Learning from the brain

This thesis was typeset using (R) Markdown, $\LaTeX$ and the `bookdown` R-package
ISBN: xxx-xx-xxxx-xxx-x
Printing: Acme Press, Inc.

Learning from the brain

Best practices for the use of neuroimaging data in psychology research

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op maandag 21 oktober 2021, te 14 uur

door

Lukas Snoek

geboren te Hoevelaken

Promotiecommissie:

Promotor:	dr. H.S. Scholte	Universiteit van Amsterdam
Copromotor:	dr. S. Oosterwijk	Universiteit van Amsterdam
Overige leden:	prof. dr. R.E. Jack	University of Glasgow
	prof. dr. R.W. Goebel	Maastricht University
	prof. dr. B.U. Forstmann	Universiteit van Amsterdam
	prof. dr. A.H. Fischer	Universiteit van Amsterdam
	prof. dr. D. Borsboom	Universiteit van Amsterdam
	prof. dr. A.G. Sanfey	Radboud University

Faculteit: Faculteit der Maatschappij- en Gedragswetenschappen

Awesome thesis title

Some funky subtitle of my fancy thesis

Lukas Snoek

Contents

Chapter 1

Introduction

The first chapter of the thesis, which introduces your PhD project. The filler-text below was created with the postmodernism generator[1].

Something about the coming about of this thesis. More of a "lessons learned" rather than a coherent research topic.

1.1 The brain is not a dictionary

Something about looking at populations of neurons/voxels/areas rather than simple one-to-one relationships. Shared states.

1.2 The brain (probably) does not care about your hypothesis

Facial expression models.

[1] http://www.elsewhere.org/journal/pomo

1.3 Interpretability and prediction are a trade-off (for now)

A plea for prediction but a cautionary tale for interpreting predictive models (confounds)

1.4 Exploration should be embraced more

Something about the "context of discovery" (cf. TCM), preregistration, and confirmation vs. exploration (Morbid curiosity.)

1.5 Proper generalization is hard

Within and between subject variance is not noise, but unexplained variance (AU limitations).

1.6 Psychology is complex, so it needs complex models

Which need to be fit on complex, large datasets. (AOMIC)

Chapter 2

Shared states: using MVPA to test neural overlap between self-focused emotion imagery and other-focused emotion understanding

This chapter has been published as: Oosterwijk, S.*, Snoek, L.*, Rotteveel, M., Barrett, L. F., & Scholte, H. S. (2017). Shared states: using MVPA to test neural overlap between self-focused emotion imagery and other-focused emotion understanding. *Social cognitive and affective neuroscience, 12*(7), 1025-1035.

* Shared first authorship

Abstract

The present study tested whether the neural patterns that support imagining “performing an action”, “feeling a bodily sensation” or “being in a situation” are directly involved in understanding *other people’s* actions, bodily sensations and situations. Subjects imagined the content of short sentences describing emotional actions, interoceptive sensations and situations (self-focused task), and processed scenes and focused on *how* the target person was expressing an emotion, *what* this person was feeling, and *why* this person was feeling an emotion (other-focused task). Using a linear support vector machine classifier on brain-wide multi-voxel patterns, we accurately decoded each individual class in the self-focused task. When generalizing the classifier from the self-focused task to the other-focused task, we also accurately decoded whether subjects focused on the emotional actions, interoceptive sensations and situations of *others*. These results show that the neural patterns that underlie self-imagined experience are involved in understanding the experience of other people. This supports the theoretical assumption that the basic components of emotion experience and understanding share resources in the brain.

2.1 Introduction

To navigate the social world successfully it is crucial to understand other people. But how do people generate meaningful representations of other people’s actions, sensations, thoughts and emotions? The dominant view assumes that representations of other people’s experiences are supported by the same neural systems as those that are involved in generating experience in the self (e.g., Gallese et al., 2004; see for an overview Singer, 2012). We tested this principle of self-other neural overlap directly, using multi-voxel pattern analysis (MVPA), across three different aspects of experience that are central to emotions: actions, sensations from the body and situational knowledge.

In recent years, evidence has accumulated that suggests a similarity between the neural patterns representing the self and others. For example, a great variety of studies have shown that observing actions and sensations in other people engages similar neural circuits as acting and feeling in the self (see for an overview Bastiaansen et al., 2009). Moreover, an extensive research program on pain has demonstrated an overlap between the experience of physical pain and the observation of pain in other people, utilizing both neuroimaging techniques (e.g., Lamm et al., 2011) and analgesic interventions (e.g., Rütgen et al., 2015; Mischkowski et al., 2016). This process of "vicarious experience" or "simulation" is viewed as an important component of empathy (Carr et al., 2003; Decety, 2011; Keysers & Gazzola, 2014). In addition, it is argued that mentalizing (e.g. understanding the mental states of other people) involves the same brain networks as those involved in self-generated thoughts (Uddin et al., 2007; Waytz & Mitchell, 2011). Specifying this idea further, a constructionist view on emotion proposes that both emotion experience and interpersonal emotion understanding are produced by the same large-scale distributed brain networks that support the processing of sensorimotor, interoceptive and situationally relevant information (Barrett & Satpute, 2013; Oosterwijk & Barrett, 2014). An implication of these views is that the representation of self- and other-focused emotional actions, interoceptive sensations and situations overlap in the brain.

Although there is experimental and theoretical support for the idea of self-other neural overlap, the present study is the first to directly test this process using MVPA across three different aspects of experience (i.e. actions, interoceptive sensations and situational knowledge). Our experimental design consisted of two different tasks aimed at generating self- and other-focused representations with a relatively large weight given to either action information, interoceptive information or situational information.

In the *self-focused* emotion imagery task (SF-task) subjects imagined performing or experiencing actions (e.g., *pushing someone away*), interoceptive sensations (e.g., *increased heart rate*) and situations (e.g., *alone in a park at night*) associated with emotion. Previous research has demonstrated that processing linguistic descriptions of (emotional) actions and feeling states can result in neural patterns of activation associated with, respectively, the representation and generation of actions and internal states (Oosterwijk et al., 2015; Pulvermüller & Fadiga, 2010). Furthermore, imagery-based inductions of emotion have been successfully used in the MRI scanner before (Oosterwijk et al., 2012; Wilson-Mendenhall et al., 2011), and are seen as robust inducers of emotional experience (Lench et al., 2011). In the *other-focused* emotion understanding task (OF-task), subjects viewed images of people in emotional situations and focused on actions (i.e., *How* does this person express his/her emotions?), interoceptive sensations (i.e., *What* does this person feel in his/her body) or the situation (i.e., *Why* does this person feel an emotion?). This task is based on previous research studying the neural basis of emotion oriented mentalizing (Spunt & Lieberman, 2012).

With MVPA, we examined to what extent the SF- and OF-task evoked similar neural patterns. MVPA allows researchers to assess whether the neural pattern associated with one set of experimental conditions can be used to distinguish between another set of experimental conditions. This relatively novel technique has been successfully applied to the field of social neuroscience in general (e.g., Gilbert et al., 2012; Brosch et al., 2013; Parkinson et al., 2014), and the field of self-other neural overlap in particular. For example, several MVPA studies recently assessed whether experiencing pain and observing pain in others involved similar neural patterns (Corradi-Dell'Acqua et al., 2016; Krishnan et al., 2016). Although there is an ongoing discussion about the specifics of shared representation in pain based on these MVPA results (see for an overview Zaki et al., 2016), many authors emphasize the importance of this technique in the scientific study of self-other neural overlap (e.g., Corradi-Dell'Acqua et al., 2016; Krishnan et al., 2016).

MVPA is an analysis technique that decodes latent categories from fMRI data in terms of multi-voxel patterns of activity (Norman et al., 2006). This technique is particularly suited for our research question for several reasons. First of all, although univariate techniques can demonstrate that tasks activate the same brain regions, only MVPA can statistically test for shared representation (Lamm & Majdandžić, 2015). We will evaluate whether multivariate brain patterns that distinguish between mental events in the SF-task can be used to distinguish, above chance level, between mental events in the OF-task. Second, MVPA analyses are particularly useful in research that is aimed at examining distributed representations (Singer, 2012). Based on our constructionist framework, we indeed hypothesize that the neural patterns that will represent self- and other focused mental events are distributed across large-scale brain networks. To capture these distributed patterns, we used MVPA in combination with data-driven univariate feature selection on whole-brain voxel patterns, instead of limiting our analysis to specific regions-of-interest (Haynes, 2015). And third, in contrast to univariate analyses that aggregate data across subjects, MVPA can be performed within-subjects and is thus able to incorporate individual variation in the representational content of multivariate brain patterns. In that aspect within-subject MVPA is sensitive to individual differences in how people imagine actions, sensations and situations, and how they understand others. In short, for our purpose to explicitly test the assumption that self and other focused processes share neural resources, MVPA is the designated method.

We tested the following two hypotheses. First, we tested whether we could classify *self-imagined* actions, interoceptive sensations and situations above chance level. Second, we tested whether the multivariate pattern underlying this classification could also be used to classify the how, what and why condition in the *other-focused* task.

2.2 Methods

Subjects

In total, we tested 22 Dutch undergraduate students from the University of Amsterdam (14 females; M_{age} = 21.48, $s.d._{age}$ = 1.75). Of those 22 subjects, 13 subjects were tested twice in 2 sessions about 1 week apart. Half of those sessions were used for the model optimization procedure. The other half of the sessions, combined with an additional nine subjects (who were tested only once), constituted the model validation set (see Model optimization procedure section). In total, two subjects were excluded from the model validation dataset: one subject was excluded because there was not enough time to complete the experimental protocol and another subject was excluded due to excessive movement (>3 mm within data acquisition runs).

All subjects signed informed consent prior to the experiment. The experiment was approved by the University of Amsterdam's ethical review board. Subjects received 22.50 euro per session. Standard exclusion criteria regarding MRI safety were applied and people who were on psychopharmacological medication were excluded a priori.

Experimental design

Self-focused emotion imagery task

The self-focused emotion imagery task (SF-task) was created to preferentially elicit *self-focused* processing of action, interoceptive or situational information associated with emotion. Subjects processed short linguistic cues that described actions (e.g., *pushing someone away*; *making a fist*), interoceptive sensations (e.g., *being out of breath*; *an increased heart rate*), or situations (e.g., *alone in a park at night*; *being falsely accused*) and were instructed to imagine performing or experiencing the content. The complete instruction is presented in the Supplementary Materials; all stimuli used in the SF-task are presented in Supplementary Table A.1. Linguistic cues were selected from a pilot study performed on an independent sample of subjects (n = 24). Details about this pilot study are available on request. The descriptions generated in this pilot study were used as qualitative input to create short sentences that described actions, sensations or situations that were associated with negative emotions, without including discrete emotion terms. The cues did not differ in number of words, nor in number of characters ($F < 1$).

The SF-task was performed in two runs subsequent to the other-focused task using the software package Presentation (Version 16.4, www.neurobs.com). Each run presented 60 sentences on a black background (20 per condition) in a fully randomized event-related fashion, with a different randomization for each subject. Note that implementing a separate randomization for each subject prevents inflated false positive pattern correlations between trials of the same condition, which may occur in single-trial designs with short inter-stimulus intervals (Mumford et al., 2014). A fixed inter-trial–interval of 2 seconds separating trials; 12 null-trials (i.e. a black screen for 8 seconds) were mixed with the experimental trials at random positions during each run (see Figure 2.1).

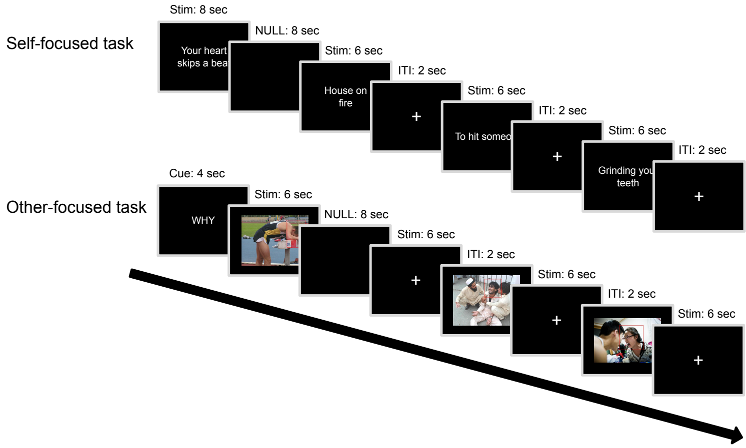


FIGURE 2.1 Overview of the self-focused and other-focused task.

Other-focused emotion understanding task

The other-focused emotion understanding task (OF-task) was created to preferentially elicit *other-focused* processing of action, interoceptive or situational information associated with emotion. Subjects viewed images of people in negative situations (e.g. a woman screaming at a man, a man held at gunpoint). A red rectangle highlighted the face of the person that the subjects should focus on to avoid ambiguity in images depicting more than one person. Image blocks were preceded by a cue indicating the strategy subjects should use in perceiving the emotional state of the people in the images (Spunt & Lieberman, 2012). The cue *How* instructed the subjects to identify actions that were informative about the person's emotional state (i.e., *How* does this person express his/her emotions?). The cue *What* instructed subjects to identify interoceptive sensations that the person could experience (i.e., *What* does this person feel in his/her body). The cue *Why* instructed subjects to identify reasons or explanations for the person's emotional state (i.e., *Why* does this person feel an emotion?). The complete instruction is presented in the Supplementary Materials.

Stimuli for the OF-task were selected from the International Affective Picture System database (IAPS; Lang, 2005; Lang et al., 1997), the image set developed by the Kveraga lab (http://www.kveragalab.org/stimuli.html; Kveraga et al., 2015) and the internet (Google images). We selected images based on a pilot study, performed on an independent sample of subjects (n = 22). Details about this pilot study are available on request.

The OF-task was presented using the software package Presentation. The task presented thirty images on a black background in blocked fashion, with each block starting with a what, why or how cue (see Figure 2.1). Each image was shown three times, once for each cue type. Images were presented in blocks of six, each lasting 6 seconds, followed by a fixed inter trial interval of 2 seconds. Null-trials were inserted at random positions within the blocks. Both the order of the blocks and the specific stimuli within and across blocks were fully randomized, with a different randomization for each subject.

2.3 Procedure

Each experimental session lasted about 2 hours. Subjects who underwent two sessions had them on different days within a time span of 1 week. On arrival, subjects gave informed consent and received thorough task instructions, including practice trials (see the Supplementary Materials for a translation of the task instructions). The actual time in the scanner was 55 minutes, and included a rough 3D scout image, shimming sequence, 3-min structural T1-weighted scan, one functional run for the OF-task and two functional runs for the SF-task. We deliberately chose to present the SF-task after the OF-task to exclude the possibility that the SF-task affected the OF-task, thereby influencing the success of the decoding procedure.

After each scanning session, subjects rated their success rate for the SF-task and OF-task (see Supplementary Figure A.1). In the second session, subjects filled out three personality questionnaires that will not be further discussed in this paper and were debriefed about the purpose of the study.

2.4 Image acquisition

Subjects were tested using a Philips Achieva 3T MRI scanner and a 32-channel SENSE headcoil. A survey scan was made for spatial planning of the subsequent scans. Following the survey scan, a 3-min structural T1-weighted scan was acquired using 3D fast field echo (TR: 82 ms, TE: 38 ms, flip angle: 8°, FOV: 240 × 188 mm, 220 slices acquired using single-shot ascending slice order and a voxel size of 1.0 × 1.0 × 1.0 mm). After the T1-weighted scan, functional T2*-weighted sequences were acquired using single shot gradient echo, echo planar imaging (TR = 2000 ms, TE = 27.63 ms, flip angle: 76.1°, FOV: 240 × 240 mm, in-plane resolution 64 × 64, 37 slices (with ascending acquisition), slice thickness 3 mm, slice gap 0.3 mm, voxel size 3 × 3 × 3 mm), covering the entire brain. For the SF-task, 301 volumes were acquired; for the OF-task 523 volumes were acquired.

2.5 Model optimization procedure

As MVPA is a fairly novel technique, no consistent, optimal MVPA pipeline has been established (Etzel et al., 2011). Therefore, we adopted a validation strategy in the present study that is advised in the pattern classification field (Kay et al., 2008;

Kriegeskorte et al., 2009). This strategy entailed that we separated our data into an optimization dataset to find the most optimal parameters for preprocessing and analysis, and a validation dataset to independently verify classification success with those optimal parameters. We generated an optimization and validation dataset by running the SF-task and OF-task twice, in two identical experimental sessions for a set of thirteen subjects. The sessions were equally split between the optimization and validation set (see Figure 2A); first and second sessions were counterbalanced between the two sets. Based on a request received during the review process, we added nine new subjects to the validation dataset. Ultimately, the optimization-set held 13 sessions and the validation-set, after exclusion of 2 subjects (see Subjects section), held 20 sessions.

In the optimization-set, we explored how different preprocessing options and the so-called 'hyperparameters' in the MVPA pipeline affected the performance of the (multivariate) analyses (visualized in Figure 2.2B; see MVPA pipeline subsection for more details). Thus, we performed the self- and cross-analyses *on the data of the optimization set* multiple times with different preprocessing options (i.e., smoothing kernel, low-pass filter and ICA-based denoising strategies) and MVPA hyperparameter values (i.e., univariate feature selection *threshold* and train/test size ratio during cross-validation). We determined the optimal parameters on the basis of classification performance, which was operationalized as the mean precision value after a repeated random subsampling procedure with 1000 iterations. A list with the results from the optimization procedure can be found in Supplementary Table A.2 and Supplementary Figure A.2. The optimal parameters were then used for preprocessing and the self- and cross-analysis within the validation-set, in which the findings from the optimization-set were replicated. All findings discussed in the @ref{shared-states-results} section

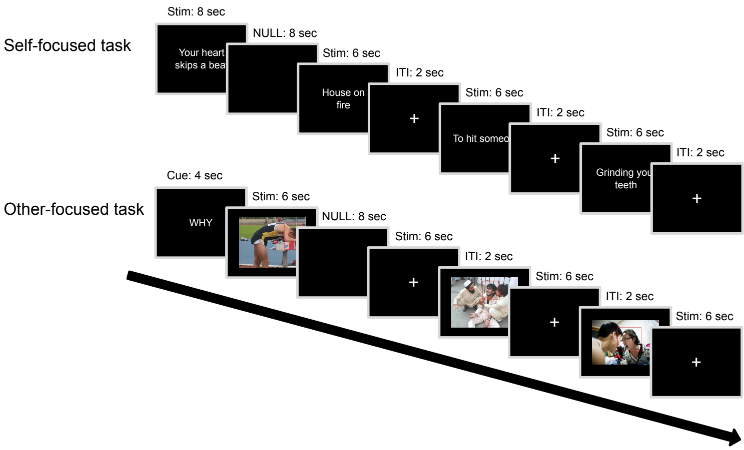


FIGURE 2.2 Mean percentage of trials successfully executed for the SF-task (left panel) and OF-task (right panel). Error bars indicate 95% confidence intervals. A one-way ANOVA of the success-rates of the SF-task (left-panel) indicated no significant overall differences, $F(2, 17) = 1.03$, p = 0.38. In the OF-task (right panel) however, a one-way ANOVA indicated that success-rates differed significantly between classes, $F(2, 17) = 17.74$, $p < 0.001$. Follow-up pairwise comparisons (Bonferroni corrected, two tailed) revealed that interoception-trials ($M = 74.00$, $SE = 2.10$) were significantly less successful ($p < 0.001$) than both action-trials ($M = 85.50$, $SE = 1.85$) and situation trials ($M = 90.00$, $SE = 1.92$).

follow from the validation-set (see Supplementary Figure A.3 for an overview of the findings from the optimization-set).

2.6 Preprocessing and single-trial modeling

2.7 Multi-voxel pattern analysis

MVPA pipeline

Cross-validation scheme and bagging procedure

Statistical evaluation

Spatial representation

2.8 Additional analyses

2.9 Univariate analysis

2.10 Code availability

Chapter 3

Results

Chapter 4

How to control for confounds in decoding analyses of neuroimaging data

Chapter 5

The Amsterdam Open MRI Collection, a set of multimodal MRI datasets for individual difference analyses

Chapter 6

Choosing to view morbid information involves reward circuitry

CHAPTER 7

Using predictive modeling to quantify the importance and limitations of action units in emotion perception

Chapter 8

Comparing models of dynamic facial expression perception

Chapter 9

Summary and general discussion

My view on going forward.

9.1 Explore!

Theories are like toothbrushes, no one likes to use someone else's.

9.2 Think *big*

Big, complex datasets to train big, complex models.

9.3 Rethink psychology education

Embrace and teach interdisciplinary.

Appendices

Appendix A

Supplement to Chapter 2

A.1 Stimuli used for SF-task

Table A.1 Stimuli used for SF-task

Class	Dutch	English translation
Action	Hard wegrennen	Running away fast
	Iemand wegduwen	Pushing someone away
	Iemand stevig vastpakken	Holding someone tightly
	Je hoofd schudden	Shaking your head
	Heftige armgebaren maken	Making big arm gestures
	Ergens voor terugdeinzen	Recoiling from something
	Je ogen dichtknijpen	Closing your eyes tightly
	Je ogen wijd open sperren	Opening your eyes widely
	Je wenkbrauwen fronsen	Frowning with your eyebrows
	Je schouders ophalen	Raising your shoulders
	Op de vloer stampen	Stamping on the floor
	In elkaar duiken	Cowering
	Je schouders laten hangen	Slumping your shoulders
	Je vuisten ballen	Tighten your fists

TABLE A.1 Stimuli used for SF-task *(continued)*

Class	Dutch	English translation
	Je borst vooruit duwen	Push your chest forward
	Je tanden op elkaar zetten	Clench your teeth
	Je hand voor je mond slaan	Put your hand in front of your mouth
	Onrustig bewegen	Moving restlessly
	Heen en weer lopen	Walking back and forth
	Je hoofd afkeren	Turning your head away
Interoception	Een brok in je keel	A lump in your throat
	Buiten adem zijn	Being out of breath
	Een versnelde hartslag	A fast beating heart
	Je hart klopt in de keel	You heart is beating in your throat
	Een benauwd gevoel	An oppressed feeling
	Een misselijk gevoel	Being nauseous
	Druk op je borst	A pressure on your chest
	Strak aangespannen spieren	Tense muscles
	Een droge keel	A dry throat
	Koude rillingen hebben	Cold shivers
	Bloed stroomt naar je hoofd	Blood is going to your head
	Een verdoofd gevoel	A numb feeling
	Je hebt tintelende ledenmaten	Tingling limbs
	Een verlaagde hartslag	A slow heartbeat
	Je hebt zware ledematen	Heavy limbs

TABLE A.1 Stimuli used for SF-task *(continued)*

Class	Dutch	English translation
	Een versnelde ademhaling	Fast breathing
	Je hebt hoofdpijn	Headache
	Je hebt buikpijn	Stomachache
	Zweet staat in je handen	Sweaty palms
	Je maag keert zich om	Your stomach churns
Situation	Vals beschuldigd worden	Being falsely accused
	Dierbare overlijdt	A loved one dies
	Vlees is bedorven	Meat that has gone off
	Je wordt bijna aangereden	You are almost hit by a car
	Iemand naast je braakt	Someone next to you vomits
	Huis staat in brand	House is on fire
	Zonder reden ontslagen worden	Being fired for no reason
	Een ongemakkelijke stilte	An uncomfortable silence
	Alleen in donker park	Alone in a dark park
	Inbraak in je huis	A house burglary
	Een gewond dier zien	Seeing a wounded animal
	Tentamen verknallen	Messing up your exam
	Je partner bedriegt je	You partner cheats on you
	Dierbare is vermist	A loved one is missing
	Belangrijke sollicitatie vergeten	Forgot a job interview
	Onvoorbereid presentatie geven	Giving a presentation unprepared
	Je baas beledigt je	Your boss offends you

TABLE A.1 Stimuli used for SF-task *(continued)*

Class	Dutch	English translation
	Goede vriend negeert je	A good friend neglects you
	Slecht nieuws bij arts	Bad news at the doctor
	Bommelding in metro	A bomb alarm in the metro

Note:

The stimulus materials presented in Table S1 were selected from a pilot study. In this pilot study we asked an independent sample of twenty-four subjects to describe how they would express an emotion in their behavior, body posture or facial expression (action information), what specific sensations they would feel inside their body when they would experience an emotion (interoceptive information), and for what reason or in what situation they would experience an emotion (situational information). These three questions were asked in random order for twenty-eight different negative emotional states, including anger, fear, disgust, sadness, contempt, worry, disappointment, regret and shame. The descriptions generated by these subjects were used as qualitative input in order to create our stimulus set of twenty short sentences that described emotional actions, sensations or situations. With this procedure, we ensured that our stimulus set held sentences that were validated and ecologically appropriate for our sample.

A.2 Instructions

Full instruction for the other-focused emotion understanding task

Translated from Dutch; task presented first.

In this study we are interested in how the brain responds when people understand the emotions of others in different ways. In the scanner you will see images that display emotional situations, sometimes with multiple people. In every image one person will be marked with a red square. While viewing the image we ask you to focus on the emotion of that person in three different ways.

With some images we ask you to focus on HOW this person expresses his or her emotion. Here we ask you to identify expressions in the face or body that are informative about the emotional state that the person is experiencing.

With other images we ask you to focus on WHAT this person may feel in his or her body. Here we ask you to identify sensations, such as a change in heart rate, breathing or other internal feeling, that the person might feel in this situation.

With other images we ask you to focus on WHY this person experiences an emotion. Here we ask you to identify a specific reason or cause that explains why the person feels what he or she feels.

Every image will be presented for six seconds. During this period we ask you to silently focus on HOW this person expresses emotion, WHAT this person feels in his/her body, and WHY this person feels an emotion.

Before you will enter the scanner we will practice. I will show you three images and will ask you to perform each of the three instructions out loud.

It is important to note that there are no correct or incorrect answers, it is about how you interpret the image. For the success of the study it is very important that you apply the HOW, WHAT or WHY instruction for each image. Please do not skip any images and try to apply each instruction with the same motivation. It is also important to treat every image separately, although it is possible that you have similar interpretations for different images.

The three instructions are combined with the images in blocks. In every block you will see five images with the same instruction. Each block will start with a cue that tells you what to focus on in that block.

Each image is combined with all three instructions, so you will see the same image multiple times. In between images you will sometimes see a black screen for a longer period of time.

Do you have any questions?

Full instruction for the self-focused emotion imagery task

Translated from Dutch; task presented second.

In this study we are interested in how the brain responds when people imagine different aspects of emotion. In the scanner you will see sentences that describe aspects of emotional experience. We ask you to try to imagine the content of each sentence as rich and detailed as possible.

Some sentences describe actions and expressions. We ask you to imagine that you are performing this action or expression. Other sentences describe sensations or feelings that you can have *inside* your body. We ask you to imagine that you are experiencing this sensation or feeling. Other sentences describe emotional situations. We ask you to imagine that you are experiencing this specific situation.

We ask you to always imagine that YOU have the experience. Thus, it is about imagining an action or expression of your body, a sensation inside your body, or a situation that you are part of.

I will give some examples now.

For each sentence you have six seconds to imagine the content. All sentences will be presented twice. In between sentences you will sometimes see a black screen for a longer period of time. For this experiment to succeed it is important that you imagine each sentence with the same motivation, even if you have seen the sentence before. Please do not skip sentences.

Do you have any questions?

A.3 Behavioral results

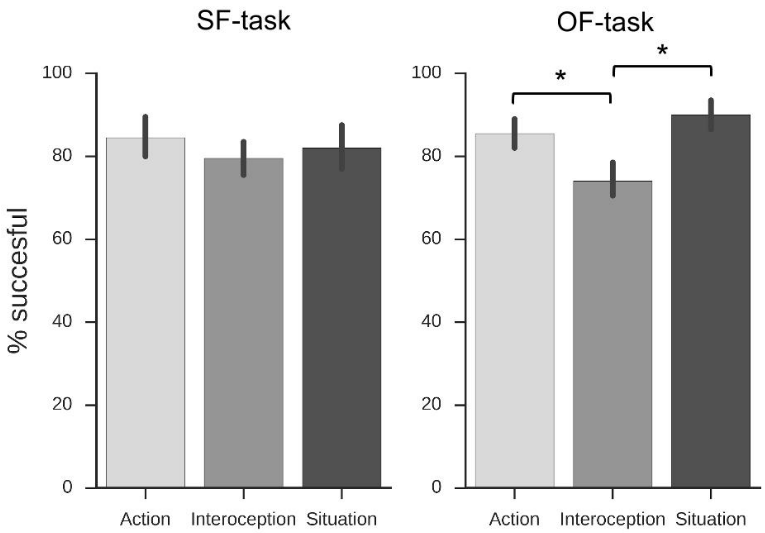


FIGURE A.1 Mean percentage of trials successfully executed for the SF-task (left panel) and OF-task (right panel). Error bars indicate 95% confidence intervals. A one-way ANOVA of the success-rates of the SF-task (left-panel) indicated no significant overall differences, $F(2, 17) = 1.03$, $p = 0.38$. In the OF-task (right panel) however, a one-way ANOVA indicated that success-rates differed significantly between classes, $F(2, 17) = 17.74$, $p < 0.001$. Follow-up pairwise comparisons (Bonferroni corrected, two tailed) revealed that interoception-trials ($M = 74.00$, $SE = 2.10$) were significantly less successful ($p < 0.001$) than both action-trials ($M = 85.50$, $SE = 1.85$) and situation trials ($M = 90.00$, $SE = 1.92$).

A.4 Optimization results

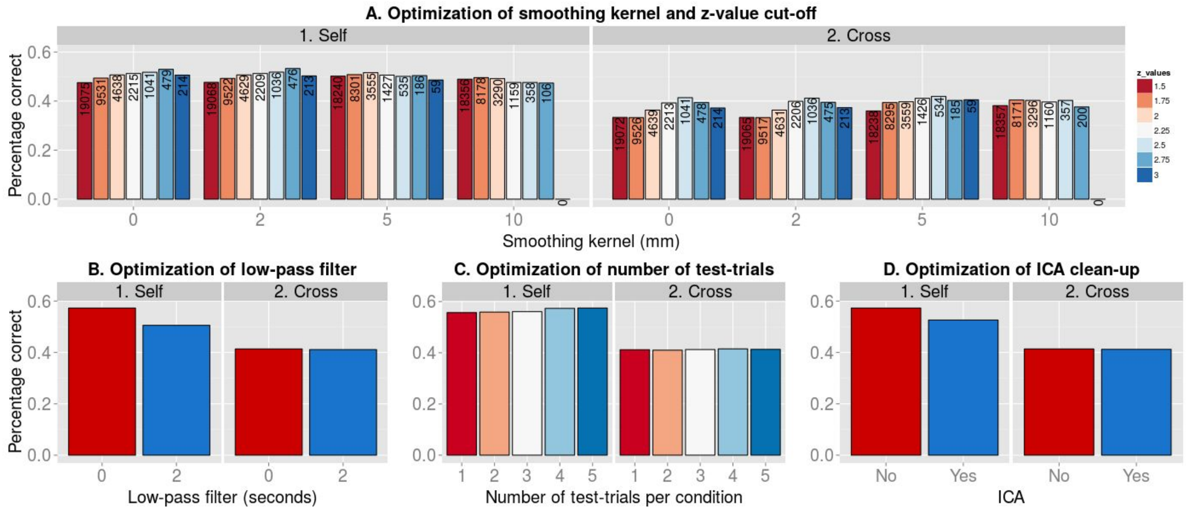


FIGURE A.2 Results of the parameter-optimization procedure. Reported scores reflect the classification scores averaged over subjects and classes (i.e. the diagonal of the confusion matrix). All optimization analyses were iterated 5000 times. **A**) Classification results for different smoothing kernels (0, 2, 5, and 10 mm) and z-value threshold for differentiation scores during feature selection (see MVPA pipeline section in the main text for a description of the particular feature selection method we employed). Numbers reflect the average number of voxels selected across iterations. **B**) Classification results of using a low-pass filter (2 seconds) or not. **C**) Classification results for different numbers of test-trials per class (1 to 5). **D**) Classification results when preprocessing the data with Independent Component Analysis (ICA) or not.

TABLE A.2 Parameters assessed in the optimization set

Parameter	Options	Final choice
Smoothing kernel	0 mm, 2 mm, 5 mm, 10 mm	5 mm
Feature selection threshold	1.5, 1.75, 2, 2.25, 2.5, 2.75, 3	2.3
Number of test-trials	1, 2, 3, 4, 5	4
Low-pass filter	2 seconds vs. none	None
ICA denoising	ICA vs. no ICA	No ICA

Note: The first set of parameters we evaluated in the optimization-set were different smoothing factors and feature selection thresholds (see MVPA pipeline section in the main text). On average, across the self- and cross-analysis, a 5 mm smoothing kernel yielded the best results in combination with a feature selection threshold of 2.25, which we rounded up to 2.3 as this number represents a normalized (z-transformed) score, which corresponds to the top 1% scores within a normal distribution. Next, the difference between using a low-pass (of 2 seconds, i.e. 1 TR) versus none was assessed, establishing no low-pass filter as the optimal choice. Next, different numbers of test-trials (1 to 5) per class per iteration were assessed. Four trials yielded the best results. Lastly, the effect of "cleaning" the data with an independent component analysis was examined (FSL: MELODIC and FIX; Salimi-Khorshidi et al., 2014). Not performing ICA yielded the best results. These parameters – 5 mm smoothing kernel, 2.3 feature selection thresholded, no low-pass filter, and four test-trials per iteration – were subsequently used in the analysis of the validation set.

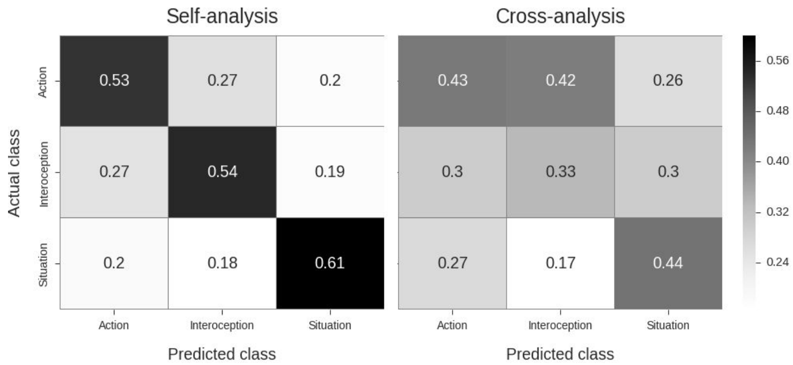


FIGURE A.3 Confusion matrices displaying precision-values yielded by the classification analysis of the optimization dataset with the final set of parameters. Because no permutation statistics were calculated for the optimization set, significance was calculated using a one-sample independent *t*-test against chance-level classification (i.e. 0.333) for each cell in the diagonal of the confusion matrices. Here, all t-statistics use a degrees of freedom of 12 (i.e. 13 subjects - 1) and are evaluated against a significance level of 0.05, Bonferroni-corrected. For the diagonal of the self-analysis confusion matrix, all values were significantly above chance-level, all $p < 0.0001$. For the diagonal of the cross-analysis confusion matrix, both the action (43% correct) and situation (44% correct) classes scored significantly above chance, $p = 0.014$ and $p = 0.0007$ respectively. Interoception was classified at chance level, $p = 0.99$, which stands in contrast with the results in the validation-set.

A.5 Bagging procedure

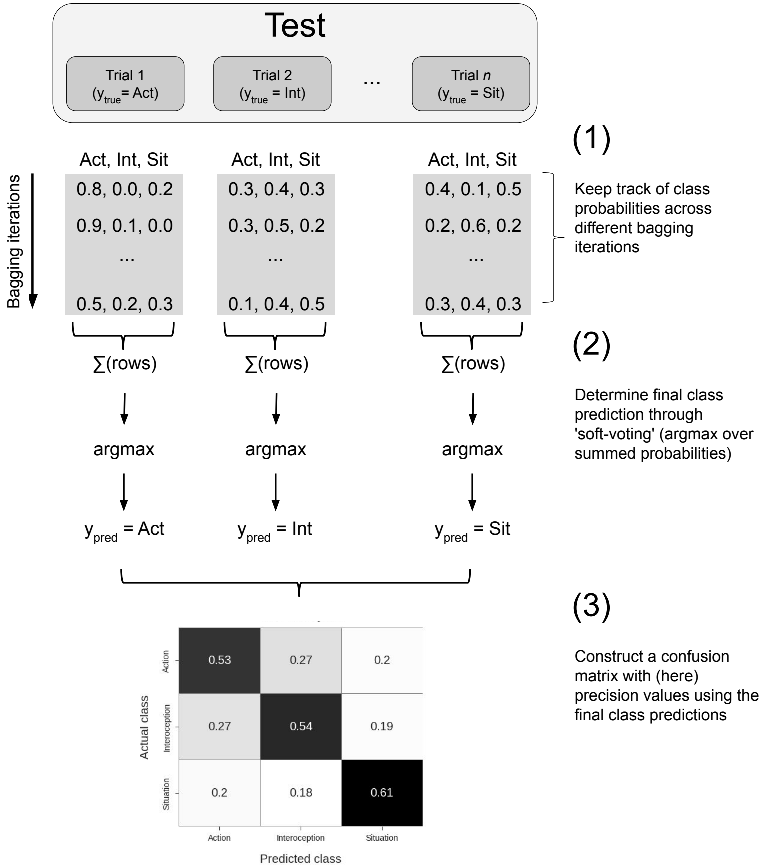


FIGURE A.4 Schematic overview of the bagging procedure. Class probabilities across different bagging iterations are summed and the class with the maximum probability determines each trial's final predicted class, which are subsequently summarized in a confusion matrix on which final recall/precision scores are calculated.

A.6 Precision vs. recall

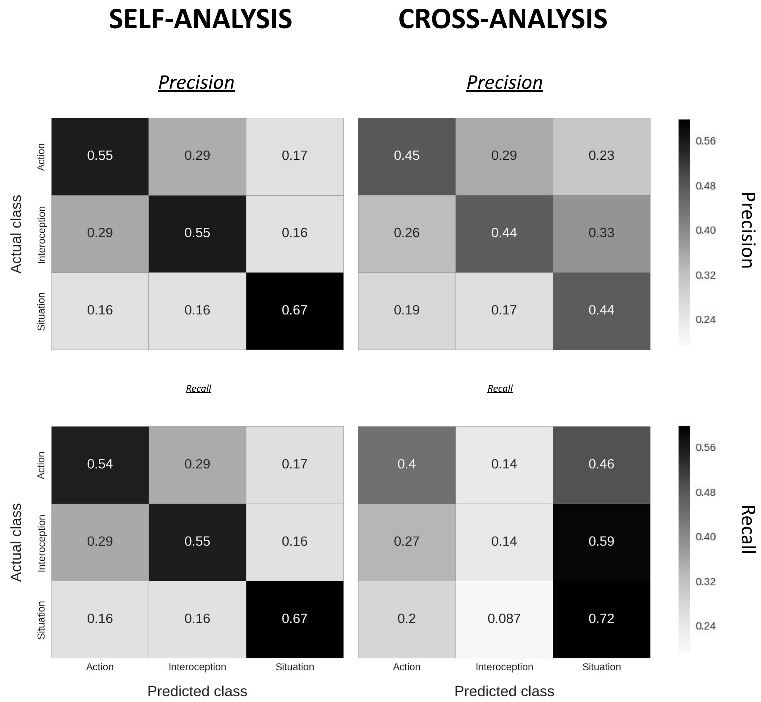


FIGURE A.5 A comparison between precision and recall confusion matrices of the self- and cross-analysis of the validation dataset. Precision refers to the amount of true positive predictions of a given class relative to all predictions for that class. Recall refers to the amount of true positive predictions of a given class relative to the total number of samples in that class. In the self-analysis, all classes were decoded significantly above chance for both precision and recall (all $p < 0.001$). In the cross-analysis, all classes were decoded significantly above chance for precision (all $p < 0.001$); for recall both action and situation were decoded significantly above chance ($p = 0.0013$ and $p < 0.001$, respectively), while interoception was decoded below chance. All p-values were calculated using a permutation test with 1300 permutations (as described in the Methods section in the main text). When comparing precision and recall scores for both analyses, precision and recall showed very little differences in the self-analysis, while the cross-analysis shows a clear difference between metrics, especially for interoception and situation. For the interoception class, the relatively high precision score (44%) compared to its recall score (14%) suggests that trials are very infrequently classified as interoception, but when they are, it is (relatively) often correct. For the situation class, the relatively high recall score (72 %) compared to its precision score (44%) suggests that situation is strongly over-classified, which is especially clear in the lower-right

A.7 Self vs. other classification

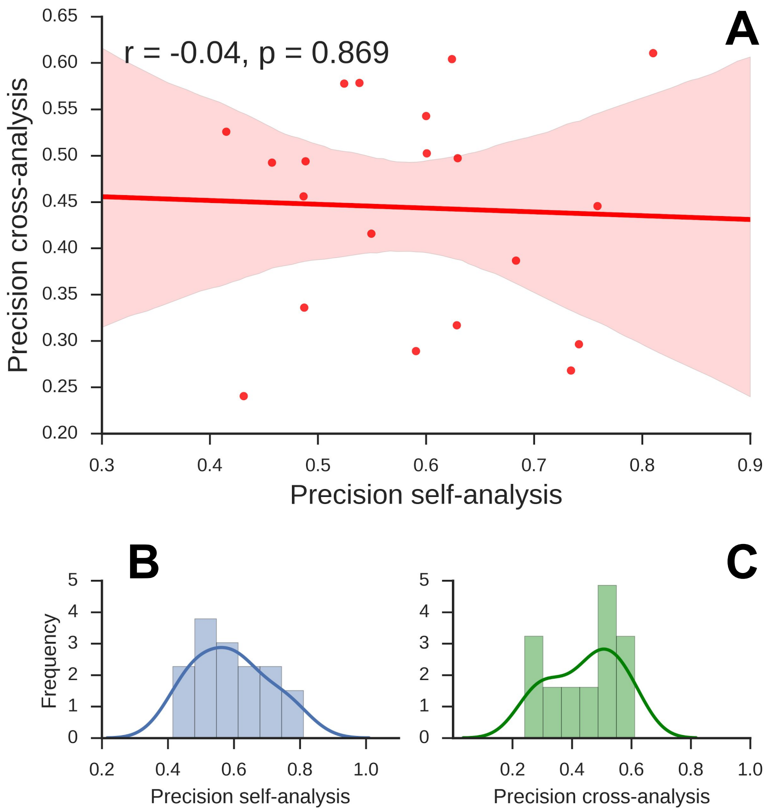


FIGURE A.6 Relation between self- and cross-analysis scores across subjects and their respective distributions. Note that the scores here represent the average of the class-specific precision scores. **A**) There is no significant correlation between precision-scores on the self-analysis and the corresponding scores on the cross-analysis, r = -0.04, p = .86, implying that classification scores in the self-analysis is not predictive of scores in the cross-analysis. **B**) The distribution of precision-scores in the self-analysis, appearing to be normally distributed. **C**) The distribution of precision-scores in the cross-analysis, on the other hand, appears to be bimodal, with one group of subjects having scores around chance level (0.333) while another group of subjects clearly scores above chance level (see individual scores and follow-up analyses in (ref:fig-shared-states-S4).

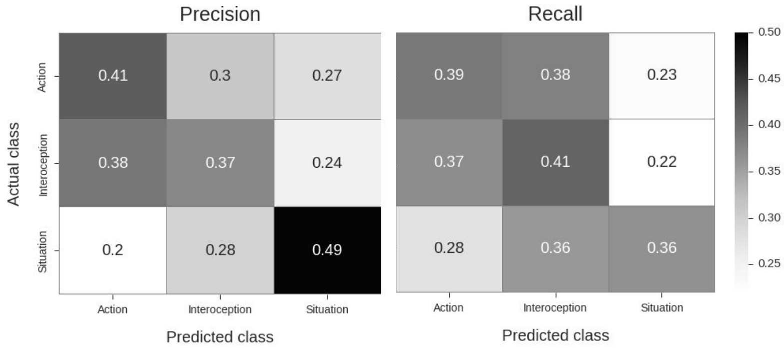


FIGURE A.7 Confusion matrices with precision (left matrix) and recall (right matrix) estimates of the other-to-self decoding analysis. The MVPA-pipeline used was exactly the same as for the (self-to-other) cross-analysis in the main text. *P*-values corresponding to the classification scores were calculated using a permutation analysis with 1000 permutations of the other-to-self analysis with randomly shuffled class-labels. Similar to the self-to-other analysis, the precision-scores for all classes in the other-to-self analysis were significant, p(action) < 0.001, p(interoception) $= 0.008$, p(situation) < 0.001. For recall, classification scores for action and interoception were significant (both $p < 0.001$), but not significant for situation ($p = 0.062$). The discrepancy between the self-to-other and other-to-self decoding analyses can be explained by two factors. First, the other-to-self classifier was trained on fewer samples (i.e. 90 trials) than the self-to-other classifier (which was trained on 120 trials), which may cause a substantial drop in power. Second, the preprocessing pipeline and MVPA hyperparameters were optimized based on the self-analysis and self-to-other cross-analysis. Given the vast differences between the nature of the self- and other-data, these optimal preprocessing and MVPA hyperparameters for the original analyses may not cross-validate well to the other-to-self decoding analysis.

A.8 Condition-average results

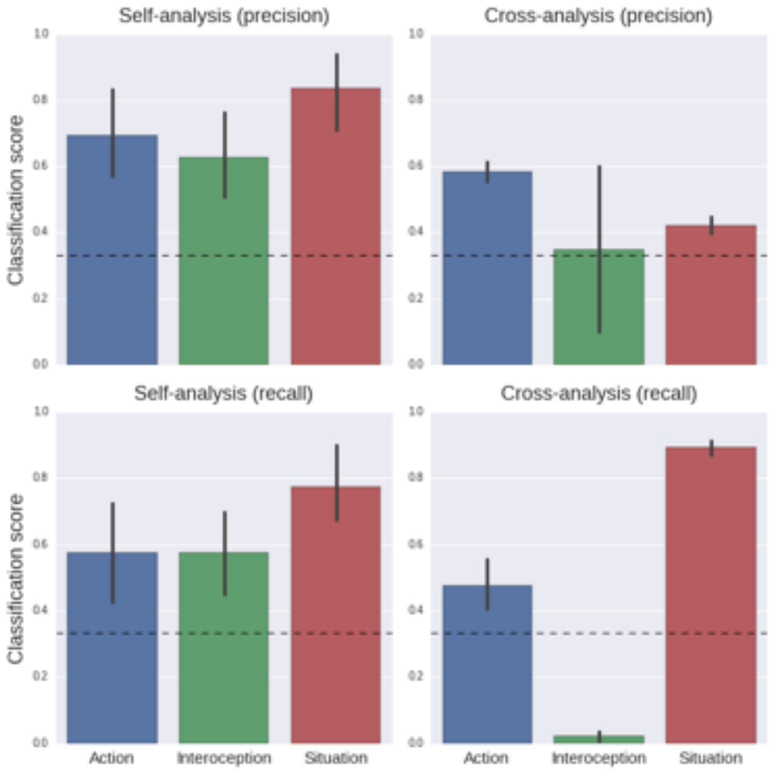


FIGURE A.8 Results of MVPA analyses using condition-average voxel patterns across subjects instead of single-trial patterns within subjects. Here, patterns are estimated in a GLM in which each condition, as opposed to each trial, is modeled with a single regressor, from which whole-brain t-value patterns were extracted. In this condition-average multi-voxel pattern analysis, condition-average patterns across subjects were used as samples. The condition-average patterns were extracted from the univariate first-level contrasts. In total, this yielded 120 samples for the self-data (3 conditions x 2 runs x 20 participants) and 60 samples for the other-data (3 conditions * 20 participants). For these analyses, the same hyperparameters were used as the original analyses reported in the main text, except with regard to the cross-validation and bagging procedure. Here, we used (stratified) 10-fold cross-validation without bagging. The upper panels show precision scores (per class) for the self- and (self-to-other) cross-analysis; the lower panels show results from the same analyses but expressed in recall-estimates (error bars indicate 95% confidence intervals). Apart from interoception in the cross-analysis (both precision and recall), all scores were significant ($p < 0.001$) in a permutation test with 1000 permutations. These results largely replicate our findings as reported in the main text. This

A.9 Individual subject scores

TABLE A.3 Mean general classification scores per subject for the self- and cross-analysis on the validation-set only.

Subject nr.	Self-analysis precision	Cross-analysis precision	Session	Part of optimization-set?
1	0.758	0.445	2	y
2	0.487	0.336	2	y
3	0.629	0.316	1	y
4	0.524	0.577	2	y
5	0.457	0.492	1	y
6	0.741	0.296	2	y
7	0.600	0.542	1	y
8	0.431	0.240	2	y
9	0.629	0.497	1	y
10	0.734	0.268	2	y
11	0.683	0.386	1	y
12	0.415	0.525	2	y
13	0.623	0.604	1	y
14	0.810	0.610	1	n
15	0.538	0.578	1	n
16	0.486	0.455	1	n
17	0.549	0.415	1	n
18	0.488	0.494	1	n
19	0.590	0.289	1	n
20	0.600	0.502	1	n

Note: Supplementary Table 3 suggests individual variability in the extent to which neural resources are shared between self- and other-focused processes. In the SF-task all subjects demonstrated a mean classification score well above .33 (i.e., score associated with chance). When generalizing the SF-classifier to the OF-task, however, the classification scores appear to be bimodally distributed (see Supplementary Figure 5C). As can be seen in Table 3, some subjects demonstrated a relatively high mean classification score (i.e., > .45), whereas other subjects demonstrated a classification score at chance level or lower. Note that there is no significant difference between the OF classification scores for subjects who participated in the experiment for the first or second time ("Session" column in table; *t*(18) = 1.73, p = 0.10), nor for subjects who were or were not part of the optimization-set ("Part of optimization-set?" column in table; *t*(18) = -.95, p = 0.35), suggesting that inclusion in the optimization-set or session ordering is not a confound in the analyses. Regarding individual variability in self-other neural overlap, it is important to note that in the field of embodied cognition, there is increasing attention for the idea that simulation is both individually and contextually dynamic (Oosterwijk & Barrett, 2014; Winkielman, Niedenthal, Wielgosz & Kavanagh, 2015; see also Barrett, 2009). To better distinguish between meaningful individual variation and variation due to other factors (e.g., random noise), future research should test a priori formulated hypotheses about how and when individual variation is expected to occur.

A.10 Brain region importance

TABLE A.4 Most important voxels in terms of their average weight across iterations and subjects.

Brain region	k	Max	Mean	Std
Frontal pole	1827	5.05	2.35	0.52
Occipital pole	1714	5.15	2.45	0.56
Supramarginal gyrus anterior	1573	7.48	2.84	0.91
Lateral occipital cortex superior	1060	4.52	2.18	0.39
Lateral occipital cortex inferior	923	4.73	2.36	0.49
Angular gyrus	856	4.52	2.24	0.40
Supramarginal gyrus posterior	806	4.49	2.29	0.45
Middle temporal gyrus temporo-occipital	798	4.00	2.33	0.48
Temporal pole	711	4.38	2.37	0.54
Precentral gyrus	568	3.54	2.14	0.31

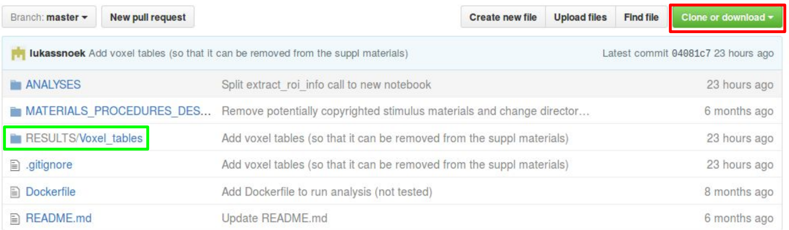
Superior temporal gyrus posterior	549	3.64	2.27	0.41
Superior frontal gyrus	510	3.83	2.18	0.38
Postcentral gyrus	489	4.61	2.43	0.60
Inferior frontal gyrus pars-triangularis	488	4.22	2.35	0.50
Inferior frontal gyrus parsoper-cularis	441	3.54	2.14	0.31
Middle temporal gyrus posterior	417	5.68	2.34	0.52
Occipital fusiform	400	4.28	2.14	0.37
Middle temporal gyrus anterior	398	5.68	2.58	0.76
Middle frontal gyrus	300	3.01	2.06	0.25
Precuneus	282	3.34	2.14	0.31

Note: Brain regions were extracted from the Harvard-Oxford (bilateral) Cortical atlas. A minimum threshold for the probabilistic masks of 20 was chosen to minimize overlap between adjacent masks while maximizing coverage of the entire brain. The column *k* represents the absolute number of above-threshold voxels in the masks. The columns *Max*, *Mean*, and *Std* represent the maximum, mean, and standard deviation from the *t*-values included in the masks. Note that the *t*-values, corresponding to the mean weight across subjects normalized by the standard error of the weights across subjects (after correcting for a positive bias when taking the absolute of the weights), were thresholded at a minimum of 1.75, referring to a *p*-value of 0.05 of a one-sided *t*-test against zero with 19 degrees of freedom (i.e. *n* – 1). Note that this *t*-value map was not corrected for multiple comparisons, and is intended to visualize which regions in the brain were generally involved in our sample of subjects. The *X*, *Y*, and *Z* columns represent the MNI152 (2mm) coordinates of the peak (i.e. max) *t*-value for each listed brain region.

A.11 General note about tables with voxel-coordinates

In order to keep the Supplementary materials concise and orderly, we chose not to include the actual tables with the peak coordinates of all significant clusters of the univariate analyses of the self- and other-task data (as these would amount to 25 pages). These tables, however, can be downloaded (as simple tab-separated-value files) from the study's Github respository[1]. The voxel tables are listed under: `SharedStates/RESULTS/Voxel_tables/*.tsv` (see green box in image below), and can be downloaded by cloning the remote Github repository locally or downloading the ZIP-file from the website (see red box in image below).

[1] https://github.com/lukassnoek/SharedStates


Branch: master
New pull request
Create new file
Upload files
Find file
Clone or download
lukassnoek Add voxel tables (so that it can be removed from the suppl materials)
Latest commit 04081c7 23 hours ago
ANALYSES
Split extract_roi_info call to new notebook
23 hours ago
MATERIALS_PROCEDURES_DES…
Remove potentially copyrighted stimulus materials and change director…
6 months ago
RESULTS/Voxel_tables
Add voxel tables (so that it can be removed from the suppl materials)
23 hours ago
.gitignore
Add voxel tables (so that it can be removed from the suppl materials)
23 hours ago
Dockerfile
Add Dockerfile to run analysis (not tested)
8 months ago
README.md
Update README.md
6 months ago

Appendix B

Supplement to Chapter 4

Appendix C

Supplement to Chapter 6

Appendix D
Supplement to Chapter 7

Appendix E

Supplement to Chapter 8

Appendix F

Data, code and materials

Bibliography

Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current opinion in neurobiology*, *23*(3), 361–372.

Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2391–2404.

Brosch, T., Bar-David, E., & Phelps, E. A. (2013). Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychological science*, *24*(2), 160–166.

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the national Academy of Sciences*, *100*(9), 5497–5502.

Corradi-Dell'Acqua, C., Tusche, A., Vuilleumier, P., & Singer, T. (2016). Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. *Nature communications*, *7*(1), 1–12.

Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emotion review*, *3*(1), 92–108.

Etzel, J. A., Valchev, N., & Keysers, C. (2011). The impact of certain methodological choices on multivariate analysis of fmri data with support vector machines. *Neuroimage*, *54*(2), 1159–1167.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in cognitive sciences*, *8*(9), 396–403.

Gilbert, S. J., Swencionis, J. K., & Amodio, D. M. (2012). Evaluative vs. trait representation in intergroup social judgments: Distinct roles of anterior temporal lobe and prefrontal cortex. *Neuropsychologia*, *50*(14), 3600–3611.

Haynes, J.-D. (2015). A primer on pattern-based approaches to fmri: Principles, pitfalls, and perspectives. *Neuron*, *87*(2), 257–270.

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature*, *452*(7185), 352–355.

Keysers, C., & Gazzola, V. (2014). Dissociating the ability and propensity for empathy. *Trends in cognitive sciences*, *18*(4), 163–166.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature neuroscience*, *12*(5), 535.

Krishnan, A., Woo, C.-W., Chang, L. J., Ruzic, L., Gu, X., López-Solà, M., Jackson, P. L., Pujol, J., Fan, J., & Wager, T. D. (2016). Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *Elife*, *5*, e15166.

Kveraga, K., Boshyan, J., Adams Jr, R. B., Mote, J., Betz, N., Ward, N., Hadjikhani, N., Bar, M., & Barrett, L. F. (2015). If it bleeds, it leads: Separating threat from mere negativity. *Social Cognitive and Affective Neuroscience*, *10*(1), 28–35.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, *54*(3), 2492–2502.

Lamm, C., & Majdandžić, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy–a critical comment. *Neuroscience Research*, *90*, 15–24.

Lang, P. J. (2005). International affective picture system (iaps): Affective ratings of pictures and instruction manual. *Technical report*.

Lang, P. J., Bradley, M. M., Cuthbert, B. N., et al. (1997). International affective picture system (iaps): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, *1*, 39–58.

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological bulletin*, *137*(5), 834.

Mischkowski, D., Crocker, J., & Way, B. M. (2016). From painkiller to empathy killer: Acetaminophen (paracetamol) reduces empathy for pain. *Social cognitive and affective neuroscience*, *11*(9), 1345–1353.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fmri data. *Trends in cognitive sciences*, *10*(9), 424–430.

Oosterwijk, S., & Barrett, L. F. (2014). Embodiment in the construction of emotion experience and emotion understanding. *Routledge handbook of embodied cognition. New York: Routledge*, 250–260.

Oosterwijk, S., Lindquist, K. A., Anderson, E., Dautoff, R., Moriguchi, Y., & Barrett, L. F. (2012). States of mind: Emotions, body feelings, and thoughts share distributed neural networks. *NeuroImage*, *62*(3), 2110–2128.

Oosterwijk, S., Mackey, S., Wilson-Mendenhall, C., Winkielman, P., & Paulus, M. P. (2015). Concepts in context: Processing mental state concepts with internal or external focus involves different neural systems. *Social neuroscience*, *10*(3), 294–307.

Parkinson, C., Liu, S., & Wheatley, T. (2014). A common cortical metric for spatial, temporal, and social distance. *Journal of Neuroscience*, *34*(5), 1979–1987.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature reviews neuroscience*, *11*(5), 351–360.

Rütgen, M., Seidel, E.-M., Silani, G., Riečansky, I., Hummer, A., Windischberger, C., Petrovic, P., & Lamm, C. (2015). Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proceedings of the National Academy of Sciences*, *112*(41), E5638–E5646.

Singer, T. (2012). The past, present and future of social neuroscience: A european perspective. *Neuroimage*, *61*(2), 437–449.

Spunt, R. P., & Lieberman, M. D. (2012). An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *Neuroimage*, *59*(3), 3050–3059.

Uddin, L. Q., Iacoboni, M., Lange, C., & Keenan, J. P. (2007). The self and social cognition: The role of cortical midline structures and mirror neurons. *Trends in cognitive sciences*, *11*(4), 153–157.

Waytz, A., & Mitchell, J. P. (2011). Two mechanisms for simulating other minds: Dissociations between mirroring and self-projection. *Current Directions in Psychological Science*, *20*(3), 197–200.

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia*, *49*(5), 1105–1127.

Zaki, J., Wager, T. D., Singer, T., Keysers, C., & Gazzola, V. (2016). The anatomy of suffering: Understanding the relationship between nociceptive and empathic pain. *Trends in cognitive sciences*, *20*(4), 249–259.

Contributions to the chapters

List of other publications

van der Maas, H.L.J., **Snoek, L.**, & Stevenson, C. (2021). How much intelligence is there in artificial intelligence? A 2020 update.

Hoogeveen, S., **Snoek, L.**, & Van Elk, M. (2020). Religious belief and cognitive conflict sensitivity: A preregistered fMRI study. *Cortex, 129*, 247–265.

van Elk, M., & **Snoek, L.** (2020). The relationship between individual differences in gray matter volume and religiosity and mystical experiences: A preregistered voxel-based morphometry study. *European Journal of Neuroscience, 51*(3), 850–865.

Van Mourik, T., **Snoek, L.**, Knapen, T., & Norris, D. G. (2018). Porcupine: a visual pipeline tool for neuroimaging analysis. *PLoS computational biology, 14*(5), e1006064.

Nederlandse samenvatting (Summary in Dutch)

Replace this with the Dutch title of your thesis

The summary goes here.

Acknowledgments

This section is optional, but theses typically include acknowledgments (*dankwoord* in Dutch) at the end. You may want to mix languages to thank people in their native tongue (though most Dutch speakers write it entirely in Dutch). But the standard language of the thesis template is English. You can switch temporarily by wrapping the text in language tags like so: `[Your Dutch text here]{lang=nl}`. This is important for things like hyphenation to work properly.