

However, we believe that there is no reason to restrict the term ‘pain signature’ to measures of topography alone. At a cerebral level, pain is likely encoded by the neuronal firing ‘intensity’ [6] in neural populations targeted by nociceptive afferents and pain-related wide dynamic range neurons [7]. Ultimately, the topography, response magnitude, and other features such as the time course of responses and interactions between brain regions are likely to be important. Conversely, we should not be too quick to uncritically accept the null hypothesis that painful and nonpainful conditions produce equivalent activation; for example, a recent study claimed that ‘pain matrix’ activation is not specific for pain [8], without fully considering potential differences in activation magnitude and functional reorganization in unique patients.

Third, Hu and Iannetti take issue with the practice of averaging across multiple trials in a condition when testing ML-derived markers or ‘signatures’. They claim that this practice, which we refer to as ‘by-condition’ testing (Figure 1), violates a ‘fundamental rule’ of ML by capitalizing on prior knowledge, leading to ‘artificially high’ prediction accuracy. This is simply not true. Testing models on a by-condition basis is common practice across many types of neuroimaging data [9], including seminal works on functional magnetic resonance imaging (fMRI) pattern recognition (e.g., [10]). Wager *et al.* [5] tested a pain-predictive fMRI ‘signature’ by averaging trials within conditions, and making a prediction about which is more painful. Crucially, in these and related studies, ‘prior knowledge’ about the painfulness of each condition is only used when assessing prediction accuracy, not when making the predictions from test data. One needs to only use prior knowledge of when events in each condition occurred and which trials belong to which conditions, not their painfulness. For example, one might test conditions involving stimulation at 47°C and 49°C, but the intensity information is not used in making the

prediction – they could equally be labeled ‘A’ and ‘B’ (Figure 1).

Averaging over trials is expected to improve prediction accuracy as long as the condition labels are meaningful, by virtue of the central limit theorem (Figure 1). Recently, Lindquist *et al.* [4] applied a ‘pain signature’ to data from six independent data sets ($N = 180$), which revealed solid predictive accuracy at the single-trial level. Based on the central limit theorem, they also estimated the number of trials one would need to average to obtain a criterion level of prediction accuracy (e.g., 90%), demonstrating how averaging over trials increases predictive power.

In summary, Hu and Iannetti initiated an important discussion about the use of ML in pain research. Understanding the nuances of which procedures are valid, and which inferences can and cannot be made, is an ongoing process. Ultimately, in our view ML is a step forward from the statistical parametric mapping and region-of-interest approaches that have dominated the field until recently. Used appropriately, it has the potential to transform the neuroscience of pain, and brain–mind mapping more broadly. We must move forward constructively and with clarity.

¹Klinik für Neurologie, Universitätsklinikum Essen, Hufelandstraße 55, 45147 Essen, Germany

²Department of Psychology and Neuroscience, University of Colorado, Boulder, 345 UCB, Boulder, CO, USA

³The Institute of Cognitive Science, University of Colorado, Boulder, 345 UCB, Boulder, CO, USA

*Correspondence: matthias.zunhammer@uk-essen.de (M. Zunhammer).

<http://dx.doi.org/10.1016/j.tins.2016.07.004>

References

- Hu, L. and Iannetti, G.D. (2016) Painful issues in pain prediction. *Trends Neurosci.* 39, 212–220
- Yarkoni, T. *et al.* (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8, 665–670
- Hutzel, F. (2014) Reverse inference is not a fallacy *per se*: cognitive processes can be inferred from functional imaging data. *Neuroimage* 84, 1061–1069
- Lindquist, M.A. *et al.* (2016) Group-regularized individual prediction: theory and application to pain. *Neuroimage*. <http://dx.doi.org/10.1016/j.neuroimage.2015.10.074>
- Wager, T.D. *et al.* (2013) An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* 368, 1388–1397

- Timmermann, L. *et al.* (2001) Differential coding of pain intensity in the human primary and secondary somatosensory cortex. *J. Neurophysiol.* 86, 1499–1503
- Treede, R.D. *et al.* (1999) The cortical representation of pain. *Pain* 79, 105–111
- Salomons, T.V. *et al.* (2016) The “pain matrix” in pain-free individuals. *JAMA Neurol.* 73, 755–756
- Pereira, F. *et al.* (2009) Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45 (1 Suppl), S199–S209
- Poldrack, R.A. *et al.* (2009) Decoding the large-scale structure of brain function by classifying mental states across individuals. *Psychol. Sci.* 20, 1364–1372

Letter Response

Issues in Pain Prediction – Beyond Pain and Gain

Li Hu^{1,2,*} and
Gian Domenico Iannetti^{2,*}

Zunhammer *et al.* [1] raise concerns that our critiques of the use of machine-learning and functional neuroimaging to predict subjective pain states are based on ‘flawed premises’. Unfortunately, Zunhammer *et al.* argue their case with some of the same spurious reasoning that we criticised in our article [2].

First, Zunhammer *et al.* question whether it matters that a ‘neurological pain signature’ [3] observed using machine-learning analysis of functional magnetic resonance imaging (fMRI) data is based on reverse inference [4]. Reverse inferences, which plague fMRI research, are a serious and well-understood logical fallacy that Zunhammer *et al.* nevertheless defend, affirming that ‘reverse inferences are not inherently flawed, and can be useful if validated successfully’. However, the fact that reverse inferences can be correct (which we clearly stated in Box 1 and the Glossary of our paper [2]) does not imply that they are correct. This seems to be particularly the case with respect to ‘pain signatures’, because not all the validations necessary to estimate the correctness of the reverse inference have been performed (e.g., testing any alleged pain signature against the responses

elicited by graded equisalient non-nociceptive somatosensory stimuli). Notably, when some of these validations have been performed, they have revealed that the brain responses elicited by transient nociceptive stimuli cannot unequivocally indicate that an individual is experiencing pain [5]. This state of affairs, therefore, renders illogical any claim that a pattern of brain activity derived with machine learning is a pain signature. While it is theoretically possible to specifically link features of brain activity to the state of being in pain, a conclusive pain signature has not been yet identified. Premature and unvalidated claims do not help progress but instead hinder it.

Second, Zunhammer *et al.* criticise our postulate that ‘signal magnitude cannot contribute meaningfully to neural signatures of pain’. However, this selective quoting misleadingly omits a crucial aspect that we made clear in our article. We do not claim that signal magnitude cannot contribute to pain signatures in general, but only when machine learning is used to identify a pain signature that cannot be identified by mass-univariate analysis [2]. Because (i) information contained in signal amplitude can be exploited by both machine learning and mass-univariate analysis [5], and (ii) the amplitude of brain activation obtained by mass-univariate analysis (i.e., the ‘pain matrix’) is largely non-pain-specific [6], it follows that the amplitude patterns identified by machine learning are unlikely to uniquely reflect pain. Our conclusion does not imply that machine learning is useless. On the contrary, compared to mass-univariate analysis, machine learning has the additional potential of identifying fine-grained spatial patterns [7] that may uniquely reflect the state of being in pain. We reiterate here what we pointed out in detail: when the objective is to use machine learning to predict pain (Objective 2 in our article), and not to identify a pain-specific signature (Objective 1 in our article), signal amplitude, along with spatial pattern and other response features, ought to be exploited.

Third, Zunhammer *et al.* argue that using prior knowledge about which variable each trial belongs to is acceptable when testing prediction performance in machine learning. However, they seem to be using a different definition of prior knowledge than we used in the article. We reiterate here that by prior knowledge we refer only to the variables that machine learning aims to guess, not to ‘when events occurred’. When machine learning is used to decode unknown experimental variables, exploiting the knowledge of these variables (e.g., by increasing the signal-to-noise of the responses by averaging trials belonging to the same variable) results in artificially high prediction accuracy, which, in turn, makes prediction performance incomparable across different studies. Interestingly, Zunhammer *et al.* agree with us that averaging across trials in the same condition can improve the prediction accuracy [8], but then conclude that ‘one need only use prior knowledge of when events in each condition occurred and which trials belong to which conditions, not their painfulness’. To assess their statement, it is important to consider an elementary principle of nociceptive physiology – that the intensity of reported painfulness largely depends on the intensity of the applied nociceptive stimulus [9]. Indeed, the information related to the condition to which each trial belongs consists of the stimulus intensity of that trial, which determines the intensity of the subsequent painful percept. Testing the prediction performance based on brain responses preliminarily averaged according to stimulus intensity (as in [3]) is therefore an incorrect strategy that unavoidably inflates accuracies of predicting pain. We reiterate that while performing machine learning using average brain responses is appropriate for encoding objectives, such as identifying a condition-specific neural signature [10] (Objective 1 in our article), it is incorrect for decoding an unknown variable (Objective 2 in our article).

On a final note, regardless of the disagreement with the criticisms of Zunhammer

et al., we find it more concerning that our colleagues used counterarguments such as ‘testing models on a by-condition basis is common practice across many types of neuroimaging data, including seminal works on fMRI pattern recognition’. Just because a method is widely used is not an argument for its validity. Uncritical acceptance of what is considered sound practice sets a dangerous precedent that can steer entire fields off course. Methodological scepticism is especially relevant to neuroimaging: a recent report showed that some procedures considered common practice in fMRI analysis are incorrect, undermining the meaningfulness of a large number of studies [11]. Using prior knowledge when testing the prediction performance is just one more example of an inappropriate analytical method. As fellow scientists, we understand the temptation to uncritically adopt the methods of other high-profile studies, especially when doing so may lead to similarly high-profile results, but, in the end, it is ultimately a disservice to the field. We agree with Zunhammer *et al.* that machine learning is a promising analytical approach in pain neuroscience research, but emphasise that it cannot be exempt from rigor.

¹CAS Key Laboratory of Mental Health, Institute of Psychology, Beijing, China

²Department of Neuroscience, Physiology and Pharmacology, University College London, UK

*Correspondence: huli@psych.ac.cn (L. Hu) and g.iannetti@ucl.ac.uk (D. Iannetti).

<http://dx.doi.org/10.1016/j.tins.2016.08.006>

References

1. Zunhammer, M. *et al.* (2016) Issues in pain prediction – more gain than pain. *Trends Neurosci.* 39, 639–640
2. Hu, L. and Iannetti, G.D. (2016) Painful issues in pain prediction. *Trends Neurosci.* 39, 212–220
3. Wager, T.D. *et al.* (2013) An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* 368, 1388–1397
4. Poldrack, R.A. (2011) Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron* 72, 692–697
5. Iannetti, G.D. *et al.* (2013) Beyond metaphor: contrasting mechanisms of social and physical pain. *Trends Cogn. Sci.* 17, 371–378
6. Mouraux, A. *et al.* (2011) A multisensory investigation of the functional significance of the ‘pain matrix’. *Neuroimage* 54, 2237–2249
7. Haynes, J.D. and Rees, G. (2006) Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534

8. Lindquist, M.A. *et al.* (2015) Group-regularized individual prediction: theory and application to pain. *Neuroimage* <http://dx.doi.org/10.1016/j.neuroimage.2015.10.074>
9. Hu, L. *et al.* (2014) Human brain responses to concomitant stimulation of Adelta and C nociceptors. *J. Neurosci.* **34**, 11439–11451
10. Norman, K.A. *et al.* (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* **10**, 424–430
11. Eklund, A. *et al.* (2016) Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7900–7905

Spotlight

Where Are Perceptual Decisions Made in the Brain?

Bijan Pesaran^{1,*} and David J. Freedman²

Recently, Katz and colleagues report the effects of silencing neurons in two brain areas during a perceptual decision-making task. In a landmark result, silencing neurons in the visual cortex impaired performance but silencing neurons whose activity correlates to the decision had no effect.

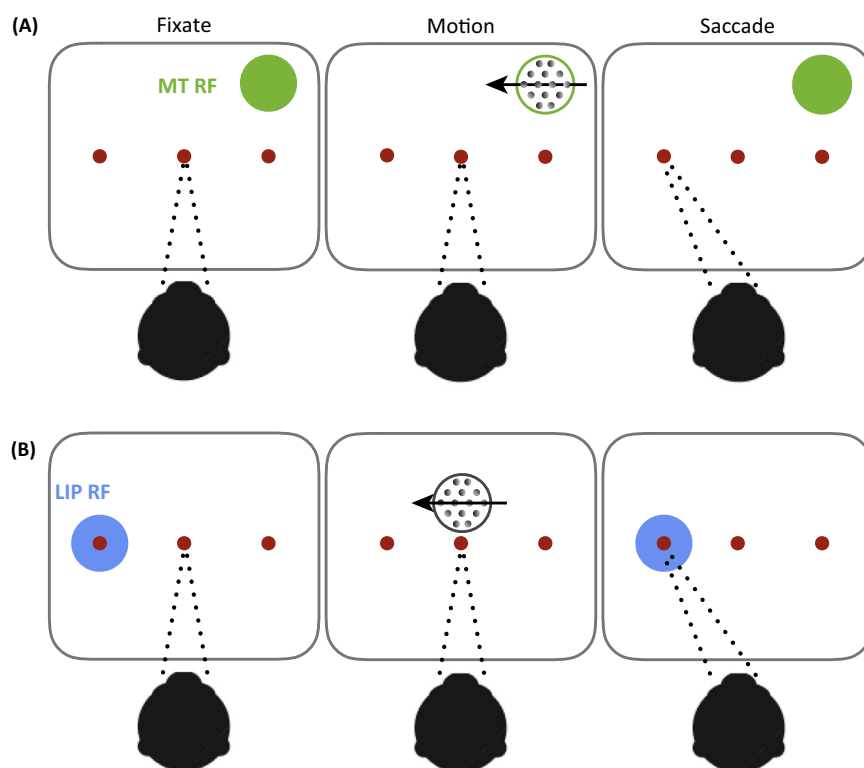
Decision-making is the process by which sensory input is transformed into perception and action. An elegant approach to the puzzle of understanding the neuronal mechanisms underlying perceptual decisions was pioneered by Newsome, Movshon and colleagues using a now widely-studied paradigm, which we refer to as the noisy motion perceptual decision task (Figure 1). In their original studies, Newsome and colleagues recorded from neurons in visual area MT that selectively respond to visual motion while monkeys indicated their perception of stimulus motion via a saccadic eye movement. Their results revealed that individual MT neurons' activity correlated with subject's decisions about motion direction (Figure 1A). This was true even when the stimulus was entirely noise and the monkeys were forced to make guesses about the direction of motion [1]

A critical question raised by these experiments was how, then, is this motion

information read out from MT neurons to guide the eye movement choices? Shadlen, Newsome and colleagues reasoned that a likely candidate was area LIP, to which MT neurons project and which is known to be involved in planning saccadic eye movements. Since LIP neurons were known to be involved in saccade planning, rather than placing the motion stimulus in the neuron's response field (RF; as was done in the MT experiments), an alternative configuration was used in which one of the saccade targets, again, instead of the motion stimulus, was placed in a neuron's RF (Figure 1B). Neuronal recordings revealed that, unlike MT neurons,

LIP neurons' activity was correlated with both the monkeys' trial-by-trial choices and reaction times [2,3]. Computational modeling suggested the animals' perceptual decision was made by the gradual accumulation of motion information in area LIP, offering a seductive view into otherwise private deliberations [3].

Given the weight of this evidence, and other work showing that silencing LIP neurons biases monkeys' eye movement choices [4], we might expect that silencing LIP neurons during the noisy motion task also alters perceptual decisions. However, until recently, studies aiming to silence LIP neurons during performance of this task,



Trends in Neurosciences

Figure 1. Noisy Motion Perceptual Discrimination Task. Monkeys are presented with a visual motion stimulus in which a field of dots appears to move either left or right; the monkeys have to report the direction of stimulus motion with a saccadic eye movement—typically with a leftward or rightward saccade for leftward or rightward motion, respectively. The amount of ‘noise’ in the stimulus is controlled by altering how much the dots appear to move in the same direction. The visual motion percept is generated by integrating the motion stimulus over time and across the field of dots. As expected, monkeys’ decisions are less accurate for noisier stimuli. (A) In the MT configuration the noisy motion stimulus is placed in the response field (RF) of the MT neurons under study and the saccade targets elsewhere. (B) In the LIP configuration one of the saccade targets is placed in the RF of the LIP neurons under study, the other saccade target is placed in the opposite visual hemifield and the noisy motion stimulus is placed centrally.