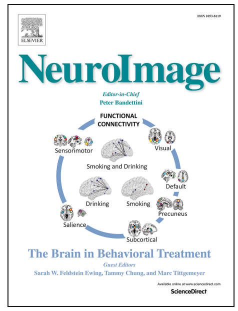
Accepted Manuscript



How to control for confounds in decoding analyses of neuroimaging data

Lukas Snoek, Steven Miletić, H. Steven Scholte

PII: S1053-8119(18)31946-3

DOI: 10.1016/j.neuroimage.2018.09.074

Reference: YNIMG 15310

To appear in: *NeuroImage*

Received Date: 30 April 2018

Revised Date: 4 September 2018

Accepted Date: 25 September 2018

Please cite this article as: Snoek, L., Miletić, S., Scholte, H.S., How to control for confounds in decoding analyses of neuroimaging data, *NeuroImage* (2018), doi: https://doi.org/10.1016/j.neuroimage.2018.09.074.

How to control for confounds in decoding analyses of neuroimaging data

Lukas Snoek[1,2]*, Steven Miletić[1]*, & H. Steven Scholte[1,2]

[1] University of Amsterdam, department of Psychology (Brain and Cognition), Amsterdam, The Netherlands
[2] Spinoza Centre for Neuroimaging, Amsterdam, the Netherlands
* Authors declare equal contributions

Corresponding author: Lukas Snoek, lukassnoek@gmail.com, Nieuwe Achtergracht 129B, 1018WS, Amsterdam, The Netherlands

Abstract

Over the past decade, multivariate "decoding analyses" have become a popular alternative to traditional mass-univariate analyses in neuroimaging research. However, a fundamental limitation of using decoding analyses is that it remains ambiguous which source of information drives decoding performance, which becomes problematic when the to-be-decoded variable is confounded by variables that are not of primary interest. In this study, we use a comprehensive set of simulations as well as analyses of empirical data to evaluate two methods that were previously proposed and used to control for confounding variables in decoding analyses: post hoc counterbalancing and confound regression. In our empirical analyses, we attempt to decode gender from structural MRI data while controlling for the confound "brain size". We show that both methods introduce strong biases in decoding performance: post hoc counterbalancing leads to better performance than expected (i.e., positive bias), which we show in our simulations is due to the subsampling process that tends to remove samples that are hard to classify or would be wrongly classified; confound regression, on the other hand, leads to worse performance than expected (i.e., negative bias), even resulting in significant below chance performance in some realistic scenarios. In our simulations, we show that below chance accuracy can be predicted by the variance of the distribution of correlations between the features and the target. Importantly, we show that this negative bias disappears in both the empirical analyses and simulations when the confound regression procedure is performed in every fold of the cross-validation routine, yielding plausible (above chance) model performance. We conclude that, from the various methods tested, cross-validated confound regression is the only method that appears to appropriately control for confounds which thus can be used to gain more insight into the exact source(s) of information driving one's decoding analysis.

Keywords: MVPA, decoding, neuroimaging, confounds, counterbalancing