Corrigendum for:

A purely confirmatory replication study of structural brain-behavior correlations

MC Keuken a, A. Ly a, WE Boekel a, E-J Wagenmakers a, L Belay, J Verhagen, S Brown b, BU Forstmann a

a University of Amsterdam, Amsterdam, The Netherlands
b University of Newcastle, Newcastle, Australia

Introduction

In our previous study, we reported a purely confirmatory replication study of structural brain-behavior correlations[1]. For all but one of the 17 findings under scrutiny, confirmatory Bayesian hypothesis tests indicated evidence in favor of the null hypothesis ranging from anecdotal (Bayes factor < 3) to strong (Bayes factor > 10). In several studies, effect size estimates were substantially lower than in the original studies.
We now discovered a mistake in the post-processing pipeline of our diffusion-weighted imaging (DWI) data analyses originally included in this replication study. This led us to recalculate and correct five of the 17 originally reported brain-behavior correlations that were based on DWI data. In short, after reanalyzing the DWI data correctly, the original conclusions for the five corrected analyses did not change.

More concretely, we discovered that an extra volume was included in the acquisition protocol which was subsequently incorrectly included in the data analyses. This extra volume was incorporated due to the Philips scanner software version R3. This volume is the average of all the acquired diffusion weighted volumes and was placed at the end of the data file. Such an extra volume can be used to calculate Apparent Diffusion Coefficient (ADC) maps. This extra volume has a b-value of 1000 and bvecs values of 0,0,0. As this is not truly a measured direction or a proper B0 volume, this volume should have been removed. The extra volume, as well as the corresponding extra entries in the bval and bvecs were removed. All DWI data processing was redone with the pre-registered parameter settings. Removing this extra volume from the analyses resulted in considerably different structural DWI measures including fractional anisotropy (FA), mean diffusivity (MD), and λ1 values from the pre-defined regions of interest (ROIs). This mistake also affected tractography results including the calculation of tract strength. Therefore, the previously reported results regarding the failed replications of Forstmann et al. (2010) and Xu et al. (2012) needed to be corrected.

After removing the additional volume from the current DWI data set, the analyses pipeline described in the original paper, i.e., section *1.1.1 DWI analyses* and *1.1.2. Probabilistic tractography* were used.

In addition to the mistake in the post-processing pipeline of the DWI data, it came to our attention that the correlation coefficients reported in the text inset of figure 6 were swapped between the two panels. Although this had no influence on the conclusion, we have taken the opportunity to correct this error.

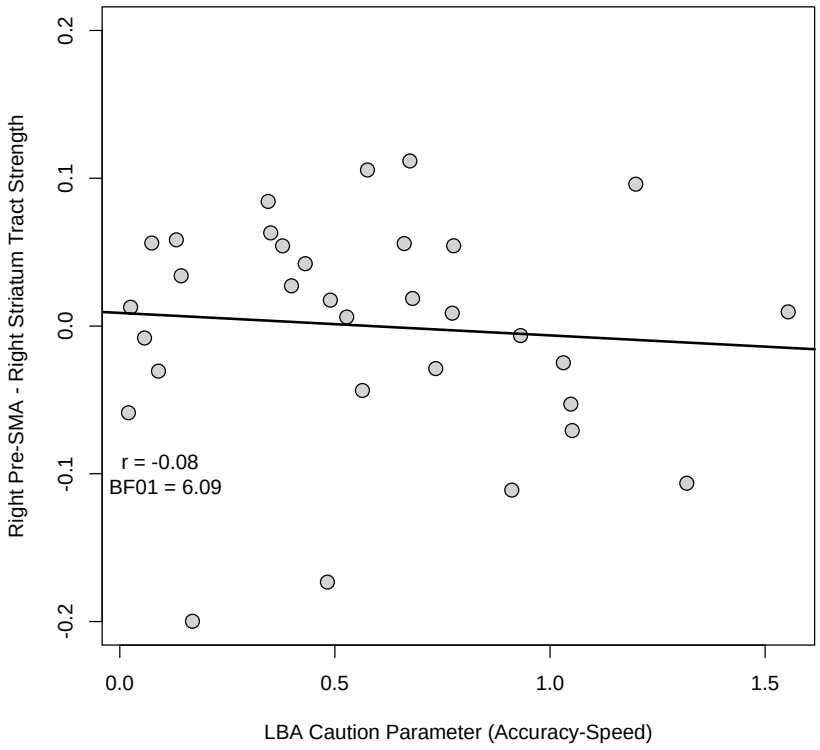
In the following we now present the corrected results, tables, and figures of the studies previously reported[1].

3.1. Replication 1: Forstmann et al. (2010)

Summary statistics One subject was removed from the analyses (>2.5SD from the mean). Below are the corrected new summary statistics for the tract strength measure between the right pre-SMA and right striatum as well as the LBA flexibility measures presented. These are based on 32 subjects and are not corrected for age and gender.

Table 1. Corrected summary statistics.

	Mean	SD	Min-Max
Tract strength	0.85	0.09	0.61, 0.95
LBA Caution parameter	0.58	0.40	0.02, 1.55



Corrected Figure 3. Scatterplot of replication 1: Forstmann et al., 2010. The relationship between LBA caution parameter (quantified by taking the difference in response caution between the accuracy and speed condition) and tract strength between right Pre-SMA and right Striatum, quantified by probabilistic tractography, corrected for age and gender.

Corrected Table 2. Results of the one-sided Bayesian hypothesis test for a positive correlation. We have included the old results to facilitate the comparison with the corrected analyses.

Data Pair					Confirmatory		Exploratory		
ROI	N_{orig}	N_{rep}	r_{orig}	r_{rep}	BF_{01}	Evidence cat.	BF_{0r}	Evidence cat.	p-value
Tract strength and LBA flexibility									
Pre-SMA to Striatum									
Corrected	9	32	.93	-0.08	6.09	Moderate (H_0)	342.09	Extreme (H_0)	0.67
Old	9	31	.93	0.03	3.90	Moderate (H_0)	180.20	Extreme (H_0)	0.43

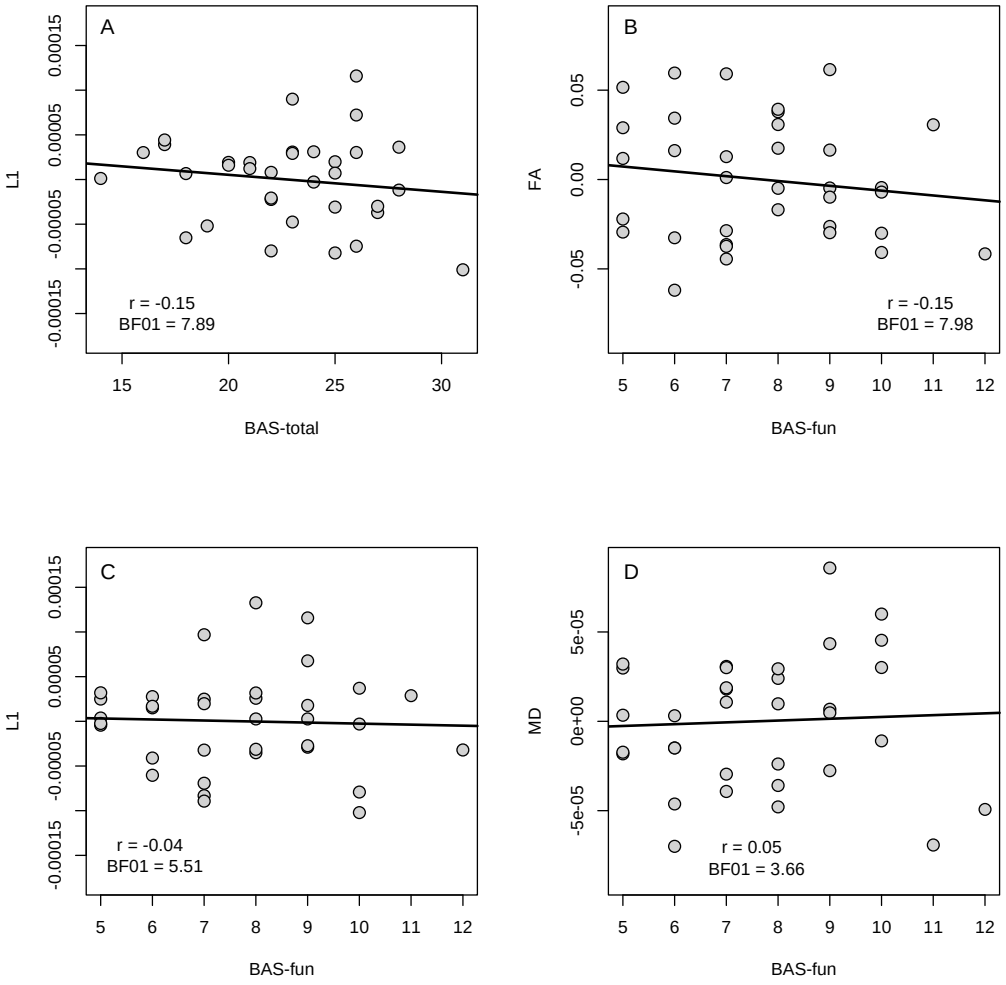
Interim conclusion By removing the extra volume in the DWI data, the correlations between the tract strength measures derived from the pre-SMA and the striatum and the LBA flexibility parameters remain absent (see[1]).

3.3 Replication 3: Xu et al. (2012)

Summary statistics One subject was removed from the λ1 BAS-total analysis as well as for the MD BAS-FUN analysis (>2.5SD from the mean). The corrected summary statistics for the DTIFit analyses are now based on 34 and 35 subjects, respectively and are not corrected for age and gender.

Table 3. Corrected summary statistics.

	Mean	SD	Min-Max
BAS-total	22.76	3.88	14, 31
BAS-fun (λ1)	7.69	1.84	5, 12
BAS-fun (FA)	7.65	1.86	5, 12
BAS-total (λ1)	11.65E4	0.50E4	10.71E4, 12.89E4
BAS-fun (FA region)	0.49	0.04	0.43, 0.55
BAS-fun (λ1 region)	11.70E4	0.55E4	10.69E4, 13.04E4
BAS-fun (MD region)	6.74E4	0.38E4	5.92E4, 7.57E4



Corrected Figure 5. Scatterplots of replication 3: Xu et al., (2012). (A) The relationship between BAS-total and λ1 in left CR and left SLF. (B-D) The relationship between BAS-FUN and (B) FA in left CR and left SLF, (C) λ1 in left CR and left SLF, and (D) MD in left SLF and left IFOF, corrected for age and gender.

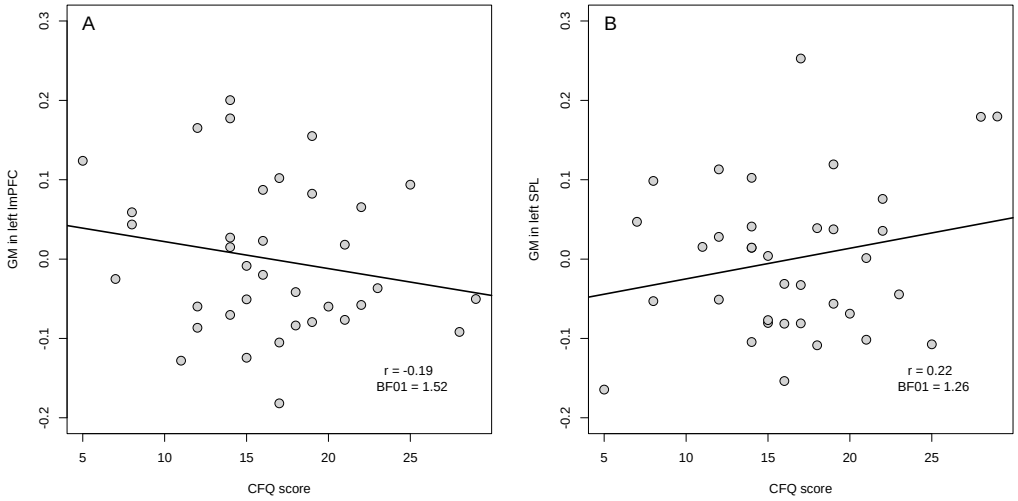
Corrected Table 4. Results of the one-sided Bayesian hypothesis test for a positive correlation. We have included the old results to facilitate the comparison with the corrected analyses.

Data Pair					Confirmatory		Exploratory		
ROI	N_{orig}	N_{rep}	r_{orig}	r_{rep}	BF_{01}	Evidence cat.	BF_{0r}	Evidence cat.	p-value
BAS-Total and λ1									
Left CR and SLF									
Corrected	51	34	.51	-.15	7.89	Moderate (H_0)	86.34	Very strong (H_0)	.80
Old	51	35	.51	-.28	11.74	Strong (H_0)	249.41	Extreme (H_0)	.95
BAS-Fun and FA									
Left CR and SLF									
Corrected	51	35	.52	-.15	7.98	Moderate (H_0)	108.08	Extreme (H_0)	.80
Old	51	36	.52	-.19	9.40	Moderate (H_0)	170.51	Extreme (H_0)	.86

BAS-Fun and λ1									
Left CR and SLF									
Corrected	51	35	.58	-.04	5.51	Moderate (H_0)	110.29	Extreme (H_0)	.59
Old	51	35	.58	-.24	10.57	Strong (H_0)	848.06	Extreme (H_0)	.92
BAS-Fun and MD									
Left SLF and IFOF									
Corrected	51	34	.51	.05	3.66	Moderate (H_0)	12.73	Strong (H_0)	.39
Old	51	36	.51	.15	2.04	Anecdotal (H_0)	4.13	Moderate (H_0)	.19

Interim Conclusion By removing the extra volume in the DWI data, the correlations between the DTIFit parameters and the BAS-FUN scores remain absent (see [1]).

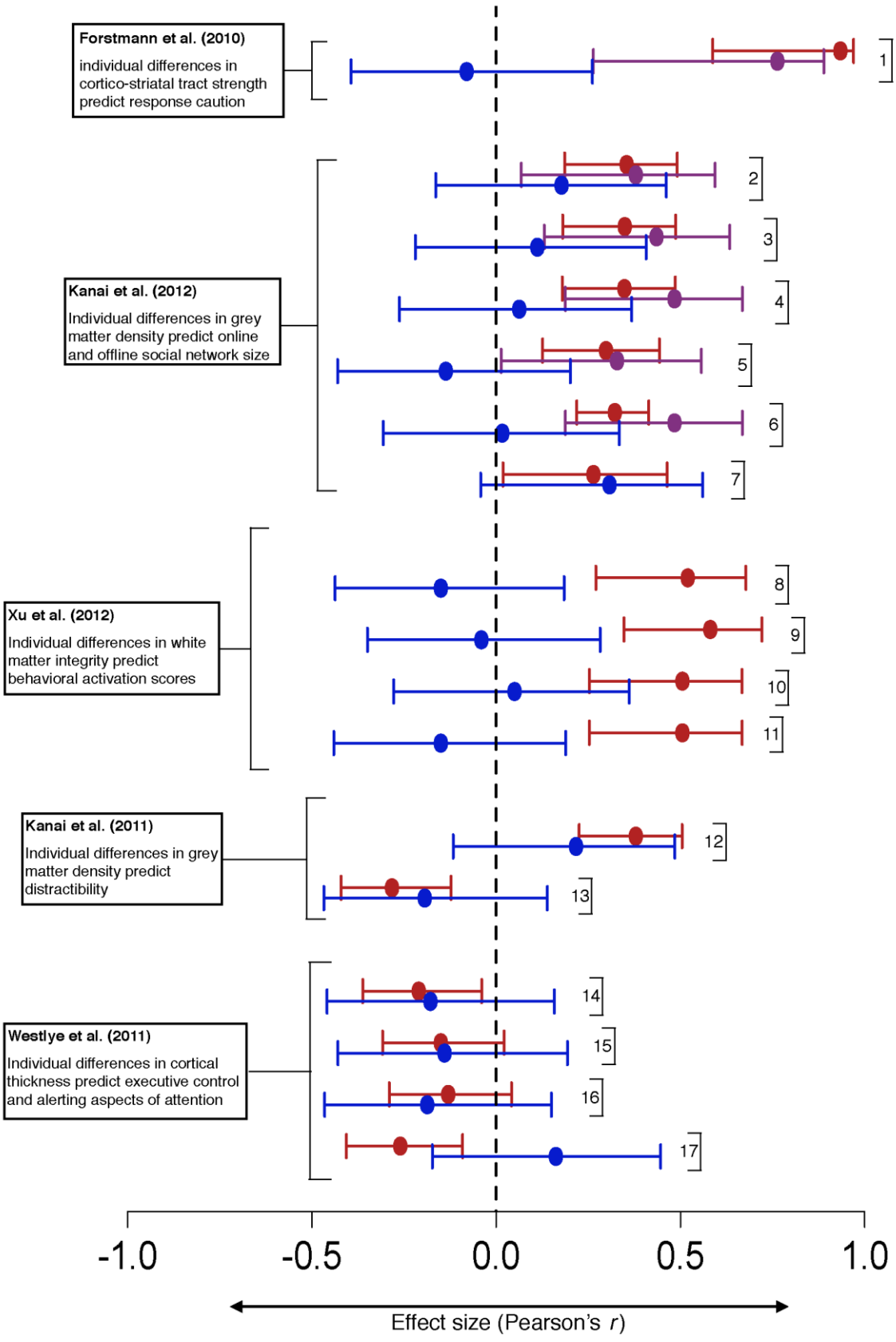
3.4 Replication 4: Kanai et al. (2011)



Corrected Figure 6. Scatterplots of replication 4: Kanai, Dong, et al., (2011). (A) The relationship between CFQ score and GM in (A) left lmPFC and (B) left SPL.

Conclusion

We have reanalyzed five previously reported structural brain-behavior correlations by removing an extra volume in the DWI data. Based on the corrected results, the original conclusions regarding the failure to replicate structural brain-behavior correlations still hold as is shown in the corrected summary figure.


Forstmann et al. (2010)
individual differences in cortico-striatal tract strength predict response caution
Kanai et al. (2012)
Individual differences in grey matter density predict online and offline social network size
Xu et al. (2012)
Individual differences in white matter integrity predict behavioral activation scores
Kanai et al. (2011)
Individual differences in grey matter density predict distractibility
Westlye et al. (2011)
Individual differences in cortical thickness predict executive control and alerting aspects of attention
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
-1.0
-0.5
0.0
0.5
1.0
Effect size (Pearson's r)

Corrected Figure 8. Summary image of our replication results. 95% confidence intervals of posterior probability distributions are shown for the original studies (red), replications within original studies (purple), and the current independent replication attempt (blue). individual effects: (1): The corrected LBA flexibility correlated to tract strength between pre-supplementary motor area and striatum. (2-6): FBN correlated to grey matter volume in (2) left middle temporal gyrus, (3) right superior temporal sulcus, (4) right entorhinal cortex, (5) left amygdala, and (6) right amygdala. (7) SNS correlated to grey matter volume in right amygdala. (8) The corrected BAS-total correlated to λ1 in left CR and SLF. (9) The corrected BAS-FUN correlated to FA in left CR and SLF. (10) The corrected BAS-FUN correlated to λ1 in left CR and SLF. (11) The corrected BAS-FUN correlated to MD in left SLF and IFOF. (12-13) CFQ correlated to grey matter volume in (12) left superior parietal lobe and (13) left middle prefrontal cortex. (14-16) Executive control correlated to cortical thickness in (14) left caudal anterior cingulate cortex, (15) left superior temporal gyrus, and (16) right middle temporal gyrus. (17) Alerting correlated to cortical thickness in left superior parietal lobe.

Acknowledgments

We thank SURFsara (www.surfsara.nl) for the support in using the Lisa Computer Cluster to process the DWI data.

Reference

1. Boekel, W. *et al.* A purely confirmatory replication study of structural brain-behavior correlations. *Cortex* **66,** 115–133 (2015).