NeuroImage 84 (2014) 1061–1069

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Review

Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data

Florian Hutzler *

Centre for Neurocognitive Research & Department of Psychology, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria

ARTICLE INFO

Article history:
Accepted 29 December 2012
Available online 11 January 2013

Keywords:
Functional specificity
Reverse inference
Cognitive process
Cognition
Brain activation

ABSTRACT

When inferring the presence of a specific cognitive process from observed brain activation a kind of reasoning is applied that is called reverse inference. Poldrack (2006) rightly criticized the careless use of reverse inference. As a consequence, reverse inference is assumed as intrinsically weak by many and its validity has been increasingly regarded as limited. Although it is undisputed that the *careless use* of reverse inference is a problematic practice, the current view of reverse inference is to the author's opinion overly pessimistic. The present manuscript provides a revised formulation of reverse inference that includes an additional conditional constraint that has been previously acknowledged, but so far not implemented: the *task-setting*. This revised formulation I.) reveals that reverse inference can have high predictive power (as demonstrated by an example estimation) and II.) allows an estimation of reverse inference on the basis of meta-analyses instead of large-scale databases. It is concluded that reverse inference cannot be disregarded as a fallacy per se. Rather, the predictive power of reverse inference can even be "decisive"—dependent on the cognitive process of interest, the specific brain region activated, and the task-setting used.

Contents

Introduction

Cognitive science attempts to model human experience and behavior and evidence from functional imaging studies can help to validate these models. One kind of reasoning that is applied is to infer the

* Fax: +43 662 8044 5126.
E-mail address: florian.hutzler@sbg.ac.at.

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2012.12.075

involvement of a specific cognitive process from observed brain activation during a task. This kind of reasoning is called *reverse inference*.

Whereas the logic of reverse inference is not problematic per se, researchers often neglect the specificity of the activation of a brain region. As pointed out by Poldrack (2006), a specific brain region can be activated by a wide range of cognitive processes. In such a case, it can be problematic to infer the involvement of a specific cognitive process from the activation of this brain region. In other words, the predictive power of reverse inference can be low. Poldrack (2006) addressed this potential fallacy of reverse inference, cautioned against the widespread careless use of reverse inference and warned that researchers should be circumspect in applying this kind of reasoning.

As a consequence, the validity of reverse inference has been increasingly regarded as limited. Brain activation patterns are considered as a weak indicator of the presence of a cognitive process (Poldrack, 2008, 2011) and reverse inference is assumed to be intrinsically weak (Fox and Friston, 2012). Today, researchers applying reverse inference are quickly regarded as falling for "the" fallacy of reverse inference—resulting in the notion of a *general fallacy of reverse inference*.

Although it is undisputed that the careless use of reverse inference is a problematic practice in neuroimaging, the current view of reverse inference is to the author's opinion overly pessimistic. The present manuscript aims to provide a revised formulation of reverse inference that includes an additional (and quite essential) conditional constraint that has been previously acknowledged, but so far not implemented: the *task-setting*.

The revised formulation I.) reveals that reverse inference can have high predictive power (up to the level of being "decisive") and II.) allows an estimation of reverse inference on the basis of meta-analyses instead of large-scale databases. Meta-analyses provide a fine-grained categorization of comparisons. This is an advantage that will become evident when the importance of an adequate classification of comparisons is discussed below.

In the following, the current formulation of reverse inference (as provided by Poldrack, 2006) will be recapitulated. On this basis, the case for a revised formulation will be made. In general, reverse inference allows the determination of the extent to which a certain brain activation is indicative of the involvement of a specific cognitive process (thereafter abbreviated as "the activation" and "the process").

Fig. 1A indicates that the activation can either co-occur with the process or can take place in the absence of the process. Thus, when we back-trace the activation to the level of processes (i.e., when drawing a reverse inference) we can follow two paths. Human reasoning intuitively accounts for the first path: the probability of an activation *in the presence* of the process. We acknowledge the so-called hit-rate when we refer to studies that manipulated the process of interest and when we analyze whether these studies observed the activation in question. The second path is more often neglected: the probability of an activation *in the absence* of the process (i.e., the false-alarm rate). We seldom discuss studies that *did not* manipulate the process of interest but nevertheless resulted in the activation in question.

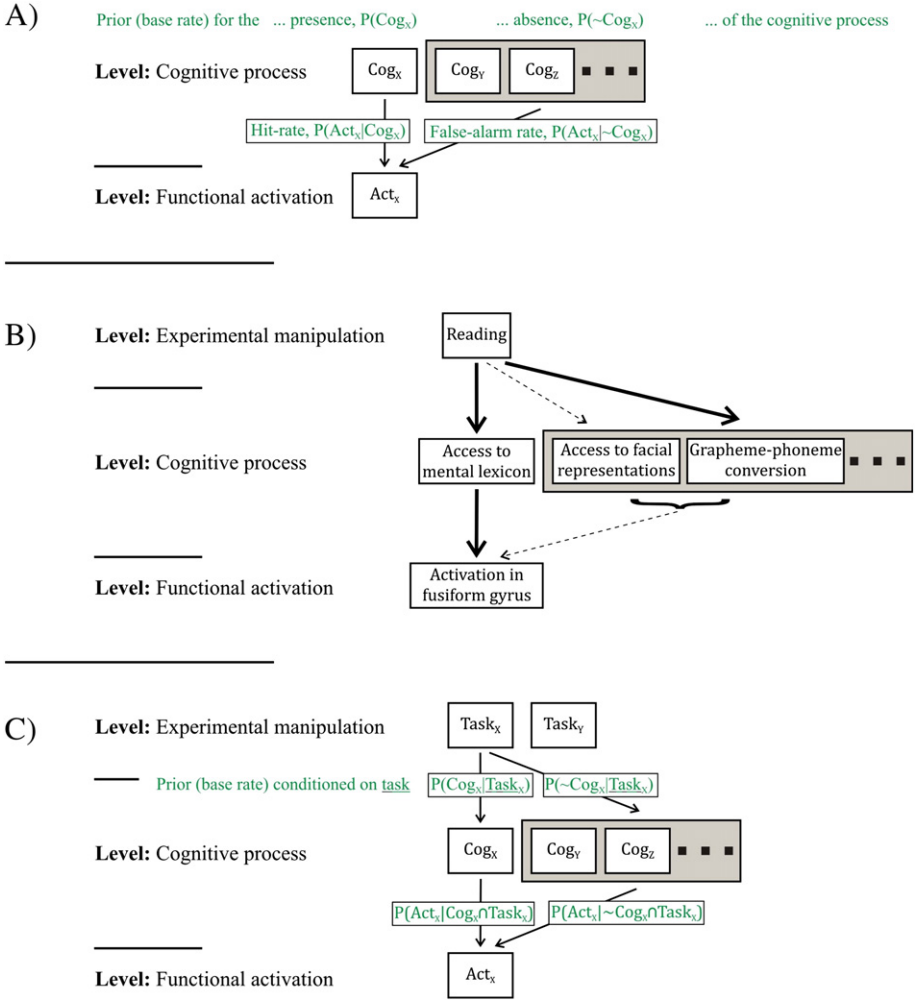

Fig. 1. Estimation of reverse inference, schematic depiction of A.) the formulation as realized in Poldrack (2006), B.) an experiment of thought illustrating the necessity for a conditionalization by task, and C.) the revised formulation as used in the present study.