the size of the image within the visual field, from a 32x32 to a 20x20 size and adding a background using an inpainting technique (Telea, 2004). We expected the network to show invariance to object scaling when the quality of the images was not excessively compromised.

In order to be able to compare the results of our study to previous literature, in the first set of experiments, we used an experimental setting similar to Kheradpisheh et al. (2016b). We considered three types of variation (rotation in plane, translation over horizontal and vertical axis, and scaling) of different levels of difficulty, from no-variation to high variation. For each of these conditions (variation type and difficulty level), we created a database by randomly selecting 300 training images and 150 test images for each of the object categories. Then, we applied the corresponding variation to each database and fed our pre-trained model with the varied images. Finally, for each condition the network was evaluated on the corresponding test images.

Experiments 4 to 6 investigated the performance of the network when tested under various deformations of the input images. Following Huiping et al. (2003) and Dodge & Karam (2016), we hypothesized that the network would show robustness to moderate deformations, but that the accuracy would drop to 0 after a certain threshold. In Experiment 4, we decreased the quality, that is, the resolution, of the images by progressively reducing the number of pixels in the input images. In Experiment 5, the contrast of the test images was changed from 0 (grey image) to 1 (original image) in steps of 0.05. Additionally, in Experiment 6, noise was injected to the network by randomly selecting an increasing number of pixels in each image and changing their value with a random value between 0 and 1 taken from a uniform distribution. The number of corrupted pixels varies between 0 and 800, being 32*32*3=3072 the size of the input, in steps of 25. Since there are three color channels, this implies that the maximum percentage of noise injected corrupted at most 75% of the input pixels.

Lastly, Experiment 7 consisted in studying the activations of the neurons to new stimuli when varying the amount of dropout in the network. Before starting the training procedure, the dropout was varied between 0 and 0.9 in steps of 0.1 in the first dense layer. We hypothesized that the sparsity of the neuronal activations to novel images would increase as a function of the level of dropout, as found by Baldi & Sadowski (2014).

Results

Geometric invariances

Biological background for invariant object recognition

The ability to recognize objects under different viewing conditions is characteristic to the brain (DiCarlo et al., 2012). Although IT neurons show some tolerance to object deformation, individual neurons need not be invariant: in the visual stream, there are neurons which are view-independent and neurons whose response depends on the orientation of the object in a given image (Dicarlo & Cox, 2007). It is hypothesized that invariance is obtained by the hierarchical combination of these neurons, in which invariant features are progressively extracted (Rolls & Deco, 2002; Tanaka, 1996). In this framework, cells at higher layers pool input from lower layer cells, becoming more tolerant to changes (Riesenhuber & Poggio, 1999). Selectivity and invariance of object representations indeed increase along the visual stream (Franco, Rolls, Aggelopoulos & Jerez, 2007; Rust & DiCarlo, 2010). In a recent study (Cichy, Khosla, Pantazis, Torralba & Oliva, 2017), a marker of neural processing of spatial information was found in MEG data and compared to the development of spatial layout descriptions in computational models. Analogously to the visual stream, a gradual emergence of invariant representation was found to appear hierarchically in the neural network layers (Cichy et al., 2017).

A possible explanation for human view invariance, proposed by Biederman (1987), is that the brain represents objects by dividing their parts into 3-dimensional view-independent geometric primitives called geons that have clearly distinguishable properties in respect to symmetry,

roundness and size. Recognition of an object would occur by computing the geon descriptions of its parts and comparing them to the stored descriptions. This theory of recognition by shape makes recognition under disrupted viewing conditions easier. In support of this theory are the facts that elements that are essential to the perception of geons, such as borders, were proven to be highly relevant for object recognition in humans and that no visual priming effect was found when using distinct sets of geons between trials (Biederman, 2000). Nevertheless, the set of qualitative shape properties chosen by Biederman is arbitrary and there is no evidence for a structural description of geons in the brain. Furthermore, the theory fails to differentiate between distinct objects of the same type (Dickinson, 1999).

Another account for visual object recognition proposes that the brain stores representations as collections of different views of the object, and that recognition occurs through interpolation between those views and depends on the distance to the closest viewpoints (Spetch & Friedman, 2003). Alternatively, it was proposed that the brain may incorporate both approaches by relying on structural descriptions of the parts of the objects as well as on viewpoint-specific features (Tarr & Bülthoff, 1998).

Experiment 1: Translation invariance

Following Kheradpisheh et al. (2016b), for the translation experiment, we selected four levels of variation defined by the percentage of translation of the images in the horizontal and vertical axes. The images were translated of a random number of pixels between ±1% of the total image size in the no-variation condition, and between ±20%, ±40%, ±60% in the conditions of variation levels 1, 2 and 3 respectively. We hypothesized that the convolution strides in the first layers of the network would create robustness to translations, and that the performance of the network would decrease when the object would start falling out of the receptive visual field. Since the objects depicted in the CIFAR-10 dataset occupy a large portion of the image, we had to decrease their size, thus reducing their quality, in order to vary their position without excessively losing visual information. This led to a lower general accuracy.

Taken together, as illustrated in Figure 5b, our network shows invariance when the position of the object within the visual field is varied in the first three conditions. Indeed, when the images are translated of up to 40% of their size, the performance of the network is stable and does not decrease significantly, similarly to what has been reported in Kheradpisheh et al. (2016b). Yet, in the highest variation condition the accuracy drops considerably, as opposed to the aforementioned study. This is likely due to the fact that, with high amounts of variation, the objects in our database fall out of the receptive field.

In order to compare our results to neurophysiological data, we computed the total drop in accuracy from the no-variation to the maximum variation condition. The performance drop is of approximately 6.3% in the present experiment, whereas it was reported to be around 3% for humans (Kheradpisheh et al., 2016b). However, it is to keep in mind that only four categories were used in said study.

In general, translation invariance is the most robust type of image variation for DNNs and humans when using uniform or natural backgrounds (Kheradpisheh et al., 2016b). For human subjects, this could follow from the fact that the brain represents objects in a rectangular coordinate system, making translations easy for the brain to overcome (Hinton, 2014b). In opposition, when using natural scenes in which more than one object was present, this translation invariance was shown to decrease in humans (Rolls & Deco, 2002).

Experiment 2: Rotation transformation

For the rotation experiment, we defined the levels of variation by the range of the angle of rotation of the images. The images were rotated by a random angle between ±1◦ in the no-variation condition, and between ±15◦, ±30◦, ±60◦, ±90◦ in the conditions of variation levels 1, 2, 3 and 4 respectively. We hypothesized that the accuracy of the network would not decrease when the rotation angle was within the range of ±15◦, the angle of rotation for data augmentation used in the training phase.

Figure 5a illustrates the performance under the various rotation conditions. The network shows moderate robustness under the various levels of variation, especially in the first level. As a matter of fact, when rotating the images of an angle within the data augmentation angle of ±15◦, the performance decreases of less than 0.1% with respect to the no-variation condition. The recognition accuracy does not immediately drop when increasing the rotation angle outside the training range and the decrease follows a trend that is comparable to that depicted in the study by Kheradpisheh et al. (2016b). However, the performance decreases significantly in the last two levels of variation and is considerably lower than

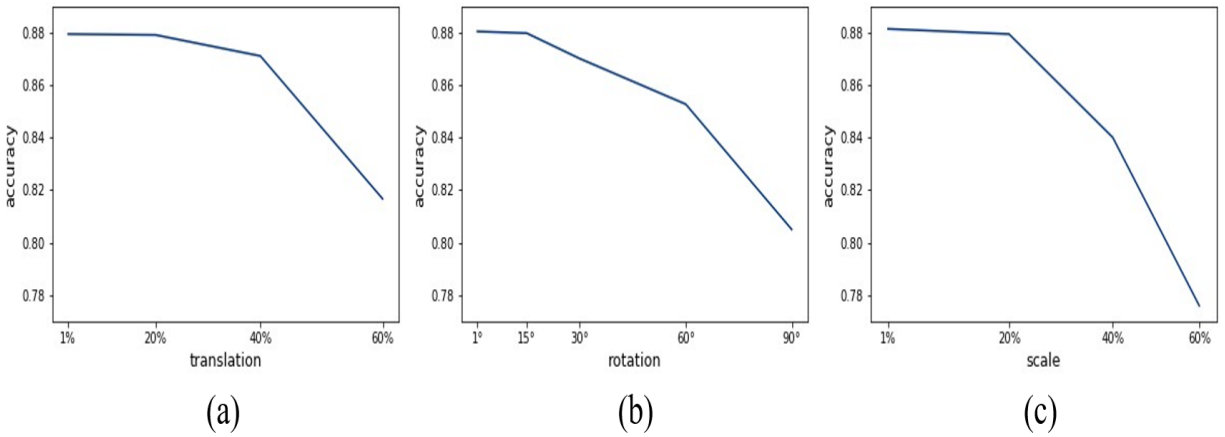


Figure 5. Empirical distribution of recognition accuracy over the various types of image deformations. The accuracy is averaged over all the images in the test set. (a) Performance of the network over image translation. The x-axis indicates the percentage of translation in the horizontal and vertical axes. (b) Performance of the network over image rotation. The x-axis indicates the angle of rotation, in degrees. (c) Performance of the network over object scaling. The x-axis indicates the size of the rescaled object with respect to the original.

that reported for the AlexNet (Krizhevsky et al., 2012) and the Very Deep (Simonyan & Zisserman, 2014) networks. The total drop in accuracy from the no-variation to the high variation condition is around 7.5% in our experiment. Nevertheless, it is to consider that in the aforementioned study only four categories were used. Moreover, this gap could be explained by the difference in the datasets used, and in the networks considered that are much deeper than our model.

In contrast with the hypotheses by Rolls (1992), the difficulties in recognizing rotated objects hold true when considering human subjects, either in terms of their response time (Murray, Jolicoeur, McMullen & Ingleton, 1993) or of their performance (Spetch & Friedman, 2003), which decreases of around 5% (Kheradpisheh et al., 2016b), for orientations to which the subjects were not trained on. This could suggest that recognition of a rotated object in the human visual system occurs through linear interpolation of two-dimensional learned views rather than by building a three-dimensional model (Bülthoff & Edelman, 1992). As a matter of fact, the first approach would explain the increase in recognition time and performance error proportionally to the amount of rotation of the object. On the contrary, the response time for recognition of rotated objects was found to diminish with practice. This suggests a shift from a mental rotation approach to a more orientation invariant approach, that could make use of geons, in which the object features are learned independently of their orientation, (Murray et al., 1993), or, more directly, suggests that the increase in the number of views with practice would lead to more uniform responses (Bülthoff & Edelman, 1993). The number of required views depends on the object and could be compared to the number of samples needed by a neural network in order to be able to generalize (Murray et al., 1993).

Experiment 3: Scaling invariance

The size of the objects contained in each test image was progressively reduced in order to test the scaling invariance of the network. The four levels of variation were defined by the size of the new image with respect to the original image. The size of the images was reduced to a random quantity within 1% of the original image size in the no- and within 20%, 40%, 60% in the conditions of variation levels 1, 2 and 3 respectively. We expected the network to show scale invariance to a certain degree, but to drop significantly in the highest variation level, as reported in Kheradpisheh et al. (2016b). Indeed, as illustrated in Figure 5c, the shape of the curve resembles that of the accuracy of the AlexNet (Krizhevsky et al., 2012) and the Very Deep (Simonyan & Zisserman, 2014) models tested in the paper. As hypothesized, the network shows robustness to image scaling when the quality of the images is not excessively compromised.

Regarding a comparison to biological data, the total drop in accuracy from the first to the last condition is of nearly 10% for human subjects (Kheradpisheh et al., 2016b) and similar (around 10.5%) for our model. It was observed that human performance in size invariant tasks significantly improves with practice, but the improvement is specific to each object and does not transfer to novel objects (Furmanski & Engel, 2000).

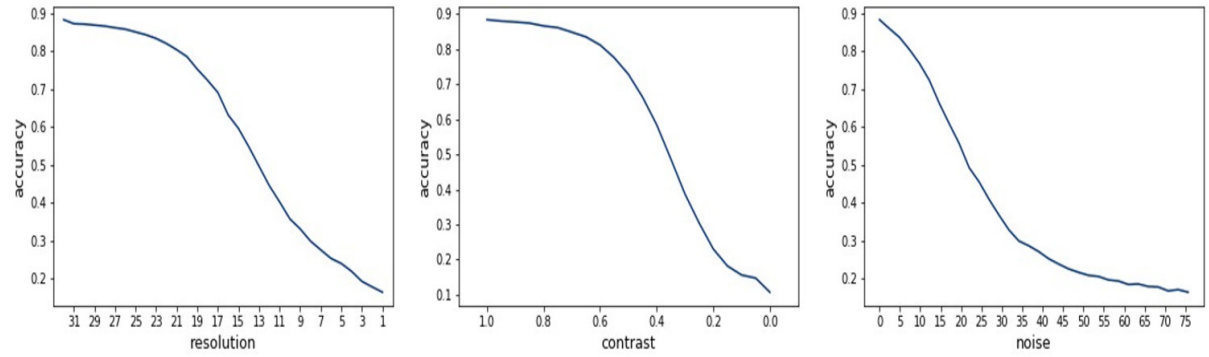


Figure 6. Empirical distribution of recognition accuracy over the various types of image distortions. The accuracy is averaged over all the images in the test set. (a) Performance of the network over image resolution. The x-axis indicates the length of the side of the input test images, in pixels. (b) Performance of the network as a function of contrast. The x-axis indicates the contrast of the modified image with respect to the original image. (c) Performance of the network as a function of noise. The x-axis indicates the percentage of noise that is randomly added to the input pixels.

This finding corroborates the hypothesis that recognition occurs in object specific mechanisms in late areas of the visual stream and is consistent both with the geons theory and with image-based models (Furmanski & Engel, 2000; Tarr & Bülthoff, 1998).

In general, when testing our model, the accuracies were higher for the translation and the rotation variations compared to the scaling condition. The same was observed for other DCNNs and for human subjects, both for their accuracy and their recognition time (Kheradpisheh et al., 2016b). This suggests that translations and rotations are easier to tolerate and need less processing time than scaling variations.

Quality, contrast and noise

Experiment 4: Decreasing the quality of the input

We decreased the resolution of the images by progressively reducing the number of pixels in the input images used for testing, resulting in blurred images. The performance of the network as a function of image resolution is illustrated in Figure 6a. The pattern is similar to the accuracy of the networks tested in Figure 2 of Dodge and Karam (2016): for the first levels of blurring the accuracy does not diminish significantly, however, the network is sensitive to high blurring levels and the performance gradually decreases to chance level when the resolution is reduced to one pixel. The significant reduction in accuracy could be due to the fact that the reduction of quality also removes textures in the input images, which may be a crucial feature used by neural networks for model recognition (Dodge & Karam, 2016).

Experiment 5: Modifying the contrast of the images

Similarly, the contrast of the test images was gradually decreased, from a factor of 1 (original image) to a factor of 0 (grey image) in steps of 0.05, obtaining the performance depicted in Figure 6b. The recognition accuracy over contrast shows a greater robustness with respect to the other deformations, confirming the results obtained in Dodge & Karam (2016) and in Geirhos et al. (2017). As a matter of fact, the accuracies in the contrast experiment of the latter study range from approximately 91 − 94% for VGG-16 (Simonyan & Zisserman, 2014), GoogLeNet (Szegedy et al., 2015) and human subjects and 84% for the AlexNet (Krizhevsky et al., 2012) model when the contrast factor is 1 to chance level for the contrast factor of 0.1. Likewise, the performance of our model starts from the original accuracy of approximately 88.4%, it drops when the contrast factor decreases to less than 0.4 until reaching chance level for a contrast factor of 0.1. A similar performance is achieved by human observers in this task (Geirhos et al., 2017).

In humans, the contrast gain control system evolved as a sophisticated contrast normalization technique and is responsible for the robustness to contrast variations (Geisler & Albrecht, 1995). In order to achieve a greater contrast invariance, images could be normalized in the first layers of the

network, the training set could be augmented with low contrast images or a mechanism similar to the contrast gain control present in humans could be included in the architecture of the network (Geirhos et al., 2017).

Experiment 6: Adding noise to the input images

Noise was added to the test images in various percentages by replacing a random set of the pixels with values drawn from a uniform distribution. The percentage of noise varies between 0% and 75%. When adding noise to the input images, the accuracy rapidly decreases following the trend in Figure 6c, approaching chance level when more than 25% of the input pixels are replaced. This rapid drop in the network performance is in accordance with the studies presented in Dodge and Karam (2016) and Geirhos et al. (2017): after the first 10% of noise is added, the accuracy drops of 12% in our model, whereas it drops of approximately 47% in VGG-16 (Simonyan & Zisserman, 2014) and GoogLeNet (Szegedy et al., 2015) and of 50% in AlexNet (Krizhevsky et al., 2012). In contrast, the drop-in accuracy for human subjects was of only 5%. Handling noise is very challenging for artificial models and drastic differences were found between DNNs and humans in this task, with human subjects outperforming artificial models (Dodge & Karam, 2017). A possible explanation could be that, since the noise was picked from a uniform distribution, it has a high frequency, thus even small changes in the input and in the first layers of the network propagate considerably in higher layers, significantly modifying the output of the model (Dodge & Karam, 2016).

In conclusion, the performance under blur and noise is reduced independently of the artificial model taken into consideration, suggesting that this depends on the architecture and training of the networks (Dodge & Karam, 2016). Therefore, the obvious solution of modifying the model accordingly or training the model on blurred or noisy images arises naturally, although this could consequently compromise the performance of the network with high quality images.

Dropout and Sparsity

Experiment 7: Correlating dropout and sparse representations

Evidence of sparse representations in early visual areas (Simoncelli, 2005; Lennie, 2003; Berry, Warland & Meister, 1997; Reinagel, 2001) and the efficiency of sparse networks (Olshausen & Field, 1996; Olshausen & Field, 2004) motivate the study of the emergence of sparsity in trained networks. Therefore, a possible correlation between dropout and sparsity was tested by varying the amount of dropout in the network. We trained the ten networks resulting from increasing the dropout level in the first dense layer from 0 to 0.9 in steps of 0.1 and hence analyzed the empirical distribution of neuronal responses of the resulting networks to the images in the test set.

In accordance with Baldi and Sadowski (2014), we found that high levels of dropout contribute to sparse representations: the activations of neurons in the first dense layer of the network are significantly closer to 0 in Figure 7b compared to Figure 7a. In particular, Figure 5, representing the mean activations of each layer of the network when presenting the 5000 test images, can be compared to Figure 11.1 in Baldi and Sadowski (2014): there is a clear prevalence of neurons with activations that are close to 0. This correlation derives from the tendency of dropout of preventing each neuron to rely excessively on other units (Hinton et al., 2012), which is achieved by minimizing the variance across neuronal activations. Therefore, sparse representations are favored.

Discussion

In this paper, we compared the processing of information in the visual system to the behavior of a Convolutional Neural Network by analyzing the aspects involved in the implementation of the computational model and in the functioning of the backpropagation algorithm for object recognition. The network's response was tested when presenting input images that were different from the images employed for the training of the network. The test images were modified either with geometric deformations, by varying the rotation, position and size of the objects within the image, or by compromising the extent of visual information transmitted from the input when changing the quality, contrast and amount of noise. The results were compared to neural data obtained from behavioral and neuroimaging studies in which the subjects' response times, accuracies and neural activations were recorded following the presentation of images with the various types of deformations. Furthermore, the fundamental characteristics of the architecture of the network as well as the training process were discussed in comparison to the structure of the visual stream and to the synaptic

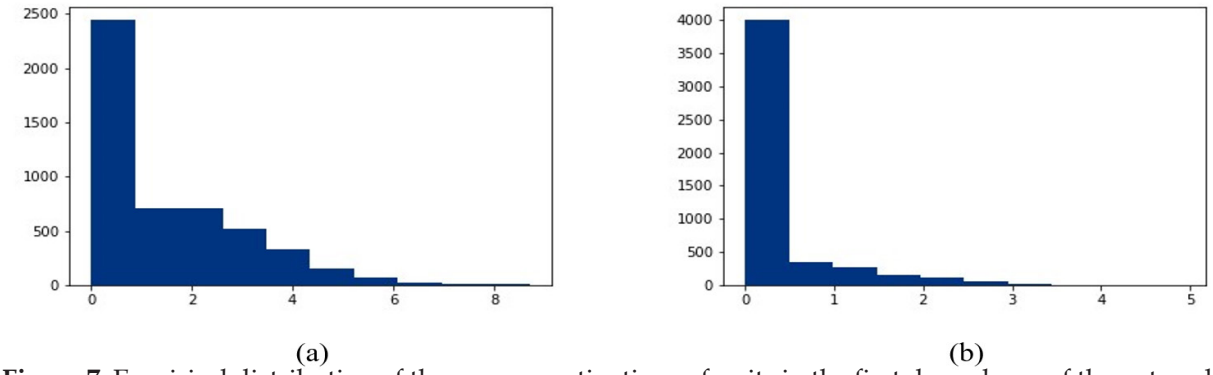


Figure 7. Empirical distribution of the average activations of units in the first dense layer of the network when presenting the 5000 images in the test set, either when using no dropout (a) or a 0.7 dropout level (b). The x-axis indicates the amount of activation; the y-axis represents the number of units corresponding to each amount of activation.

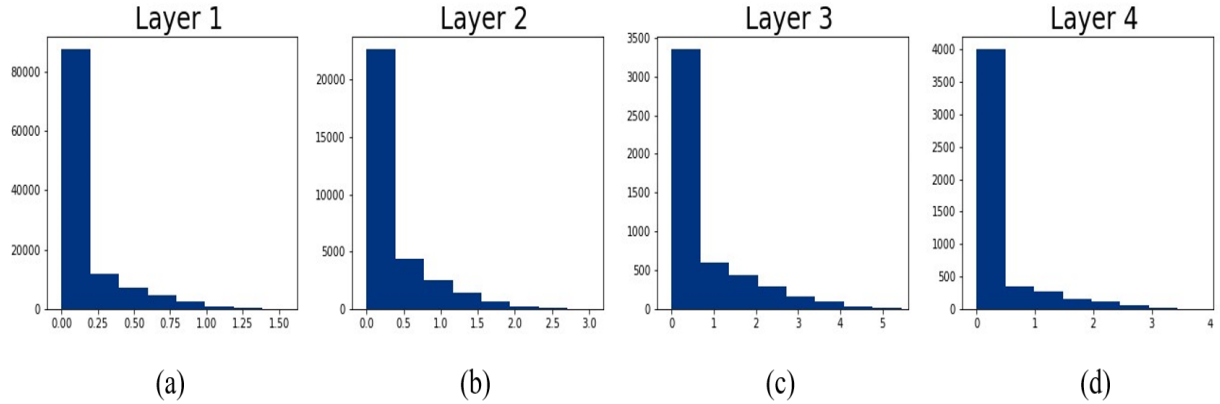


Figure 8. Empirical distribution of the average activations of units in each layer of the network when presenting the 5000 images in the test set. The x-axis indicates the amount of activation; the y-axis represents the number of units corresponding to each amount of activation. (a),(b),(c) Average neuronal activations in the three convolutional layers. (b) Average neuronal activation in the first dense layer. The prevalence of small activations is due to the tendency of dropout of favoring sparse representations.

update processes that are thought to be employed by the brain for learning.

In summary, our results indicate that our computational model is invariant to different kinds of deformations in limited amounts. In many of our experiments and in previous research (Dodge & Karam, 2016, Dodge & Karam, 2017; Geirhos et al., 2017; Kheradpisheh et al., 2016a,b), when increasing the level of image deformation, the accuracy of the networks decreases more rapidly than with human subjects. It was demonstrated that the brain responds differently to distinct kinds of image deformations: for instance, size invariance signals appear earlier than position invariance signals (Isik, Meyers, Leibo & Poggio, 2014) and rotation invariance signals (Dill & Edelman, 1997) in the brain, suggesting that some, but not all, mechanisms for invariant object recognition could be built-in (Nishimura, Scherf, Zachariou, Tarr & Behrmann, 2014). Characterizing the amount of image transformation independently of which type of variation is applied and, therefore, correlating each task to the others is not straightforward. Nevertheless, our results are comparable to the conclusions of other studies, that found both DNNs and the human brain to be more robust to rotation and translation than to scaling of the test images, in which the most amount of visual information is lost (Kheradpisheh et al., 2016b). Indeed, a considerable correlation between computational models and the human brains in terms of categorization accuracy was found (Kheradpisheh et al., 2016b). This suggests that the tasks that have the greatest computational complexity likewise represent the most challenging image variations for humans.

On the one hand, the present results illustrate that the performance of the network rapidly decreases when lowering the quality of the input images, adding noise or modifying the contrast, yet many of the images that are misclassified by the CNN are still recognizable by humans (Geirhos et al., 2017). Moreover, it was found that artificial networks can be easily mislead with low noise percentages (Goodfellow, Shlens & Szegedy, 2014) or fooled into

falsely recognizing objects in images of pure noise (Nguyen, Yosinski & Clune, 2014), even though these conditions are carefully chosen and unlikely to occur. On the other hand, there exist examples of the opposite situation, in which images that have very poor resolution, or a significant amount of noise were successfully classified by artificial networks and not by humans (Wright, Yang, Ganesh, Sastry & Ma, 2009). The present study, along with previous results on different DNNs (Ghodrati et al., 2014; Pinto et al., 2011; Dodge & Karam, 2016, Dodge & Karam, 2017; Geirhos et al., 2017), suggest that, even though neural networks have reached human classification abilities on known benchmarks, there is still a gap between Convolutional Neural Networks and the human visual system when the images are distorted (Dodge & Karam, 2017). This gap could in part be explained by the greater exposition of humans to image transformations compared to DNNs through experience and evolution. Nevertheless, it is true that humans overcome DNNs in their ability to generalize to unseen distortions (Geirhos et al., 2017). These dissimilarities give insights into the aspects that need to be improved in order to bridge the gap between neuroscience and Deep Learning and suggest a starting point for future research, in which for instance the training data could be augmented with distorted images.

It could be argued that the chosen model and training dataset are rather simple compared to the complexity of the human visual system architecture and of the real world tasks: deeper networks have proven to reach higher performance accuracies in different recognition tasks (Krizhevsky et al., 2012; Szegedy et al., 2015; Simonyan & Zisserman, 2014; Yamins & DiCarlo, 2016). Nevertheless, this straightforward design allows for tight experimental control without excessively affecting the performance. This permitted to focus on investigating and replicating the architecture and the behavior of the brain's visual system rather than optimizing its performance on a specific task. As a matter of fact, the chosen network is inspired by neural processes and the single components are, where possible, subject to biological constraints as opposed to various more complex artificial models. Several preliminary experiments were performed in order to select the number of layers, the connectivity between neurons, the dropout level and the activation functions, based on the network's performance as well as the artificial units' activations in comparison with biological data.

This procedure would have been excessively expensive in terms of computational cost if considering deeper architectures. Using different combinations of neural network architectures and datasets while considering the tradeoff between experimental control and model complexity is a crucial next step for future research in order to achieve more accurate and biologically plausible results.

Moreover, it would have been interesting to analyze further aspects of the network's response other than the performance accuracy, as, for instance, the neural unit activation in comparison to brain data obtained in fMRI studies (Kheradpisheh et al., 2016a; Güçlü & van Gerven, 2015). Additionally, it is crucial to test the robustness of artificial networks by constructing alternative experimental designs, with different recognition tests or types of images, i.e. synthetic images (Pinto et al., 2008).

In the following, additional similarities between DNNs and the human visual system along with aspects that question the biological plausibility of the current implementation of neural networks, from the learning procedure to the architecture, will be discussed. For instance, neurophysiological aspects such as spike timing dependent plasticity, dendritic computation, local excitatory-inhibitory networks may explain how gradient descent methods could be implemented in the brain (Marblestone et al., 2016).

An important assumption that makes the comparison between artificial and real networks possible is that the brain has developed cost functions, shaped by evolution, and is able to optimize them in order to adjust the connections across neurons and achieve its goals (Marblestone et al., 2016). Similarly, learning in computational models is based on the optimization of cost functions using backpropagation. The backpropagation algorithm is extremely powerful and is therefore commonly used in neural networks, although it has been widely believed to be biologically implausible (Crick, 1989; Stork, 1989), for various reasons.

To begin with, it requires labelled data for learning, even though almost all real data are unlabeled. Human brains have considerably more degrees of freedom, that is, connections, than seconds of life and consequently than the amount of labelled data they could possibly receive (Hinton, 2014). It is therefore impossible to learn weights for all the synapses in the brain, even though it is unlikely that all the connections need to be used (Hinton, 2016). Moreover, DNNs could employ unsupervised learning, in which learning can occur through unlabeled data and combine it with backpropagation only for fine-tuning the weights or for transfer learning. This technique consists in

exploiting previously learned representations to transfer a prior on the distribution of the input in order to learn new data more easily (Hinton, 2016). Transfer learning could simulate human's ability to learn a new task with few examples, as opposed to the thousands of examples required by current state of the art DNNs. However, for the purposes of the present study, unsupervised learning was not used since it is not necessary for the CIFAR-10 dataset (Paine et al., 2014).

Furthermore, while the standard artificial neurons do not encode precise timing, it is thought that synaptic weights in the brain change with Spike-Timing-Dependent-Plasticity (STDP): potentiate when a presynaptic spike is rapidly followed by a postsynaptic spike and depress when the opposite situation occurs (Gerstner, Kempter, van Hemmen & Wagner, 1996; Markram & Sakmann, 1995). The magnitude of the weight change decreases as the presynaptic and postsynaptic spikes separate. STDP can be seen as a spike-based formulation of the Hebbian postulate, stating that synapses are strengthened if a presynaptic neuron fires slightly before the postsynaptic one (Hebb, 1949). This theory is extended with the concept of synaptic weakening, in which synapses are weakened if the presynaptic cell is consistently not co-active with the postsynaptic neuron (Stent, 1973). However, it was proven that in the visual and motor systems information is mostly carried by the average firing rate of neurons, that is more compatible to learning update rules in artificial models, rather than by the spike timing (Baldi & Sadowski, 2014).

Moreover, artificial units need to send two different kinds of signals: the forward signal, representing its activity and used to generate a prediction, and the backward signal, that is, the derivative of the cost function used for the learning update. On the contrary, there is evidence that feedforward and feedback connections in the brain are implemented in distinct paths and real neurons only use one kind of signal, encoded with spikes (Douglas et al., 1989).

Additionally, DNNs are mainly feedforward networks and the feedback mechanisms occur only during learning, whereas there is evidence for continuous feedback processes in the brain (Bullier, McCourt & Henry, 1988; Felleman & Van, 1991; De Pasquale & Sherman, 2011; Mignard & Malpeli, 1991). The feedback and feedforward computations are implemented in two distinct phases and, since the feedback step needs to follow the feedforward step, a synchronization mechanism is needed (Bengio, Lee, Bornschein, Mesnard & Lin, 2015). This synchronization is not necessary in the case of recurrent neural networks, which are more compatible with neurophysiological processes under this point of view (Simard, Ottaway & Ballard, 1988). The need for two distinct phases and a separate network for the feedback of error is eliminated by associating each neuron with a mirror neuron that imitates the feedforward path in order to cancel the top-down component (Guerguiev, Lillicrap & Richards, 2017). This allows for a network that continuously generates predictions and feedback at the same time. Yet, there is no known biochemical mechanism that could duplicate the weight of a synapse between two cells (Baldi & Sadowski, 2016).

Furthermore, backpropagation assigns blame on a neuronal basis, depending on how each neuron contributed to the error, therefore feedback paths need exact knowledge of the downstream synapses (Bengio et al., 2015). Otherwise stated, in order to compute the global cost function, each neuron would need to know the output of every other neuron, whereas evidence of local learning rules has been found in some regions of the brain (Rolls & Deco, 2002). A global cost function requires the unlikely condition of the weights matrix to be symmetric (Grossberg, 1987), although the use of random weights has proven to work well in practice and gives a good approximation of backpropagation (Lillicrap, Cowden, Tweed & Akerman, 2014; Lillicrap, Cowden, Tweed & Akerman, 2016) when the synaptic signs do not change between feedback and feedforward connections (Liao, Leibo & Poggiom, 2015).

These discrepancies are solved if the error derivatives needed for backpropagation are encoded in the temporal change of the neuronal firing rates (Hinton & McClelland, 1988). This allows the output of a neuron to represent an error derivative at the same time, as it is also indicating the presence or absence of a feature (Whittington & Bogacz, 2015). Consequently, in the learning rule the weight update is proportional both to the presynaptic activity and to the rate of change of the postsynaptic activity, analogously to the STDP learning update (Bi & Poo, 1998). In this framework, STDP could be identified as a form of stochastic gradient descent (Hinton, 2016; Bengio et al., 2015). This learning rule can approximate the differential anti-Hebbian plasticity in which synapses updates depend on the product of the presynaptic firing rate and the time derivative of postsynaptic firing rate (Xie & Seung, 2000).

Another issue with backpropagation is that it requires for each connection to communicate with both positive and negative derivatives. In

contrast, according to Dale's Law (Strata & Harvey, 1999), real synapses do not change sign. However, employing neurons that are either entirely inhibitory or excitatory is unlikely to limit the functions that can be learned (Tripp & Eliasmith, 2016; Parisien, Anderson & Eliasmith, 2008).

Moreover, artificial neurons in the backpropagation algorithm can assume values in a continuous range. On the contrary, real neurons transmit information through binary spikes (Bengio et al., 2015). Nonetheless, backpropagation is very robust to noise, thus the network units could be rounded to 0 or 1, a technique similar to dropout, without compromising the model performance (Hinton, 2016). Finally, backpropagation involves purely linear computations, whereas dendrites can alternate linear and non-linear calculations. A learning rule similar to the one proposed by Hinton and McClelland (1988) solves this problem by including a non-linear term derived from the probability of firing in the weight update (Bengio et al., 2015).

Markedly, the architecture of DNNs is particularly effective in object recognition and resembles the architecture of cortical visual pathways (LeCun et al., 1999). When an object appears in the visual field, the information flows from the retina, through the LGN, to the primary and secondary visual cortex, then to V4 and finally to the inferotemporal cortex (Trappenberg, 2002). The information flows from the retina to the IT in 100 ms and then starts to flow backwards in order to update the synaptic connections. However, if the gaze is interrupted, the feedforward activations detected within the first 100 ms after presentation of the input are similar to the activations of units in Convolutional Neural Networks (Goodfellow et al., 2016).

Similar to the processing of visual information in the brain, a hierarchical structure naturally emerges along the layers of a Deep Neural Network, that process increasingly complex features (Cichy et al., 2016; Güçlü & van Gerven, 2015; Ba & Caruana, 2014), as occurs in successive regions of the visual stream. For instance, the primary visual cortex has a two-dimensional structure that reflects the images encoding the visual information that hit the retina. It is formed of simple and complex cells, which respond to specific shapes or movements respectively (Goodfellow et al., 2016). From these simple aspects, increasingly complex features are represented in successive brain areas, until the encoding of high-level characteristics in the inferotemporal cortex. Analogously, networks with multiple layers will automatically learn to recognize simple features, as edges and color, in the first layers and increasingly complex features, from shapes to higher-order characteristics like faces, in successive layers (Cichy et al., 2016; Güçlü & van Gerven, 2015; Ba & Caruana, 2014). Therefore, all intermediate features with different levels of complexity between the raw data and the final representation of an object can be represented in distinct layers, supporting the network's ability to generalize.

Specifically, the first layers of a DNN exhibit properties similar to early visual areas (Cichy et al., 2016). As a matter of fact, the receptive fields in the visual cortex can be accurately modelled by Gabor filters, in which the weights follow a Gabor function (Marĉelja, 1980; Jones & Palmer, 1987). A new architecture that incorporates Gabor filters into convolutional DNNs has recently been proposed, performing similarly to many known CNNs on the popular benchmarks such as MNIST, CIFAR-10, CIFAR-100 and ImageNet (Luan et al., 2017). By any means, the first layers of a neural network were proven to naturally converge to Gabor filters even when not explicitly programmed to do so (Bengio et al., 2015). Additionally, recent studies found that the penultimate and ultimate layers of a neural network are particularly predictive of V4 and IT neurons' responses respectively (Yamins et al., 2014; Cadieu et al., 2014), specifically when the network is trained with supervised methods (Khaligh-Razavi & Kriegeskorte, 2014). Accordingly, based on similar computational models (Cichy et al., 2016; Güçlü & van Gerven, 2015; Yamins & DiCarlo, 2016), we identify the three convolutional layers in our network with the primary visual cortex (V1), secondary visual cortex (V2), and the visual area V4, respectively, and the first dense layer with the inferior temporal cortex (IT).

In addition, the convolution operation in CNNs is inspired by biological processes in the visual cortex. Some regions in the visual cortex are sensitive to particular areas of the visual field and to specific features of the object such as orientation, shape, and movement in space. (Hubel & Wiesel, 1959; Hibel & Wiesel 1962). Similarly, in convolutional layers each neuronal unit is connected only to a subset of units in the previous layer and each filter is sensitive to a specific shape or feature. Moreover, biological circuits were proven to be able to perform the convolution operation (Cichy et al., 2016), that is thought to occur in simple cells (Serre et al., 2007).

In contrast, the biological plausibility of shared weights, that is, using the same weight matrix for all the input layer, has been questioned, since the brain uses local fields. However, a similar technique to

weight sharing can be approximated (Hinton, 2016): if two regions encoding low-level features in an early layer are close enough, then they share some high-level features in a higher layer, that gives top down supervision for both lower layer features. Thus, learning features in a low-level region helps creating higher level representation from which other low-level regions can extract information. By knowing the region's input as well as the desired high-level features, learning can be considerably accelerated. Therefore, it is possible to transfer information across regions in a layer without transporting weights.

Likewise, the use of the max-pooling operation is open to criticism: Hinton (2014) observed that, although pooling gives a small amount of translation invariance at each layer, it ignores the relations between the parts of each image and loses information that is currently not relevant but could be useful for future tasks. This suggests that more levels of structure are needed in order to properly disentangle the data. Nevertheless, it was proven that biological circuits are able to perform the max-pooling operation (Cichy et al., 2016). This idea has been supported by studies with both intracellular (Lampl, Ferster, Poggio & Riesenhuber, 2004) and extracellular (Gawne & Martin, 2002) recordings. In the simple and complex cells paradigm, the latter are thought to be responsible for the pooling operation (Serre et al., 2007): the size of the receptive fields indeed decreases from the simple to the complex stage.

Regarding the use of dropout in neural networks, its biological counterpart may be the neuronal refractory period, occurring after an action potential, in which the neuron is incapable of firing. Furthermore, by linking a dropout of probability p to a neuron that spikes with probability $1 - p$, Hinton (2016) demonstrated that randomly dropping units in neural networks is similar to the random noise inherent to the spiking rate of biological neurons that follows a Poisson process.

Moreover, the data augmentation technique used to increase the number of training samples in DNNs is biologically plausible. As a matter of fact, it might mimic the learning of invariant object representations in the brain, that occurs through a varied dataset consisting of distinct views of the objects under different viewing conditions. Furthermore, the variability in the input data could also derive from eye movements, such as drifts or saccades. It was shown that in models trained with high levels of data augmentation, the last layers exhibit greater similarities to the responses of IT in humans compared to networks trained without this technique (Hernández-García, Mehrer, Kriegeskorte, König & Kietzmann, 2018).

An additional technique used to train CNNs that biological circuits are able to approximate is batch normalization (Cichy et al., 2016). Indeed, homeostatic plasticity mechanisms in the brain operate a sort of synaptic scaling that minimizes the current into a neuron and is comparable to the application of batch normalization (Turrigiano & Nelson, 2004; Stellwagen & Malenka, 2006; Turrigiano, 2008). However, the normalization statistics change for every timestep and are computed having complete knowledge of the output of all neurons in each layer, which would be impossible for real neurons. Nevertheless, a more biologically plausible technique for normalization was proposed by Liao et al. (2016), that learns running estimates of the mean and variance only in local regions and is computationally efficient.

A last feature regarding the implementation of CNNs that can be compared to mechanisms in the human brain concerns activation functions. The biological counterpart of activation functions is the action potential firing, that determines the firing of a neuron as a function of its input (L. Hodgkin & F. Huxley, 1990). The rectifier function that was used in our network is biologically inspired and compatible with our current knowledge of the functioning of real neurons (Hahnloser et al., 2003).

Conclusion

In conclusion, we aimed to achieve a more complete overview of the differences and similarities between artificial networks and the brain. The question was tackled by comparing the behavior of Deep Neural Networks, inspired by neuroscience, and the human visual system. Deep architectures have the ability of representing properties of increasing complexity and abstraction in distinct layers and are therefore very expressive. As a matter of fact, DNNs have achieved superhuman abilities in many object classification tasks. Yet, there is still a significant gap between Deep Networks and the brain in terms of invariant recognition ability, that may be due in part to some limitations of current computational models. In artificial networks, the lack of the extensive feedback information that is provided to the visual system and used to continuously update and refine visual representations could explain their lower recognition accuracies under image variations. Moreover, CNNs are purely visual, whereas, in the brain, visual information is integrated with input

from many other senses, which is likely to improve its internal object representations, giving it an advantage over artificial networks. Looking at the insights from neuroscience and focusing on the issues discussed in the present paper, Deep Learning can be improved even further towards more efficient and more biologically plausible networks.

References

A. Olshausen, B., & Field, D. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature. 381, 607–609.

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., S. Corrado, G., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., & Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Ba, J. L. & Caruana, R. (2014). Do deep nets really need to be deep? Advances in Neural Information Processing Systems, 27, 2654–2662.

Baldi, P. & Sadowski, P. (2014). The dropout learning algorithm. Artificial Intelligence, 210, 78–122.

Baldi, P. & Sadowski, P. (2016). A theory of local learning, the learning channel, and the optimality of backpropagation. Neural Networks, 83, 51–74.

Bengio, Y. & Delalleau, O. (2011). On the expressive power of deep architectures. Springer-Verlag: Berlin, Heidelberg.

Bengio, Y. & Lecun, Y. (2007). Scaling learning algorithms towards AI. MIT Press.

Bengio, Y., Lee, D.-H., Bornschein, J., Mesnard, T., & Lin, Z. (2015). Towards biologically plausible deep learning.

Berry, M. J., Warland, D. K., & Meister, M. (1997). The structure and precision of retinal spike trains. Proceedings of the National Academy of Sciences, 94(10), 5411–5416.

Bi, G.-q. and Poo, M.-m. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. Journal of neuroscience, 18(24), 10464–10472.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. Psychological review, 94(2), 115–147.

Biederman, I. (2000). Recognizing depth-rotated objects: A review of recent research and theory. Spatial vision, 13(2), 241–253.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning (Information Science and Statistics). Berlin, Heidelberg: Springer-Verlag.

Blumberg, J. & Kreiman, G. (2010). How cortical neurons help us see: visual recognition in the human brain. The Journal of Clinical Investigation, 120(9), 3054–3063.

Booth, M. C. & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. Cerebral cortex, 8(6), 510–523.

Boureau, Y.-L., Ponce, J., & Lecun, Y. (2010). A theoretical analysis of feature pooling in visual recognition. 27th International Conference on Machine Leaning, Haifa, Israel.

Bullier, J., McCourt, M., & Henry, G. (1988). Physiological studies on the feedback connection to the striate cortex from cortical areas 18 and 19 of the cat. Experimental Brain Research, 70(1), 90–98.

Bülthoff, H. H. and Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. Proceedings of the National Academy of Sciences of the United States of America, 89(1):60–64.

Bülthoff, H. H. & Edelman, S. (1993). Evaluating object recognition theories by computer graphics psychophysics. In, 139–164.

C. Wong, S., Gatt, A., Stamatescu, V., & D. McDonnell, M. (2016). Understanding data augmentation for classification: when to warp?

Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., Majaj, N. J., & DiCarlo, J. J. (2014). Deep neural networks rival the representation of primate IT cortex for core visual object recognition. PLOS Computational Biology, 10(12), 1–18.

Cagli, E., Dumas, C., and Prouff, E. (2017). Convolutional neural networks with data augmentation against jitter-based countermeasures – profiling attacks without pre-processing.

Chollet, F. et al. (2015). Keras. https://keras.io.

Cichy, R. M., Khosla, A., Pantazis, D., & Oliva, A. (2017). Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. NeuroImage, 153, 346–358.

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., & Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. Scientific Reports, 6.

Cohen, N., Sharir, O., & Shashua, A. (2016). On the expressive power of deep learning: A tensor analysis. JMLR: Workshop and Conference Proceedings. 49, 1-31

Crick, F. (1989). The recent excitement about neural networks. Nature. 337(6203), 129–132.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems. 2, 303-314.

De Pasquale, R. & Sherman, S. M. (2011). Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse. Journal of Neuroscience, 31(46), 16494–16506.

Dicarlo, J. & Cox, D. (2007). Untangling invariant object recognition. Trends in Cognitive Neuroscience, 11(8), 333–341.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? Neuron, 73(3), 415–434.

Dickinson, S. J. (1999). Object representation and recognition. Cognitive science.

Dill, M. & Edelman, S. (1997). Translation invariance in object recognition, and its relation to other visual transformations.

Dodge, S. F. & Karam, L. J. (2016). Understanding how image quality affects deep neural networks. 2016 Eight International Conference on Quality of Multimedia Experience (QoMEX). Portugal: Lisbon.

Dodge, S. F. and Karam, L. J. (2017). A study and comparison of human and deep learning recognition performance under visual distortions. 26th International Conference on Computer Communication and Networks (ICCCN). Canada: Vancouver.

Douglas, R. J., Martin, K. A., & Whitteridge, D. (1989). A canonical microcircuit for neocortex. Neural Computation. 1(4), 480–488.

Felleman, D. J., & Van, D. E. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cerebral cortex. 1(1), 1–47.

Franco, L., Rolls, E. T., Aggelopoulos, N. C., & Jerez, J. M. (2007). Neuronal selectivity, population sparseness, and ergodicity in the inferior temporal visual cortex. Biological Cybernetics. 96, 547–560.

Furmanski, C. S., & Engel, S. A. (2000). Perceptual learning in object recognition: Object specificity and size invariance. Vision research, 40(5), 473–484.

Gawne, T. J. & Martin, J. M. (2002). Responses of primate visual cortical v4 neurons to simultaneously presented stimuli. Journal of Neurophysiology, 88(3),1128–1135.

Geirhos, R., Janssen, D. H. J., Schütt, H. H., Rauber, J., Bethge, M., & Wichmann, F. A. (2017). Comparing deep neural networks against humans: object recognition when the signal gets weaker. Computer Vision and Pattern Recognition.

Geisler, W., & Albrecht, D. (1995). Bayesian analysis of identification performance in monkey visual cortex: nonlinear mechanisms and stimulus certainty. Vision research, 35(19), 2723–2730.

Gerstner, W., Kempter, R., van Hemmen, J. L., & Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. Nature, 383(6595), 76–78.

Ghodrati, M., Farzmahdi, A., Rajaei, K., Ebrahimpour, R., and Khaligh-Razavi, S.-M. (2014). Feedforward object-vision models only tolerate image variations compared to human. Frontiers in computational neuroscience, 8:74.

Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks, Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, USA: Fort Lauderdale.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. USA: MIT Press.

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. CoRR.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. Cognitive Science, 11(1), 23–63.

Güçlü, U. & van Gerven, M. A. J. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. Journal of Neuroscience, 35, 10005–10014.

Guerguiev, J., Lillicrap, T. P., & Richards, B. A. (2017). Towards deep learning with segregated dendrites. ELife, 6.

Hahnloser, R., Sarpeshkar, R., A. Mahowald, M., Douglas, R., & Sebastian Seung, H. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. ICLR, 405, 947–951.

Hahnloser, R. H. R., Seung, H. S., & Slotine, J.-J. (2003). Permitted and forbidden sets in symmetric threshold-linear networks. Neural Computation, 15(3), 621–638.

Hastad, J. (1986). Almost optimal lower bounds for small depth circuits. Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing, 86, 6–20, USA: New York.

Håstad, J. & Goldmann, M. (1991). On the power of small-depth threshold circuits. Computational Complexity, 1(2), 113–129.

Hebb, D. O. (1949). The organization of behavior: A neuropsychological theory. New York: Wiley.

Hernández-García, A., Mehrer, J., Kriegeskorte, N., König, P., & Kietzmann, T. C. (2018). Deep neural networks trained with heavier data augmentation learn features closer to representations in hit. 2018 Conference on Cognitive Computational Neuroscience.

Hinton, G. (2014) Can the brain do back-propagation?

Hinton, G. (2016) What is wrong with convolutional neural nets?

Hinton, G. E. & McClelland, J. L. (1988). Learning representations by recirculation. Neural Information Processing Systems, 358–366.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. Computer Vision and Neural Computing.

Hochreiter, S. (1991). Untersuchungen zu dynamischen neuronalen Netzen.

Hochreiter, S., Bengio, Y., Frasconi, P., & Schmidhuber, J. (2001). Gradient flow in recurrent nets: the difficulty of learning long-term dependencies.

Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. Journal of Physiology, 148, 574–591.

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology, 160, 106–154.

Huiping, H., Bingfang, W., & Jinlong, F. (2003). Analysis

to the relationship of classification accuracy segmentation scale image resolution.

Ioffe, S. & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. Computer Vision and Neural Computing.

Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. (2014). The dynamics of invariant object recognition in the human visual system. Journal of neurophysiology, 111(1), 91–102.

Jarrett, K., Kavukcuoglu, K., Ranzato, M., & LeCun, Y. (2009). What is the best multi-stage architecture for object recognition? 2009 IEEE 12th International Conference on Computer Vision, 2146– 2153.

Jones, J. P. & Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. Journal of Neurophysiology, 58(6), 1233–1258.

Khaligh-Razavi, S.-M. & Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain it cortical representation. PLoS Computational Biology, 10(11).

Kheradpisheh, S. R., Ghodrati, M., Ganjtabesh, M., & Masquelier, T. (2016a). Deep networks can resemble human feed-forward vision in invariant object recognition. Scientific Reports, 6(32672).

Kheradpisheh, S. R., Ghodrati, M., Ganjtabesh, M., & Masquelier, T. (2016b). Humans and deep networks largely agree on which kinds of variation make object recognition harder. Frontiers in Computational Neuroscience. 92(10).

Kingma, D. P. & Ba, J. L. (2015). Adam: A method for stochastic optimization. ICLR 2015, 2654–2662.

Koffka, K. (1935). Principles of Gestalt Psychology. New York: Harcourt, Brace.

Krizhevsky, A. (2009). Learning multiple layers of features from tiny images. Technical report.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097–1105.

L. Hodgkin, A. & F. Huxley, A. (1990). A quantitative description of membrane current and its application to conduction and excitation in nerve. Bulletin of Mathematical Biology, 52, 25–71.

Lampl, I., Ferster, D., Poggio, T., & Riesenhuber, M. (2004). Intracellular measurements of spatial integration and the max operation in complex cells of the cat primary visual cortex. Journal of Neurophysiology, 92(5), 2704–2713.

Lawrence, S., Giles, C. L., Tsoi, A. C., & Back, A. D. (1997). Face recognition: a convolutional neural network approach. IEEE Transactions on Neural Networks, 8(1), 98–113.

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324.

LeCun, Y., Haffner, P., Bottou, L., & Bengio, Y. (1999). Object recognition with gradient-based learning. Shape, Contour and Grouping in Computer Vision.

Lennie, P. (2003). The cost of cortical computation. Current Biology, 13, 493–497.

Liang, S. & Srikant, R. (2016). Why deep neural networks for function approximation?

Liao, Q., Kawaguchi, K., & Poggio, T. A. (2016). Streaming normalization: Towards simpler and more biologically-plausible normalizations for online and recurrent learning. Computer Vision and Neural Computing.

Liao, Q., Leibo, J. Z., & Poggio, T. A. (2015). How important is weight symmetry in backpropagation? Computer Vision and Neural Computing..

Lillicrap, T., Cownden, D., Tweed, D., & J. Akerman, C. (2014). Random feedback weights support learning in deep neural networks.

Lillicrap, T. P., Cownden, D., Tweed, D. B., & Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. Nature communications. 7(13276).

Loshchilov, I. & Hutter, F. (2017). Fixing weight decay regularization in adam. Computer Vision and Neural Computing.

Luan, S., Zhang, B., Chen, C., Cao, X., Han, J., & Liu, J. (2017). Gabor convolutional networks. Computer Vision and Neural Computing.

M Zur, R., Jiang, Y., Pesce, L., & Drukker, K. (2009). Noise injection for training artificial neural networks: A comparison with weight decay and early stopping. Medical Physics. 36(10), 4810–4818.

Ma, Y. & Klabjan, D. (2017). Convergence analysis of batch normalization for deep neural nets.

Marblestone, A. H., Wayne, G., & Kording, K. P. (2016). Toward an integration of deep learning and neuroscience. Frontiers in Computational Neuroscience, 10.

Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. Journal of Optical Society of America, 70(11), 1297–1300.

Markram, H.& Sakmann, B. (1995). Action potentials propagating back into dendrites triggers changes in efficacy. Society for Neuroscience.

Mignard, M. & Malpeli, J. G. (1991). Paths of information flow through visual cortex. Science, 251(4998), 1249–1251.

Mishkin, D. & Matas, J. (2015). All you need is a good init. Computer Vision and Neural Computing.

Murray, J. E., Jolicoeur, P., McMullen, P. A., & Ingleton, M. (1993). Orientation-invariant transfer of training in the identification of rotated natural objects. Memory & Cognition, 21(5), 604–610.

Nair, V. & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. ICML'10, 807–814, USA.

Nguyen, A. M., Yosinski, J., & Clune, J. (2014). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. Computer Vision and Neural Computing.

Nielsen, M. A. (2018). Neural networks and deep learning.
Nishimura, M., Scherf, K., Zachariou, V., Tarr, M., & Behrmann, M. (2014). Size precedes view: Developmental emergence of invariant object representations in lateral occipital complex. Journal of Cognitive Neuroscience. 27, 1–18.
Olshausen, B. A. & Field, D. J. (2004). Sparse coding of sensory inputs. Current Opinion in Neurobiology, 14(4), 481–487.
Paine, T. L., Khorrami, P., Han, W., & Huang, T. S. (2014). An analysis of unsupervised pre-training in light of recent advances. Computer Vision and Neural Computing.
Parisien, C., H Anderson, C., & Eliasmith, C. (2008). Solving the problem of negative synaptic weights in cortical models. Neural Computation. 20(6), 1473–1494.
Perez, L. & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning.
Pinto, N., Barhomi, Y., Cox, D. D., & DiCarlo, J. J. (2011). Comparing state-of-the-art visual features on invariant object recognition tasks. Applications of computer vision (WACV), 2011 IEEE workshop on, 463–470.
Pinto, N., Cox, D. D., & DiCarlo, J. J. (2008). Why is real-world visual object recognition hard? PLoS Computational Biology, 4(1), 27.
Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. Computer Vision and Neural Computing.
Rampasek, L. & Goldenberg, A. (2016). Tensorflow: Biology's gateway to deep learning? Cell Systems, 2(1), 12–14.
Reddi, S. J., Kale, S., & Kumar, S. (2018). On the convergence of adam and beyond. International Conference on Learning Representations.
Reinagel, P. (2001). How do visual neurons respond in the real world? Current Opinion in Neurobiology, 11(4), 437–442.
Riesenhuber, M. & Poggio, T. (1999). Hierarchical models of object recognition in cortex. Nature Neuroscience, 2, 1019–1025.
Rolls, E. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences. 335(1273), 11–21.
Rolls, E. T. & Deco, G. (2002). Computational Neuroscience of Vision. Oxford University Press.
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Neurocomputing: foundations of research. chapter learning representations by back-propagating errors. 696–699.
Rust, N. C. & DiCarlo, J. J. (2010). Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area v4 to it. The Journal of neuroscience, 30(39), 12978–12995.
S Stent, G. (1973). Stent gs. a physiological mechanism for hebb's postulate of learning. Proceedings of the National Academy of Science USA, 70, 997-1001.
Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976). Quantitative studies of single-cell properties in monkey striate cortex. i. spatiotemporal organization of receptive fields. Journal of neurophysiology, 39(6), 1288–1319.
Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. Proceedings of the National Academy of Science. U.S.A., 104(15), 6424–6429.
Simard, P., Ottaway, M., & Ballard, D. (1988). Analysis of recurrent backpropagation.
Simard, P. Y., Steinkraus, D., & Platt, J. (2003). Best practices for convolutional neural networks applied to visual document analysis. Institute of Electrical and Electronics Engineers, Inc.
Simoncelli, E. (2005). Statistical modeling of photographic images.
Simonyan, K. & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition.
Spetch, M. L. & Friedman, A. (2003). Recognizing rotated views of objects: Interpolation versus generalization by humans and pigeons. Psychonomic Bulletin & Review, 10(1), 135–140.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(1), 1929–1958.
Stellwagen, D. & Malenka, R. C. (2006). Synaptic scaling mediated by glial tnf-α. Nature. 440(7087), 1054.
Stork (1989). Is backpropagation biologically plausible? International 1989 Joint Conference on Neural Networks, 241–246.
Strata, P. & Harvey, R. (1999). Dale's principle. 50, 349–350.
Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. Computer Vision and Pattern Recognition (CVPR).
Tanaka, K. (1996). Inferotemporal cortex and object vision. 19, 109–139.
Tarr, M. J. & Bülthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. Cognition, 67(1), 1–20.
Telea, A. (2004). An image inpainting technique based on the fast marching method. Journal of Graphics Tools, 9(1), 23–34.
Trappenberg, T. (2002). Fundamentals of Computational Neuroscience. Oxford University Press.
Tripp, B. & Eliasmith, C. (2016). Function approximation in inhibitory networks. Neural Networks, 77, 95– 106.
Turrigiano, G. G. (2008). The self-tuning neuron: synaptic scaling of excitatory synapses. Cell, 135(3), 422– 435.
Turrigiano, G. G. & Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. Nature Reviews Neuroscience, 5(2), 97.
Walsh, V. & Kulikowski, J. (1988). Perceptual Constancy: Why Things Look as They Do. Cambridge University

Press.
Werbos, P. & J. Paul John, P. (1974). Beyond regression: new tools for prediction and analysis in the behavioral sciences.
Whittington, J. & Bogacz, R. (2015). An approximation of the error back-propagation algorithm in a predictive coding network with local hebbian synaptic plasticity. Cold Spring Harbor Laboratory.
Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Ma, Y. (2009). Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(2), 210–227.
Xie, X. & Seung, H. S. (2000). Spike-based learning rules and stabilization of persistent neural activity. Advances in Neural Information Processing Systems. 12, 199–208.
Yamins, D. L. K. & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. Nature Neuroscience, 19, 356–365.
Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proceedings of the National Academy of Sciense U.S.A., 111(23), 8619–8624.
Zhou, Y. T. & Chellappa, R. (1988). Computation of optical flow using a neural network. IEEE 1988 International Conference on Neural Networks, 71–78.

Abstracts

Proceedings of the Master's Programme Cognitive Neuroscience is a platform for CNS students to publish their Master theses. Given the number of submissions, we select the articles that received the best reviews, under recommendation of our editors, for the printed edition of the journal. The abstracts of the other articles are provided below, and for interested readers a full version is available on our website: www.ru.nl/master/cns/journal.

Effects of Variations in Neurotrophic 5-HT on Development

Sylvia Docq, Sabrina Hanswijk and Judith Homberg

The aim of this study was to disentangle the neurotrophic influence of 5-HT in development. To this end, heterozygous male offspring of 5-HTT KO and WT rats were subjected to a fostering paradigm. Development was subsequently assessed in terms of maternal care, eye opening, weight, reflex development, olfactory discrimination and open field behaviour. There were distinct differences in maternal care depending on genotype; WT care was characterised by increased grooming while KO care was characterised by increased nursing. During infancy, pups of 5-HTT KO origin under WT care were the swiftest to open their eyes and pups of 5-HTT KO origin under KO care were the heaviest, though no other developmental differences were detected. During adolescence and adulthood, all animals under genotype congruent care were the heaviest and in case of 5-HTT KO origin under KO care also the least anxious. These findings indicate that variations in maternal care styles have differing developmental outcomes depending on the placental 5-HT levels of the offspring.

Fitness-Induced Structural Plasticity in the Brain: Longitudinal Study with 7 Tesla MRI

Giulia Lorenzon

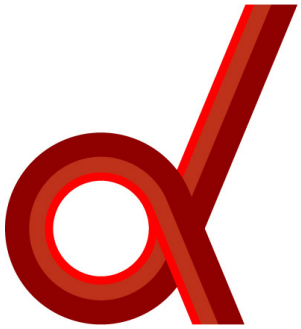
One of the main deleterious features of dementia is neuronal and structural degeneration in the brain. To face that, several attempts through cognitive and physical exercise intervention have been performed, but results are still heterogeneous and controversial. Physical exercise seems to be particularly promising since hippocampal structural plasticity has been shown to be mediated by regional increase in vascular perfusion following training. However, almost nothing is known about the effect of training intervention next to the hippocampus and in motor areas, nor what parameters really mediate training-induced plasticity. Furthermore, the difference between maintenance and actual growth effects are yet to be clearly disentangled. We addressed these questions with a longitudinal approach where 40 young healthy but physically inactive participants (24 females and 16 males, age range: 19-34 years) were enrolled in an intervention study, randomly assigned to either Training (intense aerobic exercise training on a treadmill) or Control condition. 7-Tesla MRI data as well as fitness, cognitive and vascular measures were collected at three time points (baseline, one month and four months) and VBM estimation of local amount of grey matter was performed using CAT12 toolbox. Thanks to the recently developed Sandwich Estimator (SwE) as an innovative, fast and accurate longitudinal statistical model, we investigated linear and quadratic volumetric effects in relation with fitness and cognitive improvement. Our ROI-analyses revealed significant growth on the right hippocampus, together with some interesting trends on the ventromedial side of the frontal cortex. As for cognition, we found significant volume increase in the right middle temporal gyrus associated with improvement in memory recall, along with other plasticity trends related to spatial navigation and working memory. Furthermore, whole-brain post-hoc analyses revealed additional trends towards tissue expansion spread on fronto-temporal networks. Given our longitudinal approach with 3 time points and our young sample, these results cast some light on the causality of this mechanism, suggesting that brain plasticity might be driven by or at least associated with fitness improvement, and it might consequently benefit cognition. Such findings could have relevant implications for prevention and low-cost intervention available to a broader range of people.

Memorization, Lexicalization, Semanticization, and Consolidation of Novel Words in L2.

Giacomo Tartaro, James McQueen, Atsuko Takashima

The memorization, lexicalization, semanticization, and consolidation of novel words have been mainly investigated in the first language (L1). This research explored these processes in a second language (L2) and the influence on these processes of individual differences: length of stay in an L2 environment, proficiency and vocabulary size of L2 on these four processes. Italian-English (late) bilinguals learned a set of 40 English pseudo-word/picture associations on the first day of training (remote condition), and after two days they repeated the procedure with a different set (recent condition) and then performed the testing phase. Only the vocabulary size correlated with the memorization process in the training phase. The outcome of a lexical competition task (pause detection) showed a strong competition effect for the remote condition but not for the recent condition. A primed lexical decision task showed a significant priming effect with the two conditions merged. A recognition memory task in the fMRI showed activation for the remote condition of the inferior frontal gyrus, an area which is thought to unify semantic information of different modalities. Hippocampus, involved in episodic memory and at the first stage of encoding of novel words, was active in both conditions. Overall, the consolidation process seems to not have a one-to-one correspondence with lexicalization and semanticization and a higher dependency on the hippocampus for both the Remote and Recent condition. The lexicalization of the word-form is totally in line with previous literature in L1 but not the semanticization of the meaning which nevertheless shows a trend in line with the previous literature in L1. Finally, word-learning seems to not differ between L1 and L2, at the behavioral level, with some beneficial effects of the vocabulary size in L2 on the memorization process. Also, sleep seems to be beneficial for semanticization and lexicalization effects to arise at the behavioral level.

Institutes associated with the Master's Programme Cognitive Neuroscience



Donders Institute for Brain, Cognition and Behaviour:
Centre for Cognitive Neuroimaging
Kapittelweg 29
6525 EN Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.ru.nl/donders/

Max Planck Institute for Psycholinguistics
Wundtlaan 1
6525 XD Nijmegen

P.O. Box 310
6500 AH Nijmegen
http://www.mpi.nl

Radboudumc

Radboudumc
Geert Grooteplein-Zuid 10
6525 GA Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.umcn.nl/

Baby Research Center
Montessorilaan 3
6525 HR Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.babyresearchcenter.nl