Unbiased ROI selection in neuroimaging studies of individual differences

Roberto Viviani, PhD

Department of Psychiatry III
University of Ulm, Germany

E-mail: roberto.viviani@uni-ulm.de

Abstract

Individual differences studies are those that investigate the interaction between a subject level covariate, such as sex, age, or performance, and the effect of an experimental task. Commonly, a brain region is selected on the basis of the task effect, and the signal in this region correlated with individual covariates. It is shown here that, provided that data are identically and independently distributed between subjects, the selection of the region on the basis of the task effect is unbiased in two cases: when selection is based on a one-sample *t* test of the task effect, or when the subject level covariate is centered. This result is meant to contribute to clarifying when using the same data for ROI selection and testing leads to valid tests in studies of individual differences.

Unbiased ROI selection in neuroimaging studies of individual differences

Introduction

Attention has recently been drawn to possible problems in tests in wich the selection of a region of interest (ROI) is followed by another test carried out on the ROI average data (Kriegeskorte et al. 2009, Vul et al. 2009). This work has the merit to have raised the awareness in the neuroimaging community of possible problems introduced by ROI selection procedures, and of the important problem of validity of the statistical analysis more generally. However, the claims by Vul et al. (2009) of lack of validity of a large number of neuroimaging studies have aroused controversy (Barrett 2009, Lazar 2009, Lieberman et al. 2009, Lindquist and Gelman 2009, Nichols and Poline 2009, Yarkoni 2009).

This technical note examines the procedure in which the ROI is selected on the basis of the task effect as a preliminary step to compute a correlation with individual variables. The context in which this ROI selection procedure commonly arises is when one correlates at the second level a between-subjects covariate, such as sex, age, or performance, with the signal selected for significance in a first test of activation induced by an experimental task (Figure 1). In this type of design, the procedure capitalizes on the capacity of the test for the task effect to select the relevant region, since task effects are much easier to detect than their modulation by individual differences (task effects are main effects, while individual differences are interactions). In the survey of Vul et al. (2009), ROI selection on the basis of task was adopted by about half of the reviewed studies in social and emotional neuroscience. This note will not be concerned with procedures in which ROI is selected on the basis of the between-subjects covariate itself (for the issues involved in that case, see the commentaries to Vul et al. (2009) and their reply in the same issue of the journal), or outside the context of

studies of individual differences. Hence, in this note the term ‘ROI selection’ without further qualification will always mean ‘ROI selection on the basis of the task effect’.

INSERT FIGURE 1 ABOUT HERE

For the ROI selection procedure to be valid, it is necessary that the explanatory covariate on the basis of which the ROI is selected be orthogonal to the explanatory covariate of the second test (Friston et al. 2006, Kriegeskorte et al. 2009). If this is not the case, and the same dataset is used both to select the ROI and to carry out the regression of the signal in the ROI on the individual covariate, then the measures of effect leading to ROI selection and determining the association with the individual covariate may be correlated, invalidating the analysis.

Even if orthogonality of regressors is an uncontroversial requirement to obtain unbiased ROI selection, there are three reasons to consider the properties of this procedure on a formal basis. The first is that, while Vul et al. (2009) named the resulting selection ‘independent’, they also raised doubts about the inferior guarantees offered in this respect when compared to anatomical ROI selection, mentioning an appropriate use of functional ROI selection as the customization of anatomical ROIs to counter the effect of individual variation in cortical anatomy (Vul et al. 2009, pp. 282-283; Saxe et al. 2006). With anatomical ROIs, as with data-splitting schemes or resampling, unbiased ROI selection is ensured by the fact that the selection does not take place based on the same data on which the effect of the individual differences is tested. A formal analysis may help to clarify if these doubts are justified. The second is the question raised by Kriegeskorte et al. (2009) on the role of generalized least squares (GLS) in the context of tests of within-subjects effects. In general, orthogonal regressors may not give rise to uncorrelated effect estimates in a GLS fit, because observations are weighted. Because in neuroimaging applications GLS is routinely used at the first level, one may wonder if it has an impact on ROI selection for testing the interaction with between-subjects effects. Finally, it is of interest to clarify if, beyond those just mentioned,

there is any situation in which functional ROI selection on the basis of task may result in selection bias.

Here, it is shown that ROI selection on the basis of task leads to unbiased ROI selection in most cases, although exceptions exist. Unbiasedness results from using explicitly or implicitly centered individual covariates before entering them in the analysis. Unbiased ROI selection is therefore obtained in this type of study without resorting to inefficient data-splitting schemes or anatomical ROIs. In the following, we investigate conditions for lack of bias in ROI selection in a progression of situations in which assumptions on the data are progressively relaxed. The conditions in which ROI selection bias may occur will be identified, and its impact on a study of typical size investigated with simulations.

Materials and methods

For the simulations of Figure 2 and 3, in each trial normally distributed random data with zero mean and unit variance were created for $p = 20$ subjects, with $q = 100$ data acquisitions in each (representing acquired volumes), each acquisition consisting of $k = 20\,000$ independent datasets (representing voxels). Voxel-by-voxel estimates of effects of the within-subjects covariate *A* were conducted at the first level in each subject separately using the 100 volumes per subject (hence, on a $q \times k$ dataset), and brought forth to the second level. This gave 20 volumes of 'contrast images' at the second level, one per subject (a $p \times k$ dataset), which were fitted to a model with intercept and the between-subjects covariate *B*. Tests on the intercept defined the ROI as the voxels for which the significance level was $P = 0.01$ or less, and separate histograms of the estimated effect of the between-subjects covariate *B* were drawn for the voxels within and outside the ROI. Because the null hypothesis is true in the simulated data, between-subjects effects are realizations of the same zero-centered distribution in all voxels. Hence, the histograms should show that the distribution of between-subjects effects is centered about zero both within and outside the ROI if ROI selection is unbiased. The within- and between- subjects covariates *A* and *B* were random vectors with elements drawn from the

uniform distribution in the unit interval [0, 1]. In all simulations, data at the first level had a heteroscedastic distribution to simulate misspecification of the variance structure (using ordinary least squares where a generalized least squares fit would have been more appropriate). Heteroscedasticity at the first level was obtained by multiplying the data in each subject voxel-by-voxel by $A-\bar{A}+1$, where $\bar{A}$ denotes the average of the elements of *A*, giving a coefficient in the unit interval centered about 1. In the simulations of Figure 3, heteroscedasticity affected also the second level, and was introduced by additionally multiplying all data voxel-by-voxel by $G-\bar{G}+1$, where *G* was a uniform random variable in the interval [0, 1], for the case with no association with *B*, or by $B-\bar{B}+1$, for the case where heteroscedasticity was associated with *B* (see Results for rationale). Each simulation consisted of data pooled from 10 independent repetitions of these trials.

For the simulations of Figure 4, random data were generated with a normal random field distribution so as to simulate analysis on a typical volume. Each volume was a cube of size $59 \times 59 \times 59$, smoothed with a kernel of full-width half size (FWHM) 4 voxels. To avoid edge effects, volumes images were padded at the sides with random variates for 3 times the FWHM size of the kernel prior to smoothing. The volumes thus obtained approximate a full brain analysis on a typical normalized volume in a box of size $91 \times 109 \times 91$, resampled at voxel sizes $2 \times 2 \times 2$ mm. (over 200,000 voxels), smoothed with kernel FWHM 8 mm. ROI were selected as those belonging to the largest cluster defined by the threshold $P = 0.001$, uncorrected, at the second-level test on the intercept. The simulations were otherwise conducted as in Figure 2 and 3; different degrees of heteroscedasticity at the second level were obtained by multiplying the data by $\gamma(B-\bar{B})+1$, where γ was given the values 1 (as in Figure 2 and 3), 0.5, 0.1, and 0.05.

All simulations were carried out in MATLAB R2006b (The Mathworks, Natick, MA) installed on a machine equipped with a 64-bit Athlon processor (Advanced Micro Devices,

Sunnyvale, CA) running Windows XP (Microsoft, Redmond, WA). For the generation of random numbers, the 'MATLAB5 generator' was used.

Results

Case 1: Balanced models with independent errors

Suppose we have acquired images under rest and while performing a demanding cognitive task (within-subjects factor *A*), and we are interested in the differences in activation between males and females (the between-subjects factor *B*, see Figure 1). In the individual differences study, the two relevant factors are the fixed within-subjects factor *A* cognitive task (to define the ROI), and its interaction *AB* with the between-subjects factor sex (to test if the difference in activation due to task differs in males and females). Note that it is the interaction *AB*, not the between-subjects factor *B*, that is normally of interest. In the two-step estimation procedure (Penny and Holmes 2007), one would first estimate the effect of *A*, or a predefined contrast on the levels of *A*, in each subject separately. These estimates would then be brought to the second level to be regressed on *B*. At the second level, testing for a nonzero intercept effect tests for an effect of *A* with subjects as a random factor. Testing the effect of *B* at the second level tests the interaction *AB*.

If the data are balanced in all covariates (same number of observations per level of the covariate), and one may assume that the data are independent and identically distributed except for the correlation arising in the data acquired from the same subject, then the model corresponds to a balanced split-plot ANOVA design with replications in the innermost cells. In this case, there is no need to split the data to obtain independence between effect estimates: balanced ANOVAs are known for decomposing their sums of squares into orthogonal sets.

Case 2: Balanced i.i.d. subjects

In most practical cases, however, one will be interested in the requirements for the orthogonality of *A* and *AB* in a more general model, where for example the within-subjects

covariate *A* is obtained by convolving a series of events or blocks with a canonical haemodynamic response function before forming a contrast, and the between-subjects covariate *B* is a quantitative measure such as a depression score. As a result, two assumptions no longer hold: that the data are balanced in either *A*, *B*, or *AB*, and that the acquisitions at the first level are independent. It will still be assumed that data acquired from different subjects are independent, and that the variance-covariance of the data at the first level is the same in all subjects (i.i.d subjects; this assumption is relaxed in Case 3 below).

Two very general requirements for unbiased ROI selection are that the data must be balanced in the random factor (same number of acquisitions in each subject), and that the same number of scans be acquired for any level of *A* in all subjects. Both requirements are almost invariably satisfied in neuroimaging studies. Together with the i.i.d. subjects assumption, these requirements are also necessary to implement the standard two-step estimation strategy of first- and second-level analysis.

The third crucial requirement to obtain unbiased ROI selection is that *B* be centered. This makes the covariates *A* and *AB* orthogonal (see Appendix A). Kriegeskorte et al. (2009) point out that random components at the first level may introduce dependencies in ROI selection even if the covariates are orthogonal. In the specific case of the individual differences study, it may be shown that after centering *B* this is not so. The covariance structure at the first level has no impact on the bias of ROI selection, irrespective of the estimator used (ordinary least squares, OLS, or generalized least squares, GLS). The algebraic proof is in the Appendix B, and a demonstration through the use of simulations like those of Kriegeskorte et al. (2009) is in Figure 2. The idea of these simulations is that, if selection of a ROI by the within-subjects covariate *A* does not bias estimates of the effect of *AB*, then the *AB* estimates must have the same distribution in the data within and outside the ROI chosen on *A*. Figure 1 shows that, after centering, the distribution of the *AB* interaction estimates is the same within and outside the ROI.

INSERT FIGURE 2 ABOUT HERE

The procedure typically adopted in common neuroimaging practice, however, is slightly different from the one analyzed to this point, and somewhat more complicated. Remember that at the second level, the effects of *A* and of the interaction *AB* of the whole model are respectively given by the intercept and by *B*. Usually, an experimenter would first carry out a one-sample *t* test for the effect of *A* to select the ROI. Hence, the ROI would be selected from a model at the second level containing only the intercept. Then, the average ROI signal would be regressed on *B* in a model including the intercept, perhaps after centering *B*. The model to select the ROI and the model to regress on *B* differ. The simulation of Figure 2 and the algebraic treatment in the Appendix, in contrast, considered a unified model in which both *B* and the intercept were simultaneously included for both ROI selection and regression on the between-subjects covariate.

To understand what happens when the one-sample *t* test plus regression procedure is followed, it is useful to remember that centering a covariate makes it orthogonal to the intercept, and consider the rules that specify what happens when one fits two covariates in the original or orthogonalized form, as for example listed in Draper and Smith (1998), 3rd ed., p. 436. Let X_1 and X_2 be two covariates, and $X_{2|1}$ be X_2 orthogonalized with respect of X_1. Then the following rules apply:

Rule 1: the estimate of the effect of X_1 obtained by fitting a model containing X_1 and $X_{2|1}$ is identical to the estimate obtained by fitting a model with only X_1.

Rule 2: the estimate of the effect of X_2 obtained by fitting a model containing X_1 and X_2 is identical to the estimate obtained by fitting a model with $X_{2|1}$.

Here, X_1 is the intercept, X_2 the non-centered *B*, and $X_{2|1}$ the centered *B*. When one selects the ROI using a one-sample *t* test, Rule 1 applies, meaning that one is obtaining the same estimates of the intercept as those obtained by fitting a model with a centered *B* (the significance thresholds will differ slightly since the residuals are obtained without *B*). When

one regresses on *B*, then Rule 2 applies, meaning that then it does not matter whether one centers *B* or not. Of course, if one does not centre *B*, the fit differs in the estimate of the intercept; however, one would no longer look at the intercept in the regression after having previously conducted a one-sample *t* test. In summary, the one-sample *t* test plus regression procedure is always equivalent to fitting the simultaneous model with centered *B* of the previous section, irrespective of whether one centers *B* or not, up to the degrees of freedom of the one-sample *t* test.

Case 3: Balanced subjects, not identically distributed

The results about Case 2 imply that, for a correlation to arise and bias ROI selection, assumptions on the data must be further relaxed. Here, the case is considered where the data acquired between subjects are no longer identically distributed. An example is when the magnitude of the variance of the residuals varies from subject to subject. The covariance structure is the same, but the magnitude of the associated variance component differs. More generally, heteroscedasticity between subjects may affect any variance component, but involvement of the residuals might be the case of most practical relevance.

Here, it is important to distinguish between two cases, depending on whether the differences in the random variance component are systematically correlated with the between-subjects covariate *B* or not. In the example of sex differences in the activation provoked by the cognitive task, one has a correlation of the residual variance if the first-level residual variance is systematically larger in one sex than in the other.

If there is no systematic association with the values of the between-subjects covariate *B*, there will be on average no ROI selection bias after centering *B* (Figure 3, top row). Note that here unbiasedness is in the expectation (i.e. on average across studies; see Appendix C for a formal treatment). In individual studies, chance associations may bias ROI selection, especially in studies with few participants. This differs from Case 2, because if subjects are identically distributed, *A* and *AB* are exactly orthogonal.

In contrast, if there is a systematic association between the variance structure and the between-subjects covariate *B*, ROI selection is biased, even after centering *B* (Figure 3, bottom row). One can see, however, that centering *B* still has a beneficial effect, since after centering only the source of bias due to heteroscedasticity is still present.

INSERT FIGURE 3 ABOUT HERE

To assess the practical importance of ROI selection bias heteroscedasticity, simulations were conducted on more realistic data, approximating the statistical analysis on a random field of the size of a normalized brain, selecting the ROI as the largest cluster defined by the significance level $P = 0.001$, uncorrected, as one may expect to find in a typical study. These simulations, shown in Figure 4, demonstrate that substantial degrees of correlation between first-level residual variance and the second-level covariate *B* are necessary to induce significant amounts of ROI selection bias. The data in the extreme left column, obtained with the same parameters as in Figure 3, reveal that extremely high correlations between heteroscedasticity and the covariate *B* were used to yield the robust effects reported previously. With systematic association levels below 0.4, hardly any bias can be detected in ROI selection.

INSERT FIGURE 4 ABOUT HERE

Discussion

It has been shown here that centering the between-subjects covariate leads to unbiased ROI selection in individual differences studies irrespective of the covariance structure at the first level and how it is accounted for at first-level estimation. Furthermore, the common procedure of selecting the ROI with a one-sample *t* test at the second level guarantees unbiased ROI selection by centering the between-subjects covariate implicitly. In contrast, biasedness in ROI selection may be occasioned by severe violations of the assumptions of identical distribution of data between subjects. In the simulations presented here, however, serious ROI selection bias was the consequence of very high correlations between heteroscedastic variance

and predictor variables; one wonders if such high correlations are often encountered in practice.

It is worth noting in this respect that biased ROI selection in such extreme cases is not the only problem of the model, which may be viewed as misspecified on other grounds. Because the assumptions for the multi-step estimation strategy are not satisfied, inferences at the second level may be invalid irrespective of whether or how a ROI was selected, and the ensuing problems may be of a more fundamental nature than those of ROI selection.

Centering a covariate causes the effect estimates to be parametrized as differences from the group mean. This is the same parametrization commonly used in ANOVA, and has the effect of inducing a lack of correlation in the sampling variance of the effects. The effect of centering are not limited to the between-subjects variable *B*: if the data at the first level are independent, it may be shown that centering *A* causes lack of correlation in the sampling variances of *A* and *B*, and centering both *A* and *B* has the same consequence simultaneously on *A*, *B*, and *AB*.

References

Barrett, L.F., 2009. Understanding the mind by measuring the brain. Lessons from measuring behaviour (Commentary on Vul et al., 2009). Perspectives Psychol. Sci. 4, 314-318.

Diggle, P.J., Heagerty, P., Liang, K.Y., Zeger, S.L., 2002. Analysis of Longitudinal Data (2nd ed.). Oxford University Press, Oxford.

Draper, N.R., Smith, H., 1998. Applied Regression Analysis. John Wiley & Sons (3rd ed.), New York.

Friston, K.J., Rothstein, P., Geng, J.J., Sterzer, P., Henson, R.N., 2006. A critique of functional localizers. NeuroImage 30, 1077-1087.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: The dangers of double dipping. Nat. Neurosci. 12, 535-540.

Lazar, N.A., 2009. Discussion of "Puzzingly high correlations in fMRI studies of emotion, personality, and social cognition" by Vul et al. (2009). Perspectives Psychol. Sci. 4, 308-309.

Lieberman, M.D., Berkman, E.T., Wager, T.D., 2009. Correlations in social neuroscience aren't voodoo. Commentary on Vul et al. (2009). Perspectives Psychol. Sci. 4, 299-307.

Lindquist, M.A., Gelman, A., 2009. Correlations and multiple comparisons in functional imaging. A statistical perspective (commentary on Vul et al. 2009). Perspectives Psychol. Sci. 4, 310-313.

McCulloch, C.E., Searle, S.R., 2001. Generalized, Linear, and Mixed Models. Wiley, New York.

Nichols, T.E., Poline, J.B., 2009. Commentary on Vul et al.'s (2009) "Puzzingly high correlations in fMRI studies of emotion, personality, and social cognition". Perspectives Psychol. Sci. 4, 291-293.

Penny, W.D., Holmes, A.J., 2007. Random effects analysis. In: Friston, K.J., Ashburner, J.T., Kiebel, S.J., Nichols, T.E., Penny, W.D. (eds). Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press, London.

Saxe, R., Brett, M., Kanwisher, N., 2006. Divide and conquer: A defense of functional localizers. NeuroImage 30, 1088-1096.

Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. Perspectives Psychol. Sci. 4, 274-290.

Yarkoni, T., 2009. Big correlations in little studies. Inflated fMRI correlations reflect low statistical power. Commentary on Vul et al. (2009). Perspectives Psychol. Sci. 4, 294-298.

Appendices

Appendix A. Orthogonality of A and AB

In a sample of *p* subjects with *q* acquisitions obtained in each subject, let $\mathbf{1}_p \otimes \mathbf{a}_q$ be a within-subjects covariate *A*, and $\mathbf{b}_p \otimes \mathbf{1}_q$ a between-subjects covariate *B*, where the subscript indicates the size of the column vector, and the symbol '$\otimes$' the Kronecker product. The first level covariate $\mathbf{a}_q$ may originate from the reparametrization of the model so as to encode a contrast within the factor *A*. If the between-subjects covariate *B* is centered, then $\mathbf{b}'_p\mathbf{1}_p = 0$. Then, also the inner product between the *AB* interaction $\mathbf{b}_p \otimes \mathbf{a}_q$ and the within-subjects covariate $\mathbf{1}_p \otimes \mathbf{a}_q$ is zero:

$$
\begin{aligned}
&(\mathbf{b}_p \otimes \mathbf{a}_q)'(\mathbf{1}_p \otimes \mathbf{a}_q) \\
&= (\mathbf{b}'_p \otimes \mathbf{a}'_q)(\mathbf{1}_p \otimes \mathbf{a}_q) \\
&= \mathbf{b}'_p\mathbf{1}_p \otimes \mathbf{a}'_q\mathbf{a}_q
\end{aligned}
$$

But since $\mathbf{b}'_p\mathbf{1}_p = 0$, the whole expression vanishes.

Appendix B. Balanced i.i.d subjects

The general specification of a linear mixed model is

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \varepsilon$$

where the design matrix $\mathbf{X}$ specifies the covariates of the fixed effects β, $\mathbf{Z}$ is an incidence matrix for the random effects $\mathbf{u}$, ε the error term, and $\mathrm{Var}(\mathbf{y}) = \mathbf{V}$.

Assume for the moment that the dispersion matrix $\mathbf{V}$ is known and used in a GLS estimation, so that the sampling variance of the parameter estimates is $\mathrm{Var}(\hat{\beta}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$. By assuming the data to be independent between subjects with an identical random structure at the first level, one has

$$\mathbf{V} = \mathbf{I}_p \otimes \mathbf{V}_0,$$

i.e., $\mathbf{V}$ is block-diagonal with blocks given by the first-level dispersion matrix $\mathbf{V}_0$. Hence, the sampling variance is

$$\mathrm{Var}(\hat{\beta}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} = (\mathbf{X}'(\mathbf{I}_p \otimes \mathbf{V}_0^{-1})\mathbf{X})^{-1}. \tag{1}$$

Let $\mathbf{A}_q$ be the full-rank matrix of all first-level covariates in the appropriate parametrization (in the previous section, $\mathbf{a}_q$ was a column of this matrix; we now use the symbol *A* for all these covariates). Then the columns of $\mathbf{X}$ are

$\mathbf{b}_p \otimes \mathbf{1}_q$	the between-subjects covariate *B*,
$\mathbf{b}_p \otimes \mathbf{A}_q$	its interaction with the within subjects-covariates *AB*,
$\mathbf{1}_p \otimes \mathbf{A}_q$	the within-subjects covariates *A*,
$\mathbf{1}_p \otimes \mathbf{1}_q$	the intercept.

If $\mathbf{b}_p'\mathbf{1}_p = 0$, it may be shown that the first two entries *B* and *AB* in this list are orthogonal to both last entries *A* and intercept in the outer product matrix $\mathbf{X}'(\mathbf{I}_p \otimes \mathbf{V}_0^{-1})\mathbf{X}$:

$$\begin{aligned}
B \perp A &\qquad (\mathbf{b}_p \otimes \mathbf{1}_q)'(\mathbf{I}_p \otimes \mathbf{V}_0^{-1})(\mathbf{1}_p \otimes \mathbf{A}_q) = \mathbf{b}_p'\mathbf{1}_p \otimes \mathbf{1}_q'\mathbf{V}_0^{-1}\mathbf{A}_q = 0 \\
B \perp \text{intercept} &\qquad (\mathbf{b}_p \otimes \mathbf{1}_q)'(\mathbf{I}_p \otimes \mathbf{V}_0^{-1})(\mathbf{1}_p \otimes \mathbf{1}_q) = \mathbf{b}_p'\mathbf{1}_p \otimes \mathbf{1}_q'\mathbf{V}_0^{-1}\mathbf{1}_q = 0 \\
AB \perp A &\qquad (\mathbf{b}_p \otimes \mathbf{A}_q)'(\mathbf{I}_p \otimes \mathbf{V}_0^{-1})(\mathbf{1}_p \otimes \mathbf{A}_q) = \mathbf{b}_p'\mathbf{1}_p \otimes \mathbf{A}_q'\mathbf{V}_0^{-1}\mathbf{A}_q = 0 \\
AB \perp \text{intercept} &\qquad (\mathbf{b}_p \otimes \mathbf{A}_q)'(\mathbf{I}_p \otimes \mathbf{V}_0^{-1})(\mathbf{1}_p \otimes \mathbf{1}_q) = \mathbf{b}_p'\mathbf{1}_p \otimes \mathbf{A}_q'\mathbf{V}_0^{-1}\mathbf{1}_q = 0
\end{aligned} \tag{2}$$

The matrix $\mathbf{X}'(\mathbf{I}_p \otimes \mathbf{V}_0^{-1})\mathbf{X}$ is therefore block-diagonal with *B*, *AB*, and their product in one block, and *A*, the intercept, and their product in the other. After inversion, the off-block diagonal terms remain zero, including the covariance between *AB* and *A*.

In maximum-likelihood estimation, the unknown dispersion matrix $\mathbf{V}$ is replaced by its estimate $\hat{\mathbf{V}}$. McCulloch and Searle (2001) summarize studies on the sampling variance for this case, which is given by an expression of the form

$$\mathrm{Var}(\hat{\beta}_{\mathrm{ML}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} + \mathbf{T}, \tag{3}$$

with $\mathbf{T}$ a complex function of $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ (see pp. 164-166). The discussion of the previous section may apply also to this case.

Misspecifying the dispersion matrix $\mathbf{V}_0$ at the first level, for example by using OLS, leads to the same conclusion. If a "working matrix" $\mathbf{W}$ is used in place of the unknown $\mathbf{V}^{-1}$ (Diggle et al. 2002), the sampling variance is given by

$$\mathrm{Var}(\tilde{\beta}_W) = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{V}\mathbf{W}\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}. \quad (4)$$

After replacing in this expression $\mathbf{I}_p \otimes \mathbf{V}_0$ for $\mathbf{V}$, and $\mathbf{I}_p \otimes \mathbf{W}_0$ for the misspecified matrix $\mathbf{W}_0$ at the first level, algebraic manipulations analogous to those just shown give a block diagonal structure for both $\mathbf{X}'\mathbf{W}\mathbf{X}$ and $\mathbf{X}'\mathbf{W}\mathbf{V}\mathbf{W}\mathbf{X}$, with zero inner product between *AB* and *A*. For the OLS estimator at the first level in which $\mathbf{W}_0 = \mathbf{I}_q$,

$$\mathrm{Var}(\tilde{\beta}_I) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \quad (5)$$

with block diagonal structure for $\mathbf{X}'\mathbf{V}\mathbf{X}$. In summary, it does not matter how first-level estimates are obtained with respect of the random effects; selection of ROI remains unbiased in the i.i.d. subjects case.

Appendix C, Balanced subjects, not identically distributed

It is shown here that deviations from the independent identical distribution case affecting the data acquired between subjects may bias ROI selection. Suppose, for example, that the model is fitted assuming identically distributed data between subjects (by using OLS at the second level), so that the working matrix is $\mathbf{W} = \mathbf{I}_p \otimes \mathbf{W}_0$. In reality one variance component differs from subject to subject. To model this eventuality, write

$$\mathbf{V} = \mathbf{I}_p \otimes \mathbf{V}_q + \mathbf{\Lambda}_p \otimes \mathbf{R}_\lambda$$

where $\mathbf{I}_p \otimes \mathbf{V}_q$ absorbes all variance components that are i.i.d. between subjects, with the heteroscedastic $\mathbf{\Lambda}_p \otimes \mathbf{R}_\lambda$ being the exception. $\mathbf{R}_\lambda$ is the correlation structure of the residuals, parametrized by the diagonal matrix $\mathbf{\Lambda}_p$ containing heteroscedasticity parameters at subject

level (in the simulations of the main text, where heteroscedasticity affected residual variance, $\mathbf{R}_\lambda = \mathbf{I}_q$). Replacing $\mathbf{V}$ in equation (4), the sampling variance decomposes into a sum of two terms:

$$\mathrm{Var}(\tilde{\beta}_W) = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}(\mathbf{I}_p \otimes \mathbf{V}_p)\mathbf{W}\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}(\mathbf{\Lambda}_p \otimes \mathbf{R}_\lambda)\mathbf{W}\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}.$$

The first term is the sampling variance due to identically distributed components between subjects, which was shown in the previous section to give zero covariance between *A* and *AB*. In the second term, $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} = (\mathbf{X}'(\mathbf{I}_p \otimes \mathbf{W}_0)\mathbf{X})^{-1}$ has the usual block diagonal structure. The contribution of heteroscedasticity to the covariance is therefore given by $\mathbf{X}'\mathbf{W}(\mathbf{\Lambda}_p \otimes \mathbf{R}_\lambda)\mathbf{W}\mathbf{X}$. If one computes the analogous entries in the list of products of cross terms (2) for this covariance, one finds that block diagonal structure is maintained if the weighted inner product term $\mathbf{b}_p'\mathbf{\Lambda}_p\mathbf{1}_p$ is zero. For example, after some algebraic manipulations the entry for the covariance between *A* and *AB* is

$$\mathbf{b}_p'\mathbf{\Lambda}_p\mathbf{1}_p \otimes \mathbf{A}_q'\mathbf{W}_0\mathbf{R}_\lambda\mathbf{W}_0\mathbf{A}_q, \qquad (6)$$

while the other entries are all Kronecker products between $\mathbf{b}_p'\mathbf{\Lambda}_p\mathbf{1}_p$ and a second term. If the variance of the data in the subjects, i.e. the diagonal terms of $\mathbf{\Lambda}_p$, are uncorrelated with the values of $\mathbf{b}_p$, then the weights in the inner product cancel each other on average, and for $\mathbf{b}_p'\mathbf{1}_p = 0$ the expectation of these products is zero. Hence, in the absence of a systematic association, bias can occur only if chance differences in variance are aligned with the between-subjects covariate $\mathbf{b}_p$. This becomes increasingly unlikely with increasing number of subjects.

It is also possible for the term $\mathbf{A}_q'\mathbf{W}_0\mathbf{R}_\lambda\mathbf{W}_0\mathbf{A}_q$ to be zero. Suppose that there is correlation between subjects given by $\mathbf{\Lambda}_p$, but data are i.i.d. at the first level, and OLS is used; then $\mathbf{W}_0 = \mathbf{I}_q$, $\mathbf{R}_\lambda = \mathbf{1}_q\mathbf{1}_q'$, and if the columns of $\mathbf{A}$ are centered, $\mathbf{A}'\mathbf{1}_q\mathbf{1}_q'\mathbf{A} = \mathbf{0}$.

Figure captions

Figure 1. Schematic diagram of an individual differences study. The within-subjects variable *A* codes experimental conditions (used for ROI selection), as in the figure, or expected BOLD signal at blocks or event types. The between-subjects variable *B* codes for individual characteristics (used for ROI analysis).

Figure 2. Estimates of the interaction effect *AB* inside (gray, red in the online version) and outside (black, blue in the online version) the ROI selected by an effect of *A* significant at the $P = 0.01$ level in random data. On the left, the *B* covariate was not centered, causing bias in ROI selection visible as *AB* effects being positive even if in reality here the null hypothesis holds. On the right, no bias is ROI selection after centering *B*. In these data, large variance differences correlated with *A* were introduced in the first-level data, and the effect of *A* was estimated using ordinary least squares, showing no impact on ROI selection bias.

Figure 3. The same simulation as in the previous figure, with the addition of large variance differences across subjects. In the top row, these differences were introduced randomly. As a result, ROI selection is biased only if *B* is not centered (left) but remains unbiased if *B* is centered (right), as in the case where variance is the same across subjects (Figure 2). However, if the differences in variance are correlated with *B* (bottom row), centering *B* will not remove this additional source of bias in ROI selection.

Figure 4. The top row displays correlations carried out at the ROI to test the interaction between *A* and *B* in random data, with progressively lower degrees of heteroscedasticity from left to right. The least squares lines of all ten trials of the simulation are shown to assess variability of bias at fixed heteroscedasticity levels. The middle row shows boxplots of the *t* values obtained in the regression; if ROI selection is unbiased, the boxes should be roughly centered at zero. In the bottom row, diagnostic scatterplots of the correlations between residual standard deviation in the first-level fit in the ROI and the between-subjects covariate *B*. The highest correlation level on the left was used in the simulations of Figure 3 to obtain the robust effects shown there. These plots show that in samples of this size this correlation is capable to alert to the possible ROI selection bias.

Figures

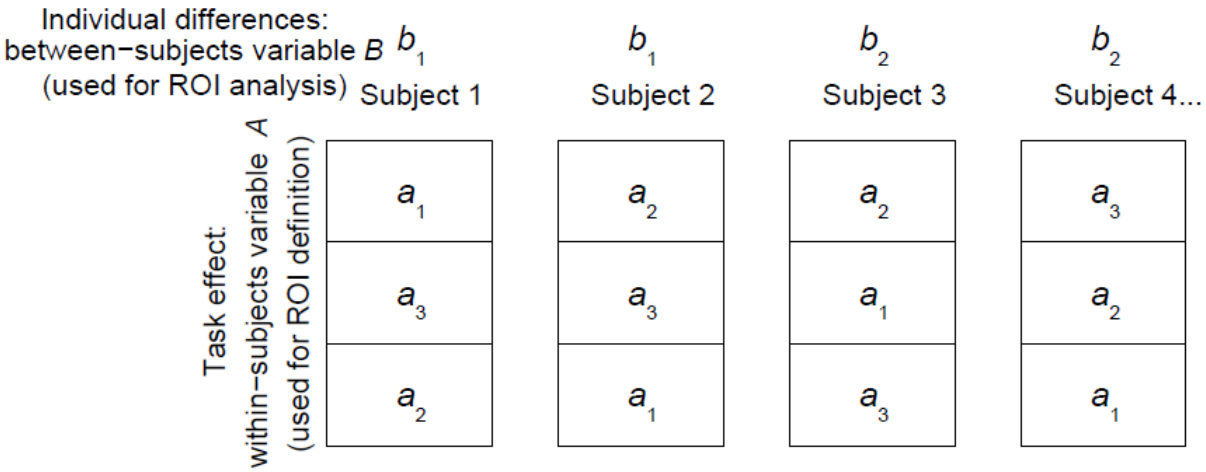

Figure 1

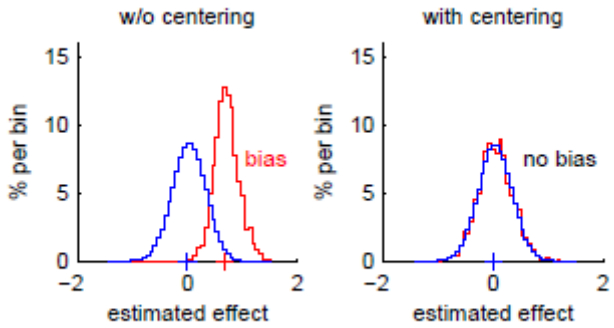
w/o centering
with centering
% per bin
15
10
5
0
bias
no bias
−2
0
2
estimated effect

Figure 2

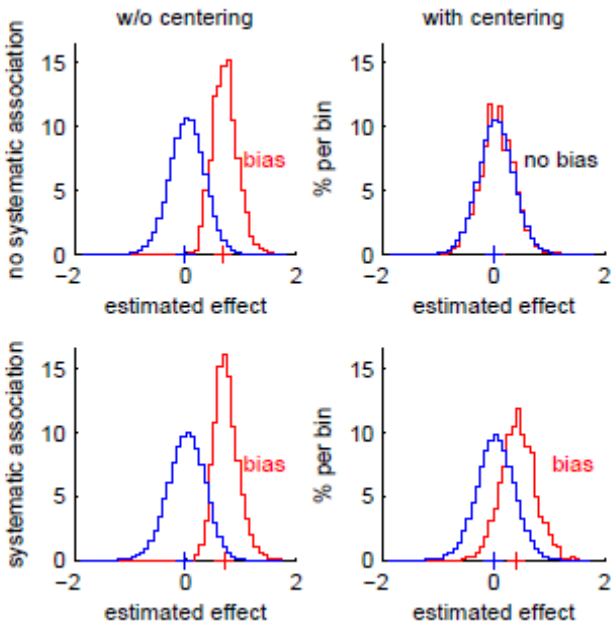
w/o centering
with centering
no systematic association
systematic association
% per bin
bias
no bias
estimated effect
15
10
5
0
−2
0
2

Figure 3

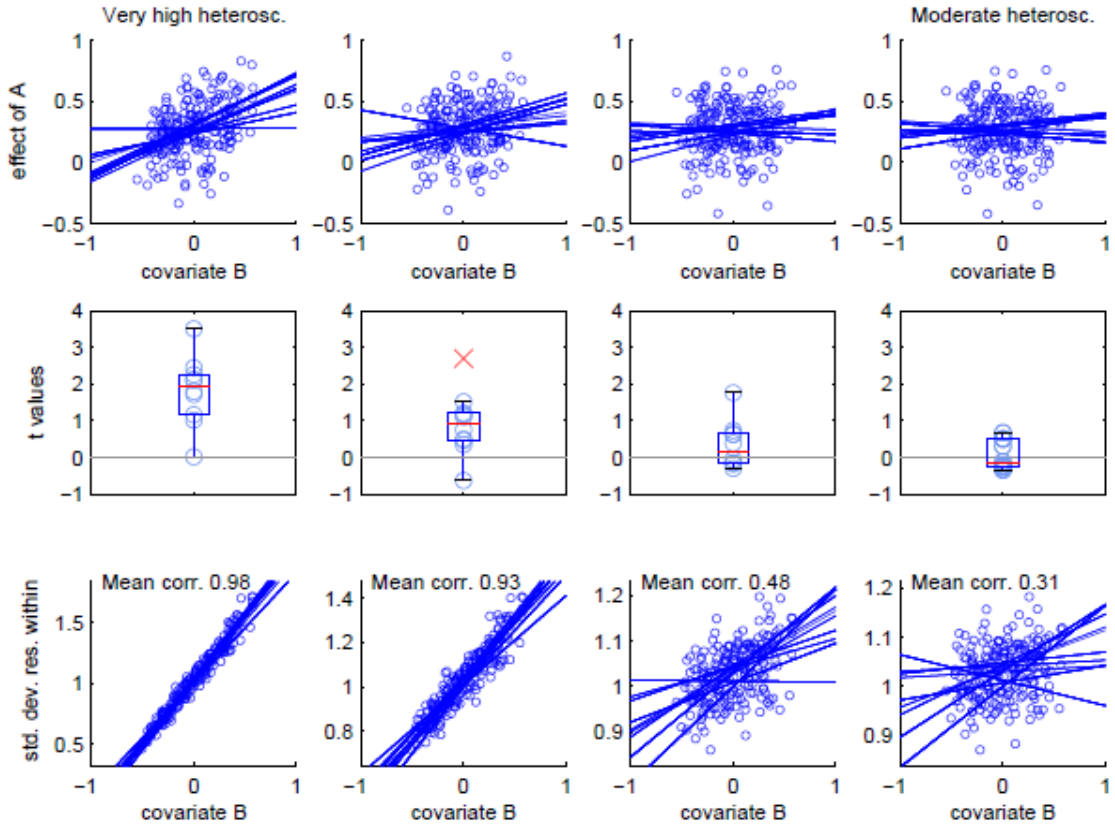
Very high heterosc.
Moderate heterosc.
effect of A
t values
std. dev. res. within
covariate B
Mean corr. 0.98
Mean corr. 0.93
Mean corr. 0.48
Mean corr. 0.31

Figure 4