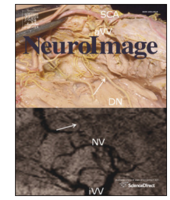
NeuroImage xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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



Interpreting response time effects in functional imaging studies

J.S.H. Taylor [a,*,1], Kathleen Rastle [a,2], Matthew H. Davis [b,3]

[a] Department of Psychology, Royal Holloway University of London, Egham Hill, Egham TW20 0EX, UK
[b] Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK

ARTICLE INFO

Article history:
Accepted 27 May 2014
Available online xxxx

Keywords:
Response time
Reading aloud
Neuroimaging
Regularity
Lexicality
Learning

ABSTRACT

It has been suggested that differential neural activity in imaging studies is most informative if it is independent of response time (RT) differences. However, others view RT as a behavioural index of key cognitive processes, which is likely linked to underlying neural activity. Here, we reconcile these views using the effort and engagement framework developed by Taylor, Rastle, and Davis (2013) and data from the domain of reading aloud. We propose that differences in neural *engagement* should be independent of RT, whereas, differences in neural *effort* should co-vary with RT. We illustrate these different mechanisms using data from an fMRI study of neural activity during reading aloud of regular words, irregular words, and pseudowords. In line with our proposals, activation revealed by contrasts designed to tap differences in neural engagement (e.g., words are meaningful and therefore engage semantic representations more than pseudowords) survived correction for RT, whereas activation for contrasts designed to tap differences in neural effort (e.g., it is more difficult to generate the pronunciation of pseudowords than words) correlated with RT. However, even for contrasts designed to tap neural effort, activity remained after factoring out the RT–BOLD response correlation. This may reveal unpredicted differences in neural engagement (e.g., learning phonological forms for pseudowords > words) that could further the development of cognitive models of reading aloud. Our framework provides a theoretically well-grounded and easily implemented method for analysing and interpreting RT effects in neuroimaging studies of cognitive processes.

Introduction

A key experimental method in both cognitive psychology and cognitive neuroscience involves asking participants to perform specific tasks on selected stimuli and collecting behavioural (accuracy, response time) and/or haemodynamic outcome measures. Statistical comparisons of these measures allow researchers to draw increasingly specific inferences concerning the underlying cognitive and neural processes that contribute to task performance.

However, despite this similarity in approach, psychologists and neuroscientists often differ in their treatment of a behavioural outcome measure – response time (RT) – that is routinely collected in these experiments. Neuroscientists have sometimes argued that RT differences confound comparisons of brain activity between conditions, and have thus employed a variety of approaches to exclude these apparently 'uninteresting' RT-associated neural responses (Binder et al., 2005; Christoff et al., 2001; Crittenden and Duncan, 2012; Graves et al., 2010) or used passive perception designs to minimise the influence of task performance (Ben-Shachar et al., 2011; Pulvermüller et al., 2012; Vinckier et al., 2007; Wright et al., 2011). In contrast, since the time of Donders (1969/1868), behavioural studies have used RT as a key dependent measure to support the inference that different types of stimuli are represented and/or processed in different ways.

In this paper we propose a framework to explain which between-condition differences in neural activity should be independent of RT. We then set out a method for both regressing out and including RT-associated variance when analysing functional magnetic resonance imaging (fMRI) data. We demonstrate the effectiveness of this approach in analysing neuroimaging data collected during reading aloud.

Response time effects in brain imaging and behavioural studies

Evoked haemodynamic responses often increase with the duration of stimulation (Boynton et al., 1996; Horner and Andrews, 2009), and hence should also increase with the time spent on task. This observation has led to concerns regarding the appropriate treatment of neuroimaging contrasts between conditions that differ in RT. The nature of the concern is that two conditions may produce differential activation not because of a qualitative difference in their underlying neural mechanisms, but because stimulus processing in one condition takes longer

* Corresponding author.
E-mail address: j.taylor@rhul.ac.uk (J.S.H. Taylor).

[1] J. S. H. Taylor was supported by a postdoctoral fellowship from the United Kingdom Medical Research Council and the Economic and Social Research Council (U.1055.04.013.00006.01) and by a research fellowship from Newnham College, University of Cambridge, UK.
[2] Kathleen Rastle was supported by an Economic and Social Research Council Research Grant (062-23-2268).
[3] Matthew H. Davis was supported by the United Kingdom Medical Research Council (MC-A060-5PQ80).

http://dx.doi.org/10.1016/j.neuroimage.2014.05.073
1053-8119/© 2014 Published by Elsevier Inc.

than that in the other. Researchers have approached this potential problem in a variety of ways. For example, Crittenden and Duncan (2012) explicitly modelled event duration (RT), allowing them to examine multiple demand network (fronto-parietal cortices) activity under various manipulations of task difficulty, independent of RT. Taking a different approach, Yarkoni et al. (2009) included trial-by-trial RT as a parametric modulator and found that activity in frontal and parietal cortices was positively correlated with RT across several different tasks (working memory, emotional processing, decision making). They suggested that, “RT variability may explain a considerable amount of variance in frontal activation in most tasks” and that this may account for “fMRI effects previously attributed to qualitative differences between experimental conditions” (p. e2457). Yet a different method was used by Binder et al. (2005); a conjunction analysis revealed brain regions in which activity correlated with RT *for all item types* during reading aloud of regular words, irregular words, and pseudowords. It was proposed that RT correlated brain activity *within stimulus type* must arise from “domain general” processing demands. Activation differences *between stimulus types* were therefore only regarded as interesting if they occurred outside of these domain general brain regions. A similar interpretation, although a different method of modelling RT, was applied by Graves et al. (2010) who included multiple psycholinguistic variables, along with RT, as parametric modulators in their analysis of neural activity in an fMRI study of reading. The authors argued that effects of the psycholinguistic variables were of greatest interest if they occurred in areas that did not show positive correlations with RT. Thus, in all these discussed cases it is assumed that differential neural activity only provides evidence of neural specialisation if activation differences cannot be explained by differences in RT.

However, these approaches overlook the information provided by RT variation in behavioural studies. For example, in the Stroop task, patients suffering from psychological disorders are typically slower to name ink colours for words relevant to their clinical condition (Williams et al., 1996), and in the Implicit Association Test, white participants are typically slower to classify black faces and positive words with the same key press than they are to classify black faces and negative words with the same key press (Phelps et al., 2000). In both of these cases, RT differences between conditions indicate underlying processing differences, and we would thus expect differences in neural activity *in regions relevant to performing the task* to correlate with these RT effects, as explicitly demonstrated by Phelps et al. (2000) for the amygdala.

This was acknowledged by Wilson et al. (2009) in their interpretation of neural activity during picture naming. They argued that where RT effects occurred in brain regions in which activity was sensitive to psycholinguistic variables of interest (such as word frequency and concept familiarity) these brain regions were “presumably involved in the stages of word production identified by the other variables in question”. However, RT effects outside of these regions were taken to reflect executive and attentional processes. Whilst this seems sensible, the psycholinguistic variables considered were by no means exhaustive, RT could simply be functioning as a proxy for variables directly relevant to picture naming, but not included in the model, for example, initial phoneme, age-of-acquisition. Similar concerns were raised by Henson (2005) who stated that, as behavioural data (such as RT) and neuroimaging data are both dependent variables, one cannot cause the other. Instead, both are better thought of as different indices of underlying cognitive processes. This was in fact the approach taken in two later studies by Wilson et al. (2010, 2014). RT was used as a proxy for syntactic complexity when examining activity in inferior frontal gyrus and anterior temporal lobe during syntactic processing in neuropsychological patients.

The effort and engagement framework

We argue that separating informative from non-informative differences in neural activity between conditions of interest is not as simple as controlling for effects of RT, or examining the overlap and separation of effects of RT and variables of interest. Instead, it is essential to have a theory that specifies whether and why differences between conditions should (or should not) be independent of RT in order to know how best to treat RT in neuroimaging studies. One framework that provides a way to relate cognitive processes to neural activity was set out by Taylor et al. (2013). We proposed that two principles govern the relationship between cognitive processes and aggregate measures of neural activity such as Blood Oxygenation Level Dependent (BOLD) fMRI: 1) engagement — stimuli that are represented by a model component or brain region should activate it more than stimuli that are not represented by a component or region; and 2) effort — within a set of stimuli that are represented by a model component or brain region; those that fit the representations less well should be more effortful to process, and thus produce greater activity, than those that fit the representations extremely well. As discussed in Taylor et al., the framework critically assumes that computational processes that are functionally separated in cognitive models can be mapped onto separate brain processes (Henson, 2005, 2006a, 2006b).

As illustrated in Fig. 1, this proposal implies an inverted u-shaped relationship between the BOLD signal and the fit between stimuli and neural representations. The upward going portion of the curve is driven by greater engagement for stimuli which fit representations than for stimuli that do not. This is consistent with the majority of ‘subtraction’ studies in which differential neural activity is seen in regions that respond more to a preferred stimulus type than to other stimulus types. For example, a region in the right fusiform gyrus responds more strongly to faces than to other visual stimuli such as houses (Kanwisher et al., 1997), reflecting greater neural engagement for represented than non-represented stimuli. In contrast, the downward going portion of the inverted u-shaped function is driven by reduced effort for stimuli that fit the representations very well as compared to those that fit less well. This is consistent with repetition suppression or familiarity effects in functional imaging studies: highly familiar stimuli typically elicit reduced activity compared to less familiar stimuli (e.g., common versus uncommon orientations of an object), potentially due to sharpening of neural responses, or other mechanisms (Grill-Spector et al., 2006).

This inverted u-shaped relationship is thus needed to account for the existing functional imaging literature (see Taylor et al., 2013 for further details) and is related to other proposals of a non-linear relationship between the BOLD signal and cognitive processing, e.g., Price and Devlin (2011). A clear advantage of our proposal is that effort and engagement readily map onto cognitive distinctions (e.g., represented vs. non-represented stimuli, processing-time differences) that can be used to guide interpretation of neuroimaging contrasts, as detailed in the following paragraph.

Our framework suggests that a stimulus type that is represented by a particular brain region should engage that region more than another

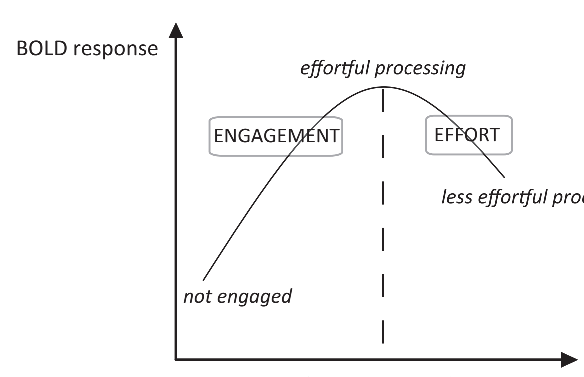


Fig. 1. Inverted U-shaped function showing how engagement and processing effort relate to blood oxygenation level dependent (BOLD) signal.

Please cite this article as: Taylor, J.S.H., et al., Interpreting response time effects in functional imaging studies, NeuroImage (2014), http://dx.doi.org/10.1016/j.neuroimage.2014.05.073