Reliability versus Validity in Resting State and Task-based fMRI

Stephen J. Gotts
Laboratory of Brain and Cognition
NIMH/NIH
Bethesda, MD

Acknowledgements:

Section on Cognitive Neuropsychology (NIMH)

Alex Martin, Chief
Adrian Gilmore
Sarah Kalinowski
Shawn Milleville

NMR Center

Vinai Roopchansingh

Section on Functional Imaging Methods

Peter Bandettini, Chief
Javier Gonzalez-Castillo
Dan Handwerker

Functional Connectivity of Spontaneous Activity at Rest (i.e. "Resting State")

Functional Connectivity of Spontaneous Activity at Rest (i.e. "Resting State")

- very popular (easy and fast to administer)
- subjects passively view a fixation cross
- fluctuations in spontaneous activity (< .1 Hz) are correlated throughout the brain in a spatially restricted manner

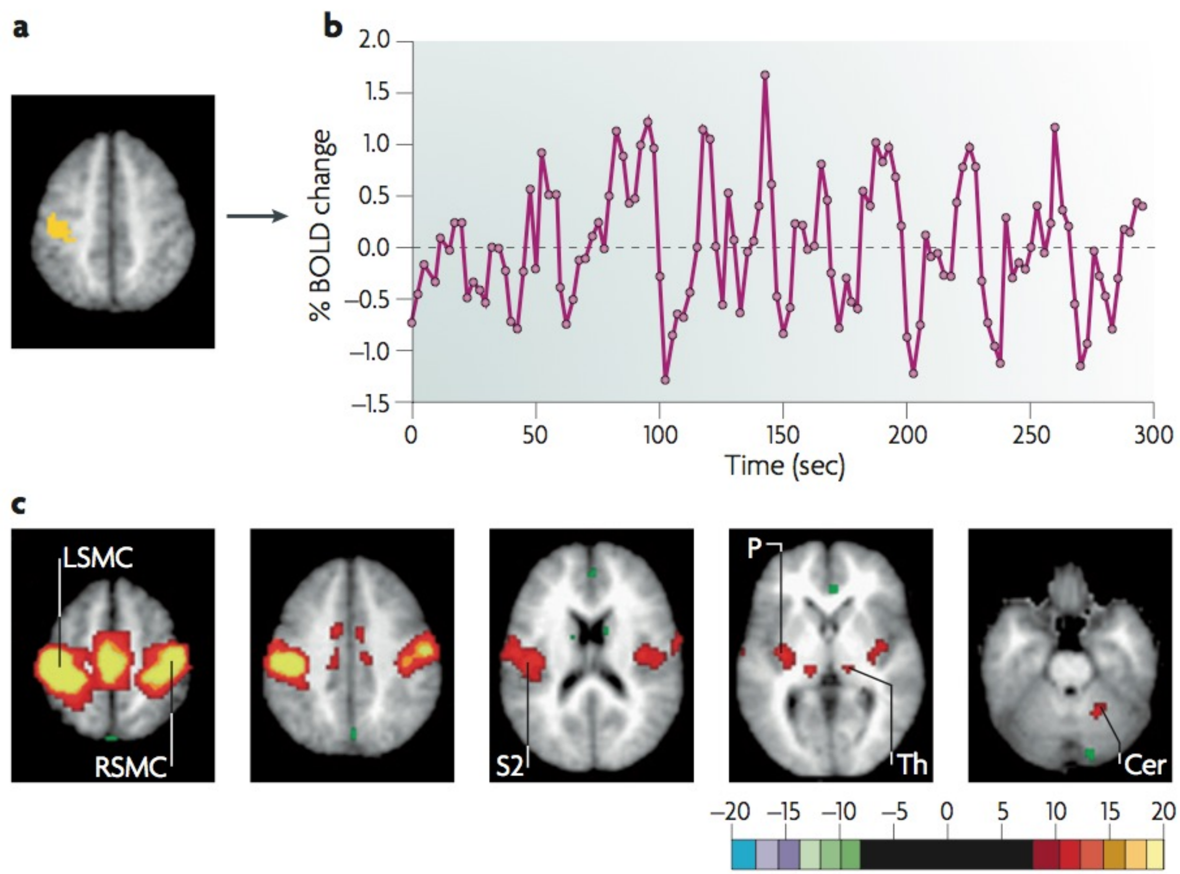
Functional Connectivity of Spontaneous Activity at Rest (i.e. "Resting State")

- very popular (easy and fast to administer)
- subjects passively view a fixation cross
- fluctuations in spontaneous activity (< .1 Hz) are correlated throughout the brain in a spatially restricted manner

For review:

Fox & Raichle (2007). *Nat Rev Neurosci*

Fox & Greicius (2010). *Front Syst Neurosci*



Today's Talk

Today's Talk

- What’s the best way to remove noise in resting-state fMRI?

Today's Talk

- What’s the best way to remove noise in resting-state fMRI?
- Test-retest reliability as a guide?

Today's Talk

- What’s the best way to remove noise in resting-state fMRI?
- Test-retest reliability as a guide?
- Validity vs. test-retest reliability

Today's Talk

- What’s the best way to remove noise in resting-state fMRI?
- Test-retest reliability as a guide?
- Validity vs. test-retest reliability
    - our recent efforts to evaluate these issues experimentally

Why are artifacts so problematic ?

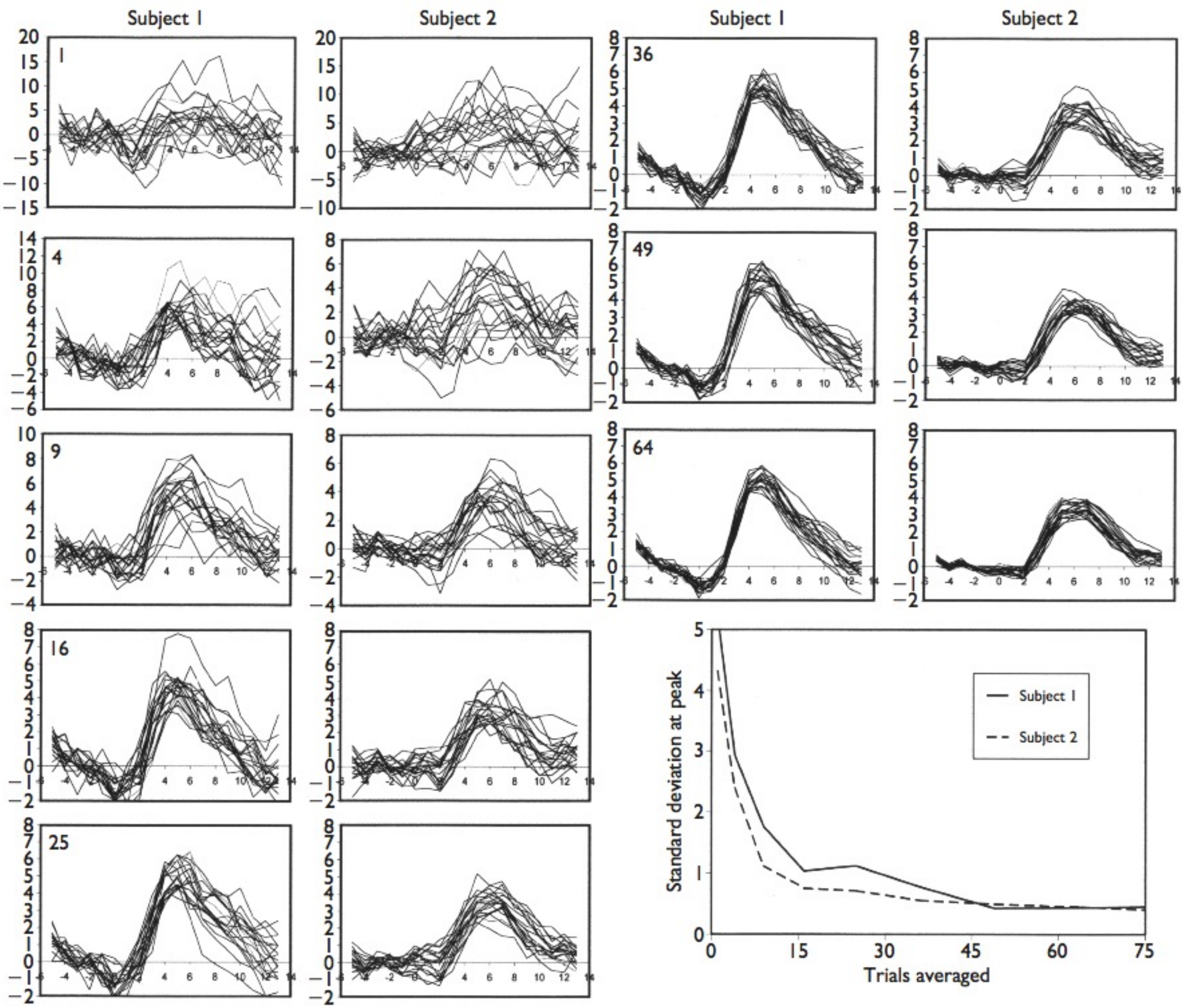
Why are artifacts so problematic ?

Task-based fMRI:

Improved Estimates of BOLD response with Trial Averaging

(e.g. Bandettini et al., 1993; Friston et al., 1995)

Huettel & McCarthy (2001). Neuroreport.



Why are artifacts so problematic ?

Task-based fMRI:

Improved Estimates of BOLD response with Trial Averaging

(e.g. Bandettini et al., 1993; Friston et al., 1995)

Why are artifacts so problematic ?

Task-based fMRI:

Improved Estimates of BOLD response with Trial Averaging

(e.g. Bandettini et al., 1993; Friston et al., 1995)

Resting-state fMRI:

No averaging!

Why are artifacts so problematic ?

Task-based fMRI:

Improved Estimates of BOLD response with Trial Averaging
(e.g. Bandettini et al., 1993; Friston et al., 1995)

Resting-state fMRI:

No averaging!

Anything that causes temporal variation in the BOLD response can influence estimates of covariation

Common Noise Sources in fMRI

Time-varying artifacts affecting the BOLD signal:

Common Noise Sources in fMRI

Time-varying artifacts affecting the BOLD signal:

- Head motion

Common Noise Sources in fMRI

Time-varying artifacts affecting the BOLD signal:

- Head motion
- Non-neural physiological artifacts (e.g. cardiac/respiration cycles, end-tidal $C0_2$, blood pressure fluctuations)

Common Noise Sources in fMRI

Time-varying artifacts affecting the BOLD signal:

- Head motion
- Non-neural physiological artifacts (e.g. cardiac/respiration cycles, end-tidal $C0_2$, blood pressure fluctuations)
- Hardware instabilities

Common Noise Sources in fMRI

Time-varying artifacts affecting the BOLD signal:

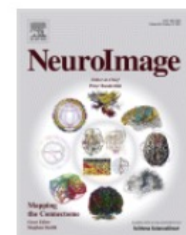
- Head motion
- Non-neural physiological artifacts (e.g. cardiac/respiration cycles, end-tidal $C0_2$, blood pressure fluctuations)
- Hardware instabilities

For good overview:

ELSEVIER

NeuroImage

Volume 80, 15 October 2013, Pages 349–359



Resting-state fMRI confounds and cleanup

Kevin Murphy[a], Rasmus M. Birn[b], Peter A. Bandettini[c, d]

Common Noise Removal Steps

Common Noise Removal Steps

Goal: Remove noise variation in BOLD signal without removing or distorting neurally-derived variation

Common Noise Removal Steps

Goal: Remove noise variation in BOLD signal without removing or distorting neurally-derived variation

Common Noise Removal Steps

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

(spatial) ICA:

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

(spatial) ICA:

- Decompose fMRI time series into statistically independent spatial components (e.g. Comon, 1994; McKeown et al., 1998; Kiviniemi et al., 2003; Beckmann et al., 2005; Perlbarg et al., 2007)

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

(spatial) ICA:

- Decompose fMRI time series into statistically independent spatial components (e.g. Comon, 1994; McKeown et al., 1998; Kiviniemi et al., 2003; Beckmann et al., 2005; Perlbarg et al., 2007)

- Can match to artifact templates and discard

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

(spatial) ICA:

- Decompose fMRI time series into statistically independent spatial components (e.g. Comon, 1994; McKeown et al., 1998; Kiviniemi et al., 2003; Beckmann et al., 2005; Perlbarg et al., 2007)
- Can match to artifact templates and discard
- Can work as long as noise sources are independent of neurally-derived BOLD components

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

(spatial) ICA:

- Decompose fMRI time series into statistically independent spatial components (e.g. Comon, 1994; McKeown et al., 1998; Kiviniemi et al., 2003; Beckmann et al., 2005; Perlbarg et al., 2007)

- Can match to artifact templates and discard

- Can work as long as noise sources are independent of neurally-derived BOLD components

- Combine with multi-echo EPI to sort BOLD from non-BOLD components (Kundu et al., 2012)

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

(spatial) ICA:

- Decompose fMRI time series into statistically independent spatial components (e.g. Comon, 1994; McKeown et al., 1998; Kiviniemi et al., 2003; Beckmann et al., 2005; Perlbarg et al., 2007)

- Can match to artifact templates and discard

- Can work as long as noise sources are independent of neurally-derived BOLD components

- **Combine with multi-echo EPI to sort BOLD from non-BOLD components (Kundu et al., 2012)**

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

Regression:

Model Nuisance Variables and Subtract from Original Timeseries

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

Regression:

Model Nuisance Variables and Subtract from Original Timeseries

Common Nuisance Regressors (varies by lab):

- Motion Parameters

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

Regression:

Model Nuisance Variables and Subtract from Original Timeseries

Common Nuisance Regressors (varies by lab):

- Motion Parameters
- Ventricles
- White Matter

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

Regression:

Model Nuisance Variables and Subtract from Original Timeseries

Common Nuisance Regressors (varies by lab):

- Motion Parameters
- Ventricles
- White Matter
- Bandpass Filtering < 0.1 Hz (??? ***Sampling Rate***)

Common Noise Removal Steps

Basic Methods: Multiple Regression vs. ICA

Regression:

Model Nuisance Variables and Subtract from Original Timeseries

Common Nuisance Regressors (varies by lab):

- Motion Parameters
- Ventricles
- White Matter
- Bandpass Filtering < 0.1 Hz (??? ***Sampling Rate***)
- "Global Signal"

How do we know if any of this stuff is working?

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

- transient head motion (Framewise Displacement, or FD)

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

Compare pre- and post-cleaning to see if artifact is gone

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

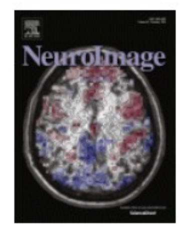
- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

Compare pre- and post-cleaning to see if artifact is gone

ELSEVIER

NeuroImage

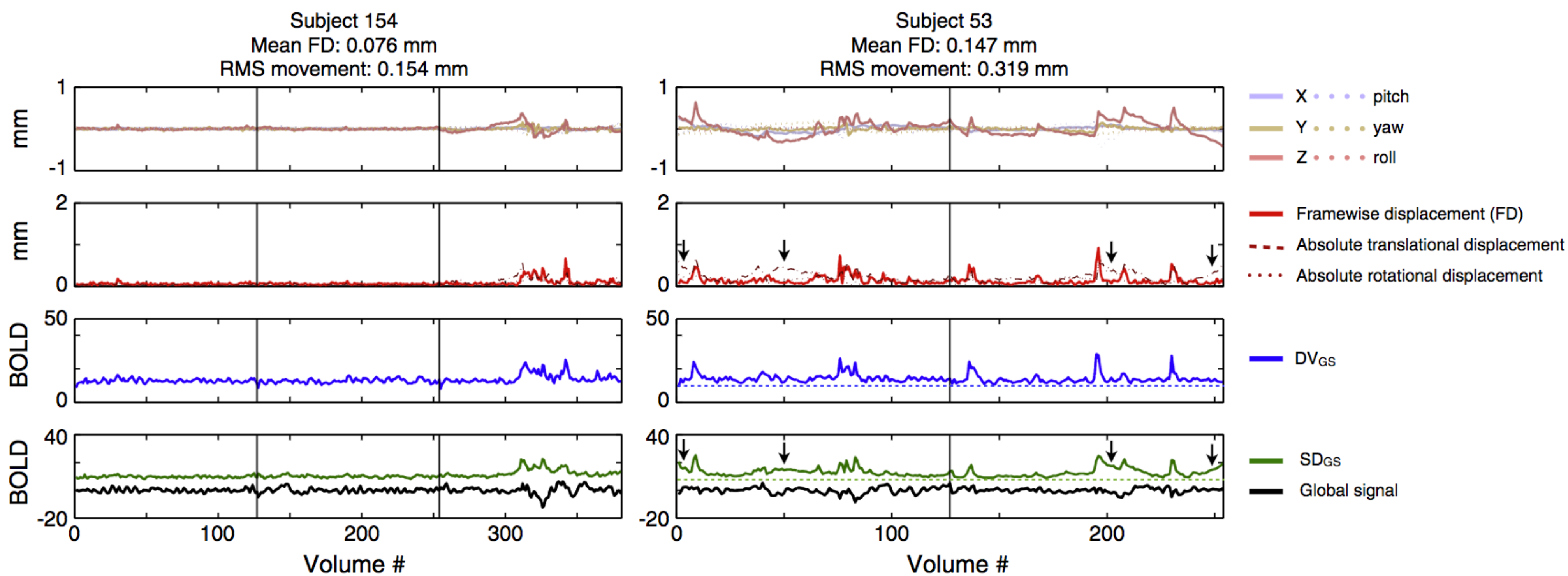
Volume 84, 1 January 2014, Pages 320-341



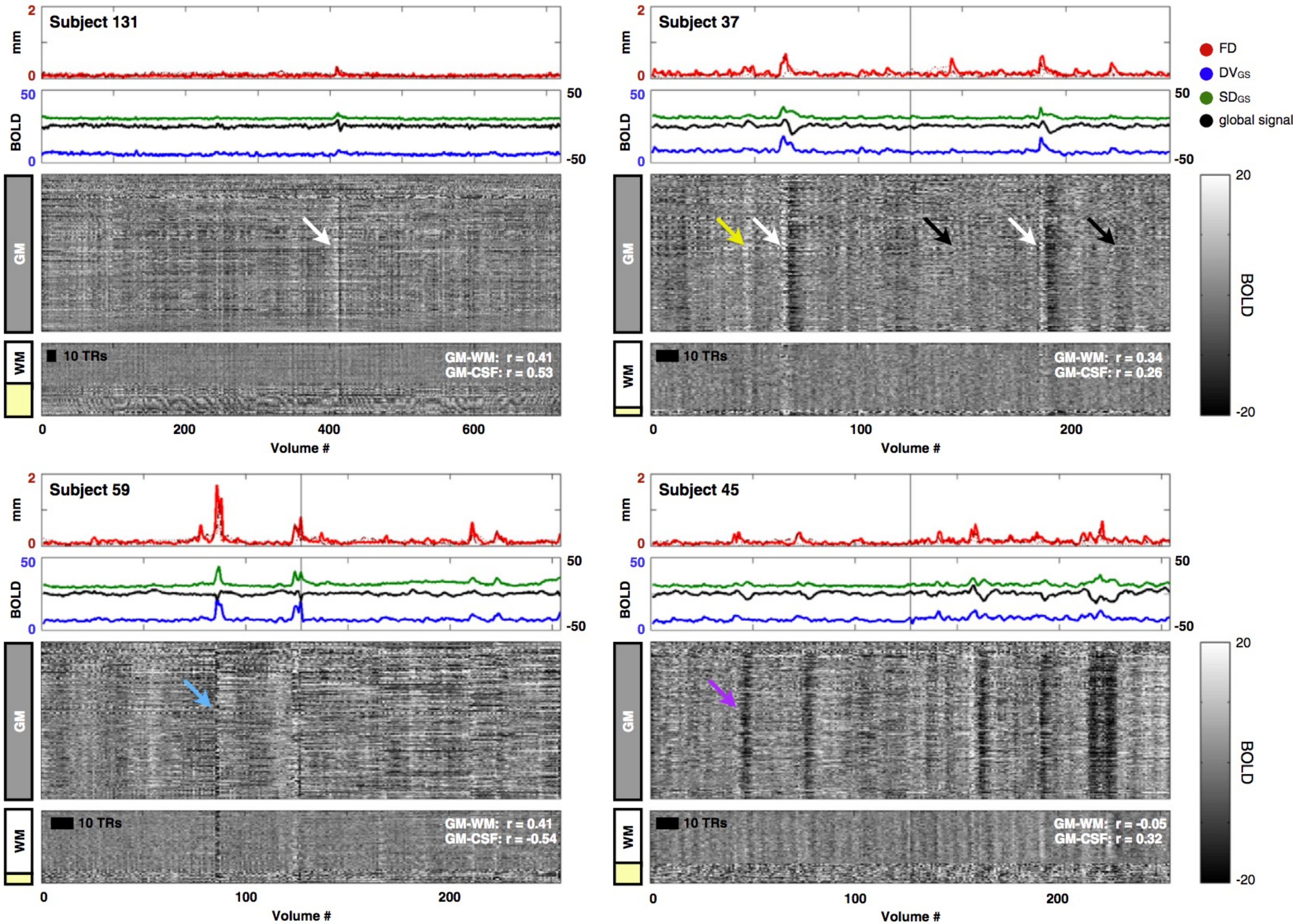
Methods to detect, characterize, and remove motion artifact in resting state fMRI

Jonathan D. Power [a], Anish Mitra [a], Timothy O. Laumann [a], Abraham Z. Snyder [a, b], Bradley L. Schlaggar [a, b, c, d], Steven E. Petersen [a, b, d, e, f, g]

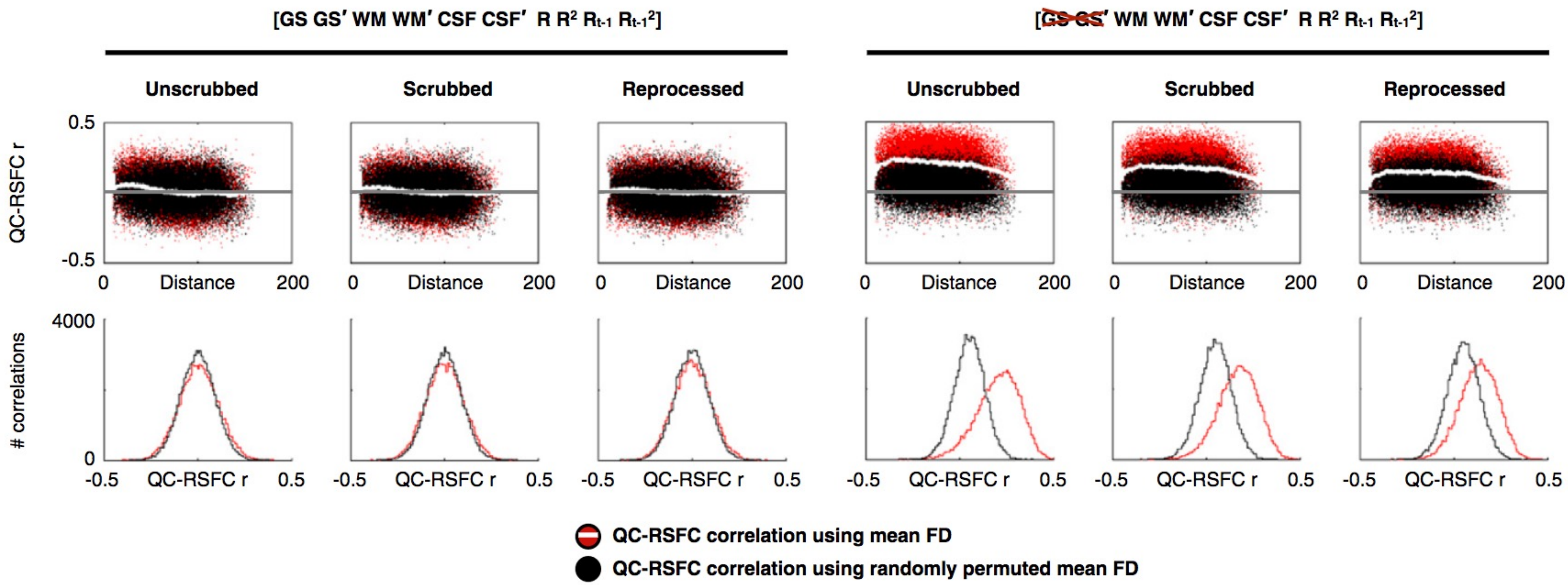
Transient head motions lead to corresponding changes in the BOLD signal

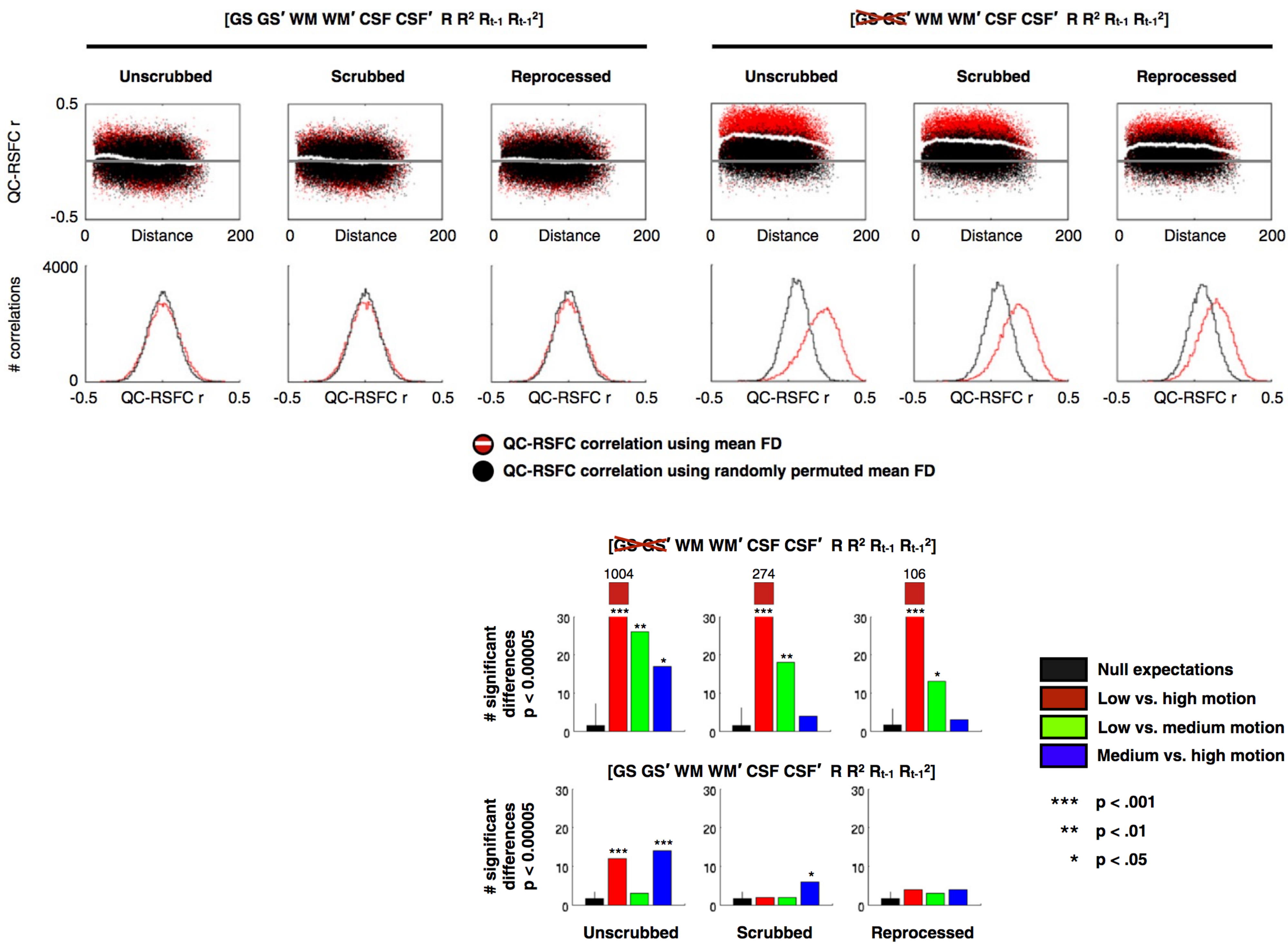


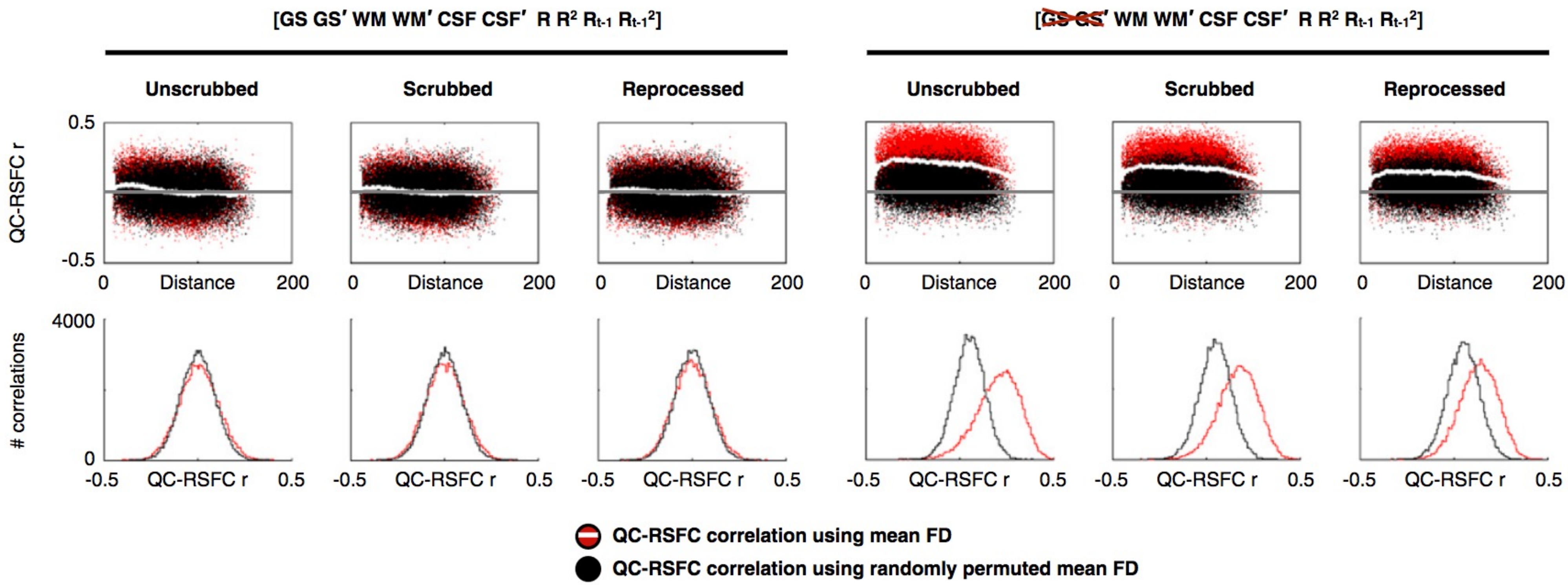
Power et al. (2014). Neuroimage

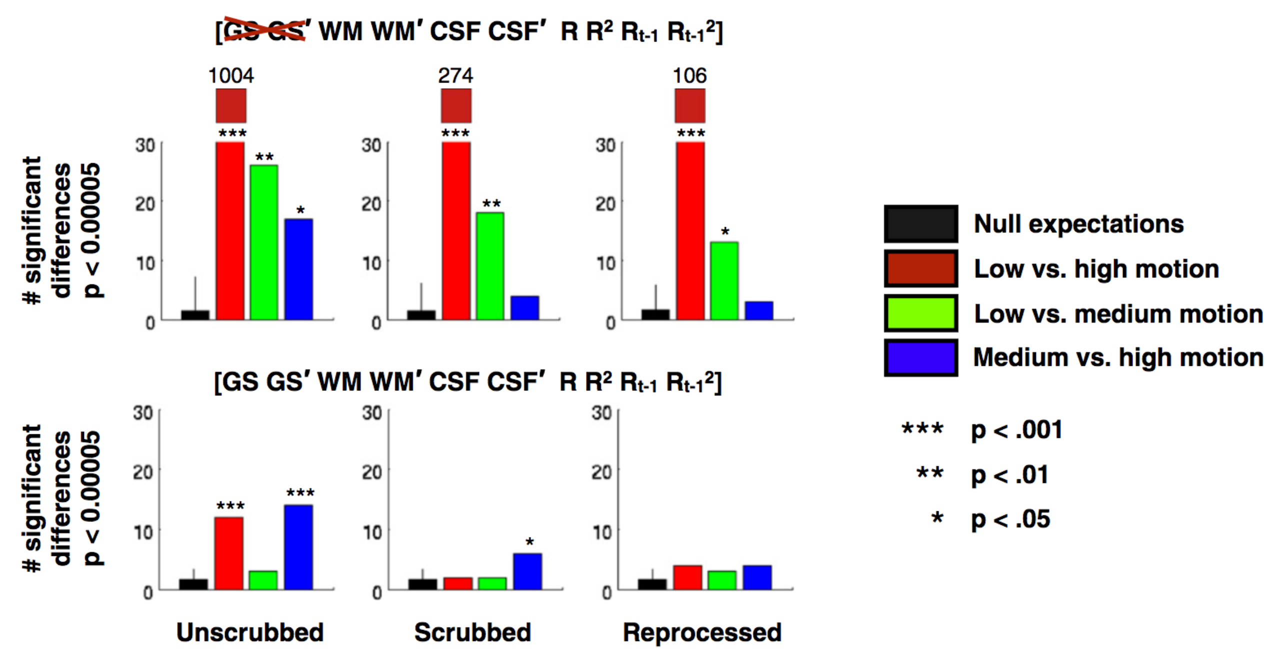


Subject 131
Subject 37
Subject 59
Subject 45
mm
BOLD
GM
WM
10 TRs
GM-WM: r = 0.41
GM-CSF: r = 0.53
GM-WM: r = 0.34
GM-CSF: r = 0.26
GM-WM: r = 0.41
GM-CSF: r = -0.54
GM-WM: r = -0.05
GM-CSF: r = 0.32
Volume #
FD
DVGS
SDGS
global signal


[GS GS′ WM WM′ CSF CSF′ R R² Rt-1 Rt-1²]
[GS GS′ WM WM′ CSF CSF′ R R² Rt-1 Rt-1²]
Unscrubbed
Scrubbed
Reprocessed
Unscrubbed
Scrubbed
Reprocessed
QC-RSFC r
0.5
-0.5
0
Distance
200
correlations
4000
0
-0.5
QC-RSFC r
0.5
QC-RSFC correlation using mean FD
QC-RSFC correlation using randomly permuted mean FD


[GS GS′ WM WM′ CSF CSF′ R R² Rt-1 Rt-1²]
[GS GS′ WM WM′ CSF CSF′ R R² Rt-1 Rt-1²]
Unscrubbed
Scrubbed
Reprocessed
QC-RSFC r
0.5
-0.5
0
Distance
200
correlations
4000
QC-RSFC correlation using mean FD
QC-RSFC correlation using randomly permuted mean FD
1004
274
106
significant differences p < 0.00005
Null expectations
Low vs. high motion
Low vs. medium motion
Medium vs. high motion
*** p < .001
** p < .01
* p < .05





Take Home:

Censor high-motion TRs,
Global Signal Regression

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

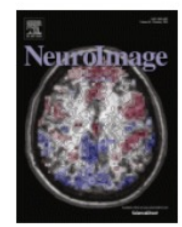
- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

Compare pre- and post-cleaning to see if artifact is gone

ELSEVIER

NeuroImage

Volume 84, 1 January 2014, Pages 320-341



Methods to detect, characterize, and remove motion artifact in resting state fMRI

Jonathan D. Power [a], Anish Mitra [a], Timothy O. Laumann [a], Abraham Z. Snyder [a, b], Bradley L. Schlaggar [a, b, c, d], Steven E. Petersen [a, b, d, e, f, g]

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

Compare pre- and post-cleaning to see if artifact is gone

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

Compare pre- and post-cleaning to see if artifact is gone

A distinct approach is to use **reliability** as a guide:

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

Compare pre- and post-cleaning to see if artifact is gone

A distinct approach is to use **reliability** as a guide: take two scans and pick the pipeline that gives best test-retest reliability

How do we know if any of this stuff is working?

We have some measures of certain artifacts (e.g.):

- transient head motion (Framewise Displacement, or FD)
- independent respiration and cardiac traces

Compare pre- and post-cleaning to see if artifact is gone

A distinct approach is to use **reliability** as a guide: take two scans and pick the pipeline that gives best test-retest reliability

Underlying assumption is that when noise is removed, "real" patterns should repeat

Test-retest reliability: agreement across repeated measurements taken by a single person or instrument on the same item, under the same conditions, and in a short period of time

Test-retest reliability: agreement across repeated measurements taken by a single person or instrument on the same item, under the same conditions, and in a short period of time

A test or measure cannot be more correlated with a different measure than it is with itself.

Test-retest reliability: agreement across repeated measurements taken by a single person or instrument on the same item, under the same conditions, and in a short period of time

A test or measure cannot be more correlated with a different measure than it is with itself.

Therefore, reliability of fMRI BOLD fluctuations and behavioral measures serves as an upper limit on our ability to measure any brain-behavior relationships.

Cerebral Cortex October 2009;19:2209-2229
doi:10.1093/cercor/bhn256
Advance Access publication February 16, 2009

The Resting Brain: Unconstrained yet Reliable

Zarrar Shehzad[1], A. M. Clare Kelly[1], Philip T. Reiss[2,3], Dylan G. Gee[1], Kristin Gotimer[1], Lucina Q. Uddin[4], Sang Han Lee[3], Daniel S. Margulies[5], Amy Krain Roy[1], Bharat B. Biswal[3,6], Eva Petkova[2,3], F. Xavier Castellanos[1,3] and Michael P. Milham[1]

[1]Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, NYU Child Study Center, New York, NY 10016, USA, [2]Division of Biostatistics, NYU Child Study Center, New York, NY 10016, USA, [3]Nathan Kline Institute for Psychiatric Research, Orangeburg, NY 10962, USA, [4]Department of Psychiatry, Stanford University School of Medicine, Stanford, CA, USA, [5]Berlin School for Mind and Brain, Humboldt Universitat, Berlin, Germany and [6]Department of Radiology, University of Medicine and Dentistry of New Jersey, Newark, NJ 07101, USA

Cerebral Cortex October 2009;19:2209-2229
doi:10.1093/cercor/bhn256
Advance Access publication February 16, 2009

The Resting Brain: Unconstrained yet Reliable

Zarrar Shehzad[1], A. M. Clare Kelly[1], Philip T. Reiss[2,3], Dylan G. Gee[1], Kristin Gotimer[1], Lucina Q. Uddin[4], Sang Han Lee[3], Daniel S. Margulies[5], Amy Krain Roy[1], Bharat B. Biswal[3,6], Eva Petkova[2,3], F. Xavier Castellanos[1,3] and Michael P. Milham[1]

6,

n,

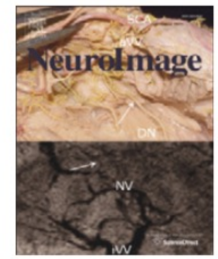
y,

NeuroImage 76 (2013) 183–201

Contents lists available at SciVerse ScienceDirect

NeuroImage

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



A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics

Chao-Gan Yan [a,b,c], Brian Cheung [b], Clare Kelly [c], Stan Colcombe [a], R. Cameron Craddock [b,d], Adriana Di Martino [c], Qingyang Li [b], Xi-Nian Zuo [e], F. Xavier Castellanos [a,c], Michael P. Milham [a,b,*]

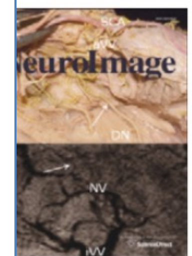
[a] Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA
[b] Center for the Developing Brain, Child Mind Institute, New York, NY, USA
[c] The Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, New York University Child Study Center, New York, NY, USA
[d] Virginia Tech Carilion Research Institute, Roanoke, VA, USA
[e] Key Laboratory of Behavioral Science, Laboratory for Functional Connectome and Development, Magnetic Resonance Imaging Research Center, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

Cerebral Cortex October 2009;19:2209-2229
doi:10.1093/cercor/bhn256
Advance Access publication February 16, 2009

The Resting ... Reliable

...eiss[2,3], Dylan
...ang Han Lee[3],
... Biswal[3,6],
...chael P. Milham[1]

ELSEVIER



A compreh...
micromove...

Chao-Gan Yan...
Adriana Di Ma... ,*

[a] Nathan Kline Institute...
[b] Center for the Develop...
[c] The Phyllis Green and...
[d] Virginia Tech Carilion Research Institute, Roanoke, VA, USA
[e] Key Laboratory of Behavioral Science, Laboratory for Functional Connectome and Development, Magnetic Resonance Imaging Research Center, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Neuroinformatics
» Brain imaging
» Functional magnetic resonance imaging
» Cognitive neuroscience

Received: 19 June 2014
Accepted: 14 October 2014
Published: 9 December 2014

An open science resource for establishing reliability and reproducibility in functional connectomics

Xi-Nian Zuo[1,2,*], Jeffrey S. Anderson[3], Pierre Bellec[4], Rasmus M. Birn[5], Bharat B. Biswal[6], Janusch Blautzik[7], John C.S. Breitner[8], Randy L. Buckner[9], Vince D. Calhoun[10], F. Xavier Castellanos[11,12], Antao Chen[2], Bing Chen[13], Jiangtao Chen[2], Xu Chen[2], Stanley J. Colcombe[11], William Courtney[10], R. Cameron Craddock[11,14], Adriana Di Martino[12], Hao-Ming Dong[1,15], Xiaolan Fu[1,16], Qiyong Gong[17], Krzysztof J. Gorgolewski[18], Ying Han[19], Ye He[1,15], Yong He[20], Erica Ho[11,14], Avram Holmes[21], Xiao-Hui Hou[1,15], Jeremy Huckins[22], Tianzi Jiang[23], Yi Jiang[1], William Kelley[22], Clare Kelly[12], Margaret King[10], Stephen M. LaConte[24], Janet E. Lainhart[5], Xu Lei[2], Hui-Jie Li[1], Kaiming Li[17], Kuncheng Li[25], Qixiang Lin[20], Dongqiang Liu[13], Jia Liu[20], Xun Liu[1], Yijun Liu[2], Guangming Lu[26], Jie Lu[25], Beatriz Luna[27], Jing Luo[28], Daniel Lurie[11,14], Ying Mao[29], Daniel S. Margulies[18], Andrew R. Mayer[10], Thomas Meindl[7], Mary E. Meyerand[30], Weizhi Nan[1,15], Jared A. Nielsen[3], David O'Connor[11,14], David Paulsen[27], Vivek Prabhakaran[31], Zhigang Qi[25], Jiang Qiu[2], Chunhong Shao[32], Zarrar Shehzad[11,14], Weijun Tang[33], Arno Villringer[34], Huiling Wang[35], Kai Wang[1,15], Dongtao Wei[2], Gao-Xia Wei[1], Xu-Chu Weng[13], Xuehai Wu[29], Ting Xu[1,11,14], Ning Yang[1,15], Zhi Yang[1], Yu-Feng Zang[13], Lei Zhang[1,15], Qinglin Zhang[2], Zhe Zhang[1,15], Zhiqiang Zhang[26], Ke Zhao[1], Zonglei Zhen[20], Yuan Zhou[1], Xing-Ting Zhu[1,15] & Michael P. Milham[11,14]

Cerebral Cortex October 2009;19:2209-2229
doi:10.1093/cercor/bhn256
Advance Access publication February 16, 2009

The Resting
Reliable

eiss[2,3], Dylan
ang Han Lee[3],
. Biswal[3,6],
chael P. Milham[1]

ELSEV

NeuroImage

6,
n,
y,

A con
micr

Chao-C
Adrian

[a] Nathan K
[b] Center for the Develop
[c] The Phyllis Green and
[d] Virginia Tech Carilion Research Institute, Roanoke, VA, USA
[e] Key Laboratory of Behavioral Science, Laboratory for Functional Connectome and Development, Magnetic Resonance Imaging Research Center, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

SCIENTIFIC DATA

OPEN An open science resource for

Zhiqiang Zhang[26], Ke Zhao[1], Zonglei Zhen[20], Yuan Zhou[1], Xing-Ting Zhu[1,15] & Michael P. Milham[11,14]

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 19 February 2015
doi: 10.3389/fnins.2015.00048

Evaluating the reliability of different preprocessing steps to estimate graph theoretical measures in resting state fMRI data

Nathassia K. Aurich[1], José O. Alves Filho[1], Ana M. Marques da Silva[1,2,3] and Alexandre R. Franco[1,2,4]*

[1] *Faculdade de Engenharia, PUCRS, Porto Alegre, Brazil*
[2] *Instituto do Cérebro do Rio Grande do Sul (InsCer-RS), PUCRS, Porto Alegre, Brazil*
[3] *Faculdade de Física, PUCRS, Porto Alegre, Brazil*
[4] *Faculdade de Medicina, PUCRS, Porto Alegre, Brazil*

Cerebral Cortex October 2009;19:2209-2229
doi:10.1093/cercor/bhn256
Advance Access publication February 16, 2009

The Resting ... Reliable

SCIENTIFIC DATA

OPEN An open science resource for

frontiers in NEUROSCIENCE
ORIGINAL RESEARCH ARTICLE
published: 19 February 2015
doi: 10.3389/fnins.2015.00048

NeuroImage 117 (2015) 67–79

Contents lists available at ScienceDirect

NeuroImage

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

ELSEVIER

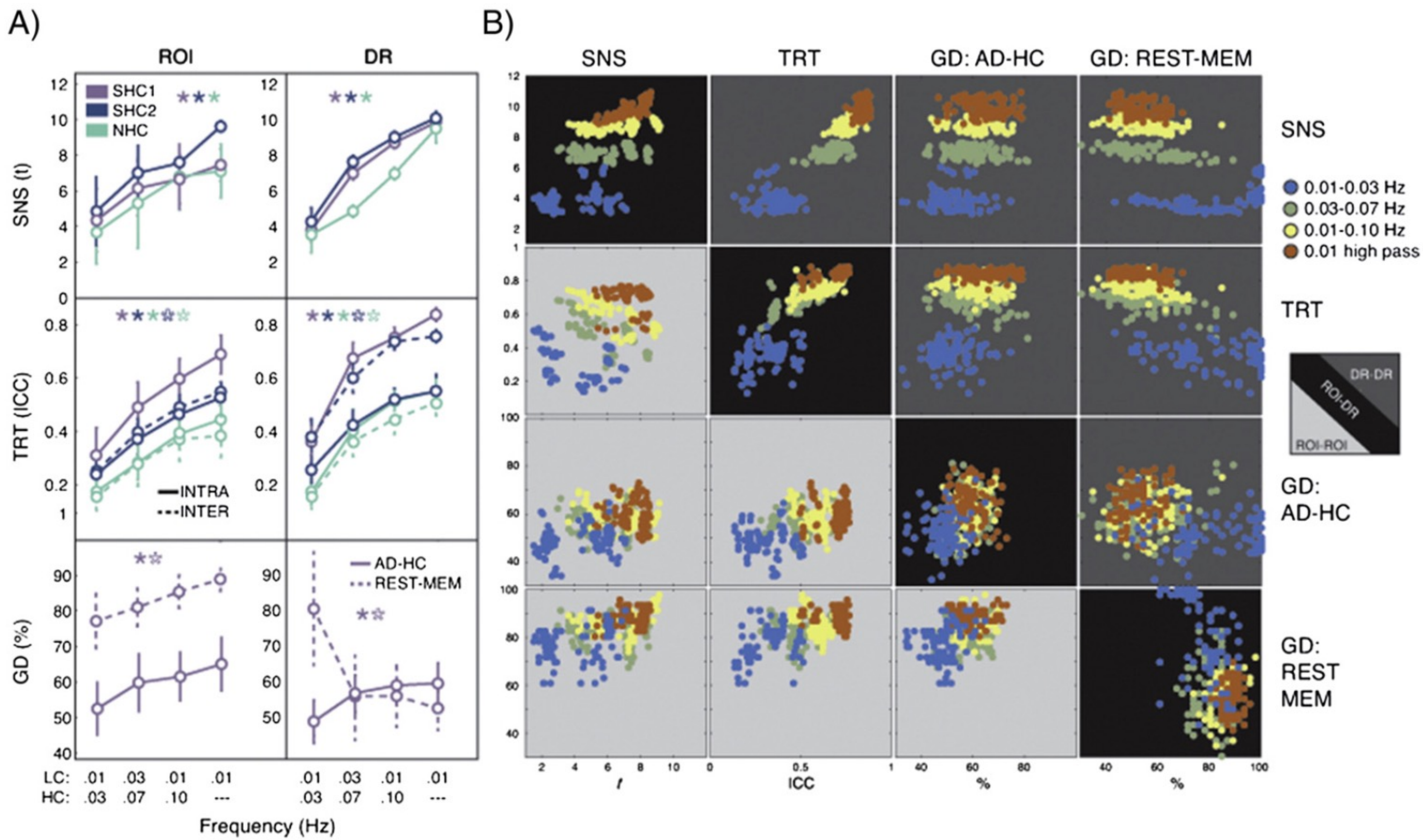
Optimization of rs-fMRI Pre-processing for Enhanced Signal-Noise Separation, Test-Retest Reliability, and Group Discrimination

CrossMark

William R. Shirer [a,*,1], Heidi Jiang [a,b,*,1], Collin M. Price [a], Bernard Ng [a,c], Michael D. Greicius [a]

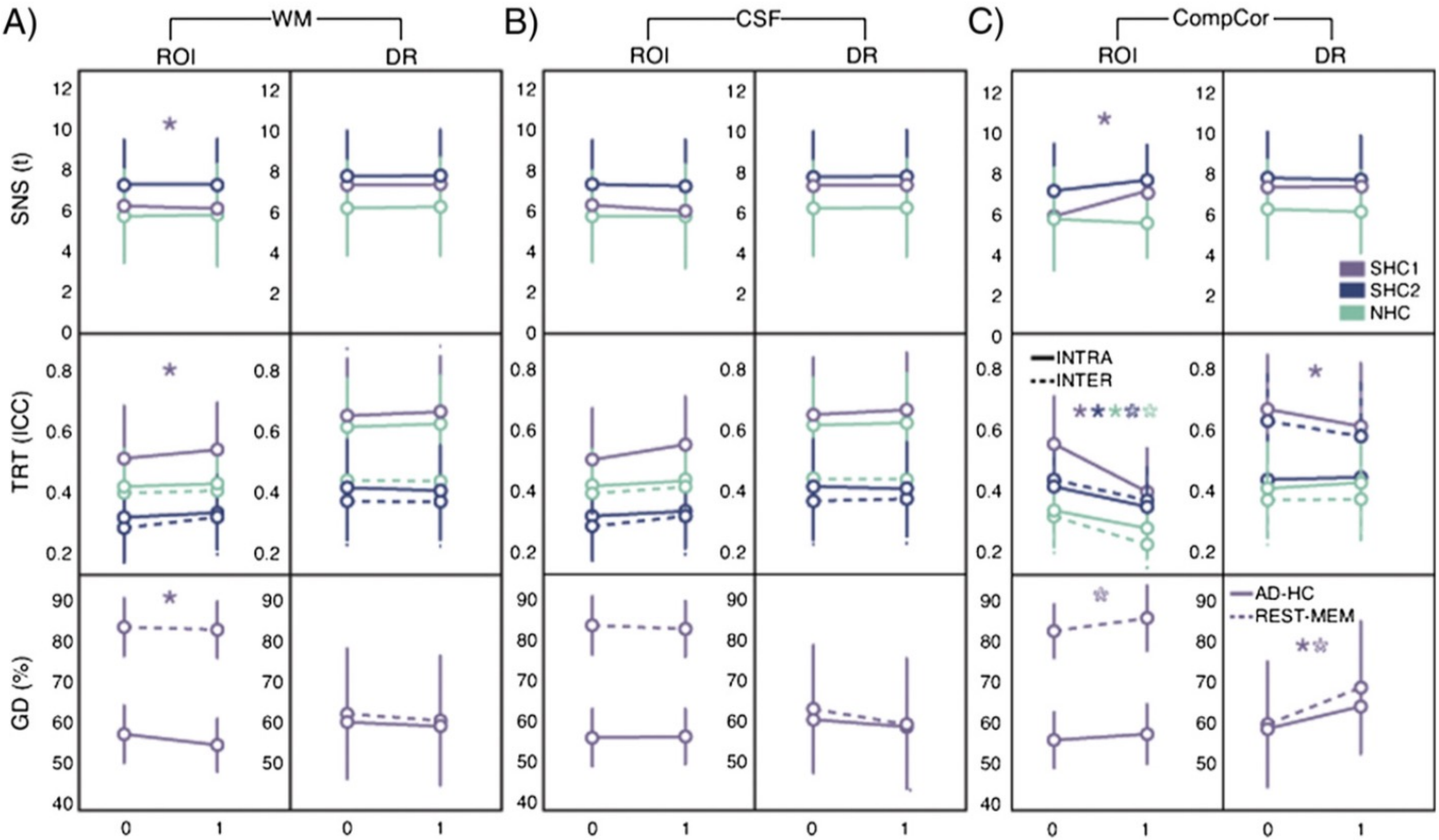
[a] *Functional Imaging in Neuropsychiatric Disorders (FIND) Lab, Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, USA*
[b] *Interdepartmental Program in Neuroscience, Northwestern University, Chicago, USA*
[c] *Parietal Team, Neurospin, INRIA Saclay, France*

Which procedures affect test-retest reliability (and other measures of data quality) ?



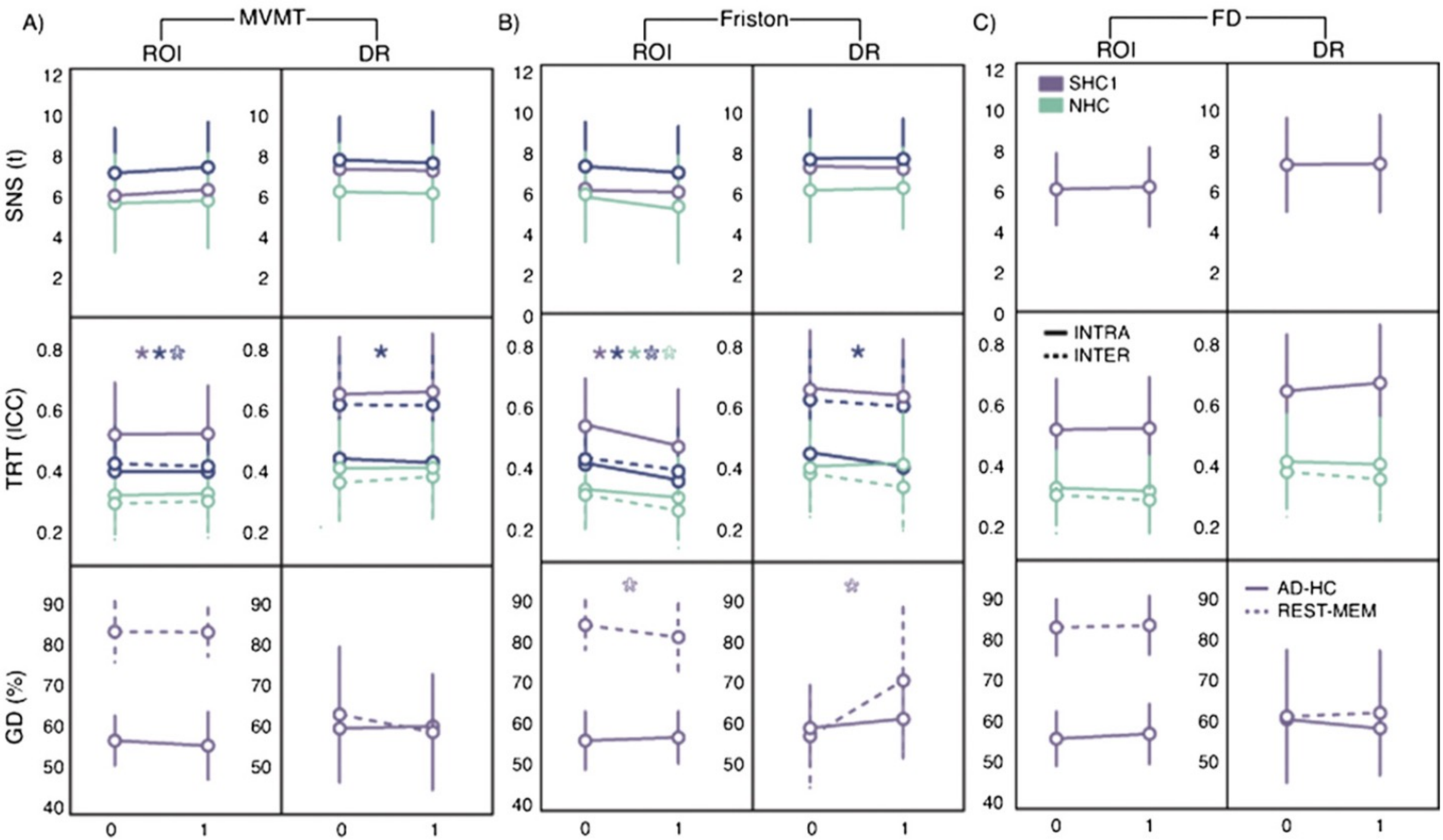
Shirer et al. (2015). Neuroimage

Which procedures affect test-retest reliability (and other measures of data quality) ?

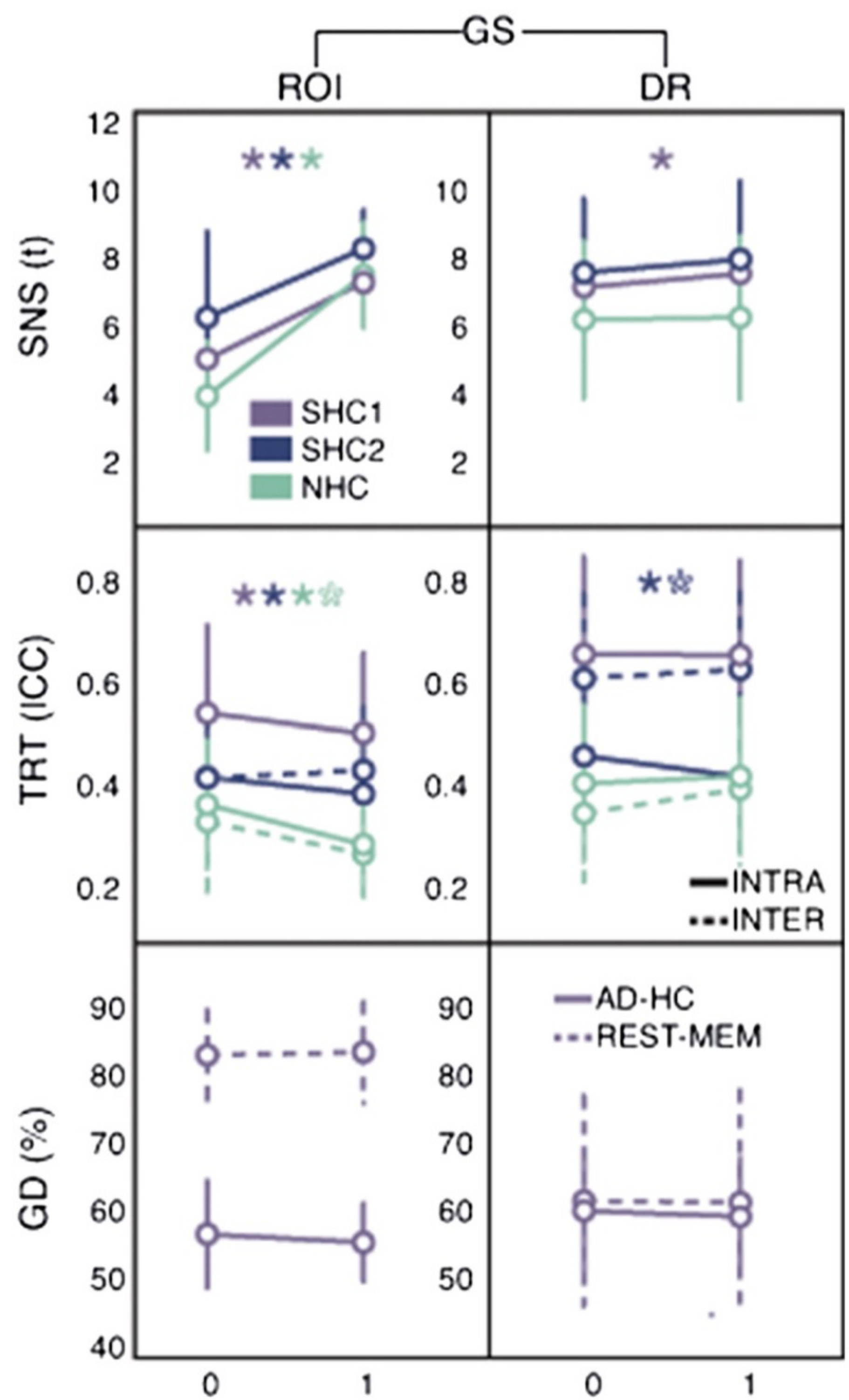


Shirer et al. (2015). Neuroimage

Which procedures affect test-retest reliability (and other measures of data quality) ?



Shirer et al. (2015). Neuroimage



Shirer et al. (2015). Neuroimage

Prediction

Prediction

If test-retest **reliability** is a good indicator of which cleaning procedures best remove noise and spare “neurogenic” signals of interest, measures of **validity** should follow the same pattern

Prediction

If test-retest **reliability** is a good indicator of which cleaning procedures best remove noise and spare “neurogenic” signals of interest, measures of **validity** should follow the same pattern

Validity in this context refers to the extent to which we can use fMRI fluctuations/covariation to predict an independent behavioral measure that indexes the ability of interest

Previous Examples of Validity

Previous Examples of Validity

- Use task-based localizer for faces versus other classes of objects and show that resting-state correlations among face-selective regions predict face processing abilities behaviorally (Zhu et al., 2011)

Previous Examples of Validity

- Use task-based localizer for faces versus other classes of objects and show that resting-state correlations among face-selective regions predict face processing abilities behaviorally (Zhu et al., 2011)

- Use lateralized brain regions to predict related lateralized behavioral abilities (e.g. language and visuospatial processing) (Gotts et al., 2013)

Previous Examples of Validity

- Use task-based localizer for faces versus other classes of objects and show that resting-state correlations among face-selective regions predict face processing abilities behaviorally (Zhu et al., 2011)

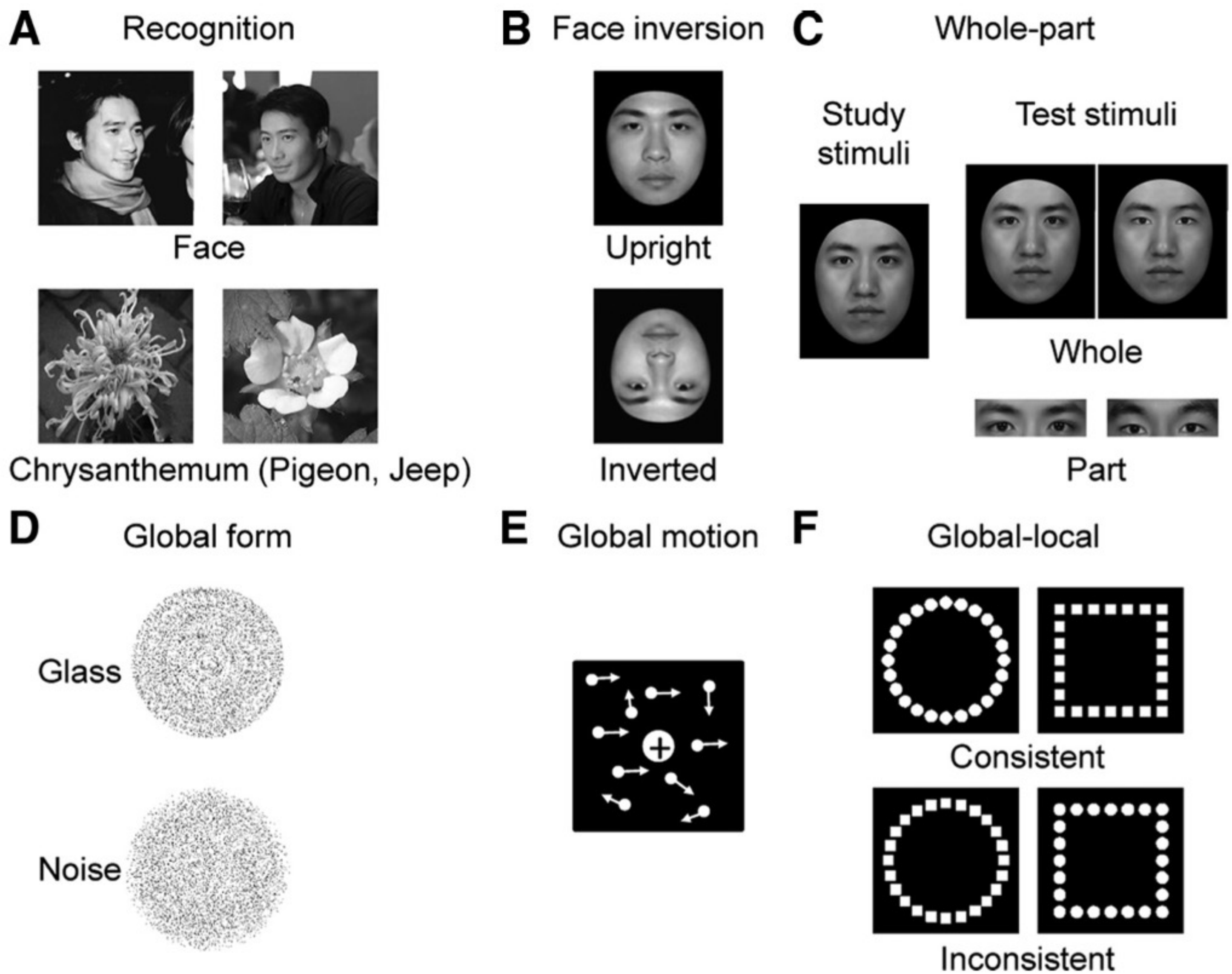
- Use lateralized brain regions to predict related lateralized behavioral abilities (e.g. language and visuospatial processing) (Gotts et al., 2013)

- Comparing clinical versus control group in resting-state correlations and predicting independent measures of clinical symptoms using the same regions (Gotts et al., 2012)

Resting-state Correlations Among Face-Selective Regions Predict Face Processing Ability Behaviorally

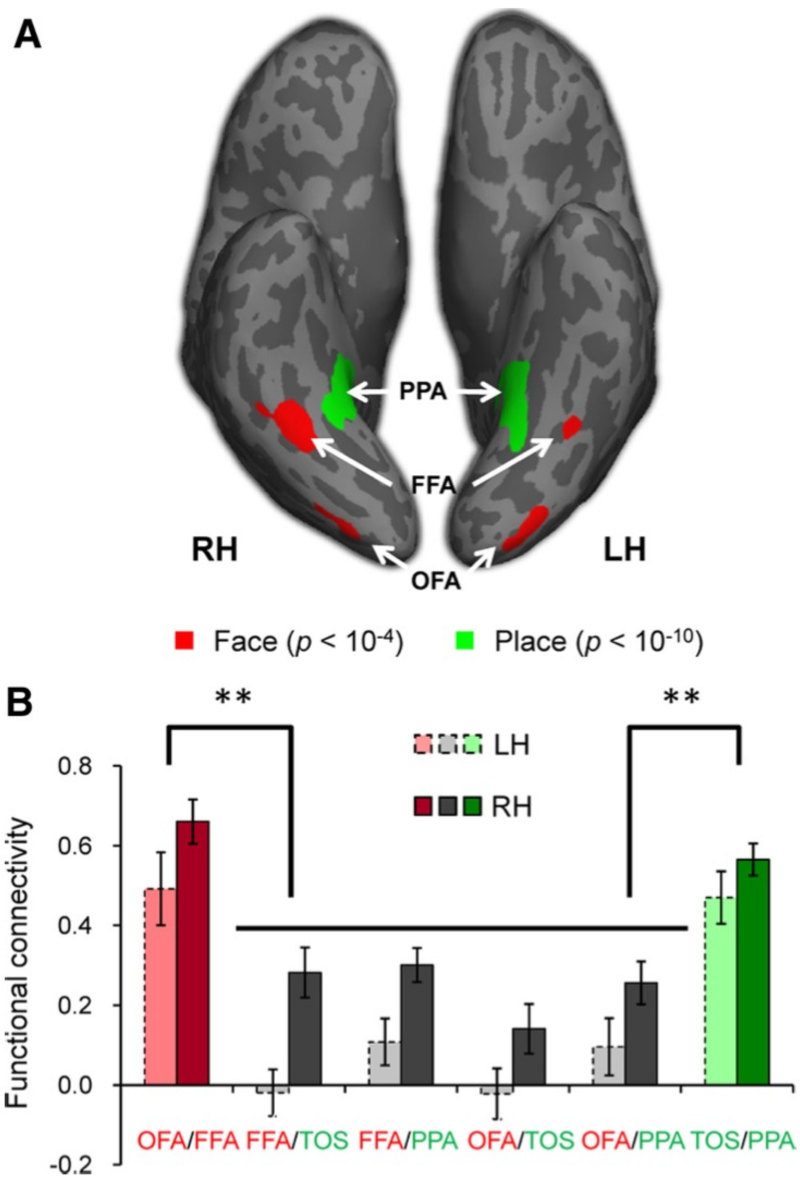
Resting-state Correlations Among Face-Selective Regions Predict Face Processing Ability Behaviorally

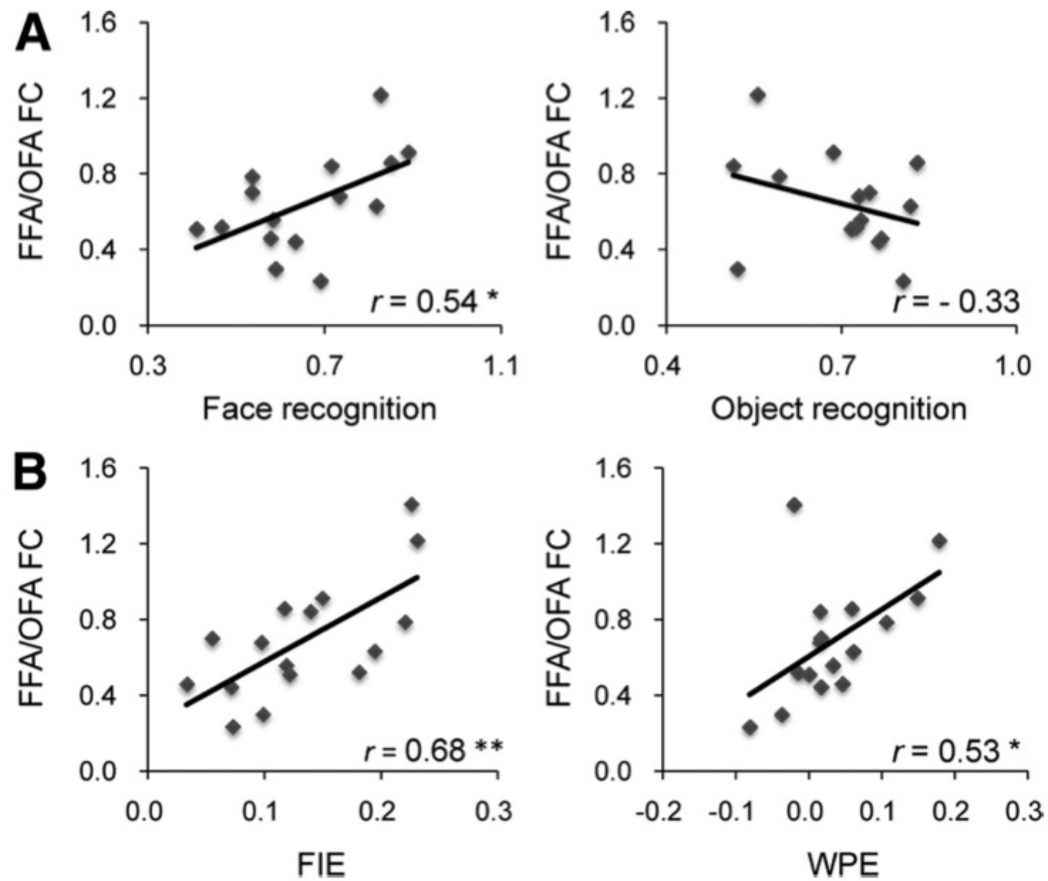
Zhu et al. (2011). J Neurosci



Resting-state Correlations Among Face-Selective Regions Predict Face Processing Ability Behaviorally

Zhu et al. (2011). J Neurosci





Inversion Effect

Whole-part Effect

Autism (ASD) vs. Typically Developing (TD)

Autism (ASD) vs. Typically Developing (TD)

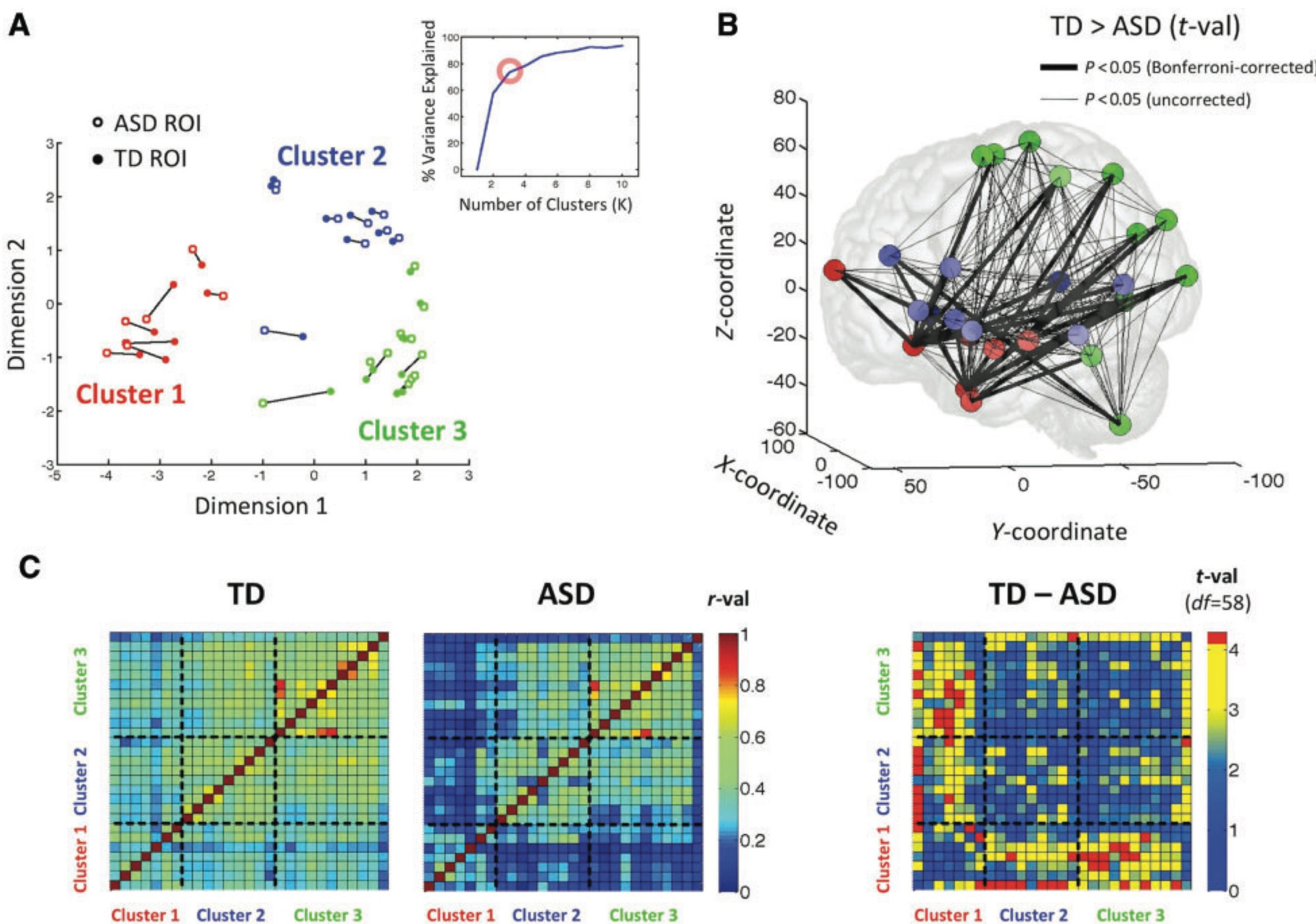
doi:10.1093/brain/aws160 Brain 2012: 135; 2711–2725 | 2711

BRAIN
A JOURNAL OF NEUROLOGY

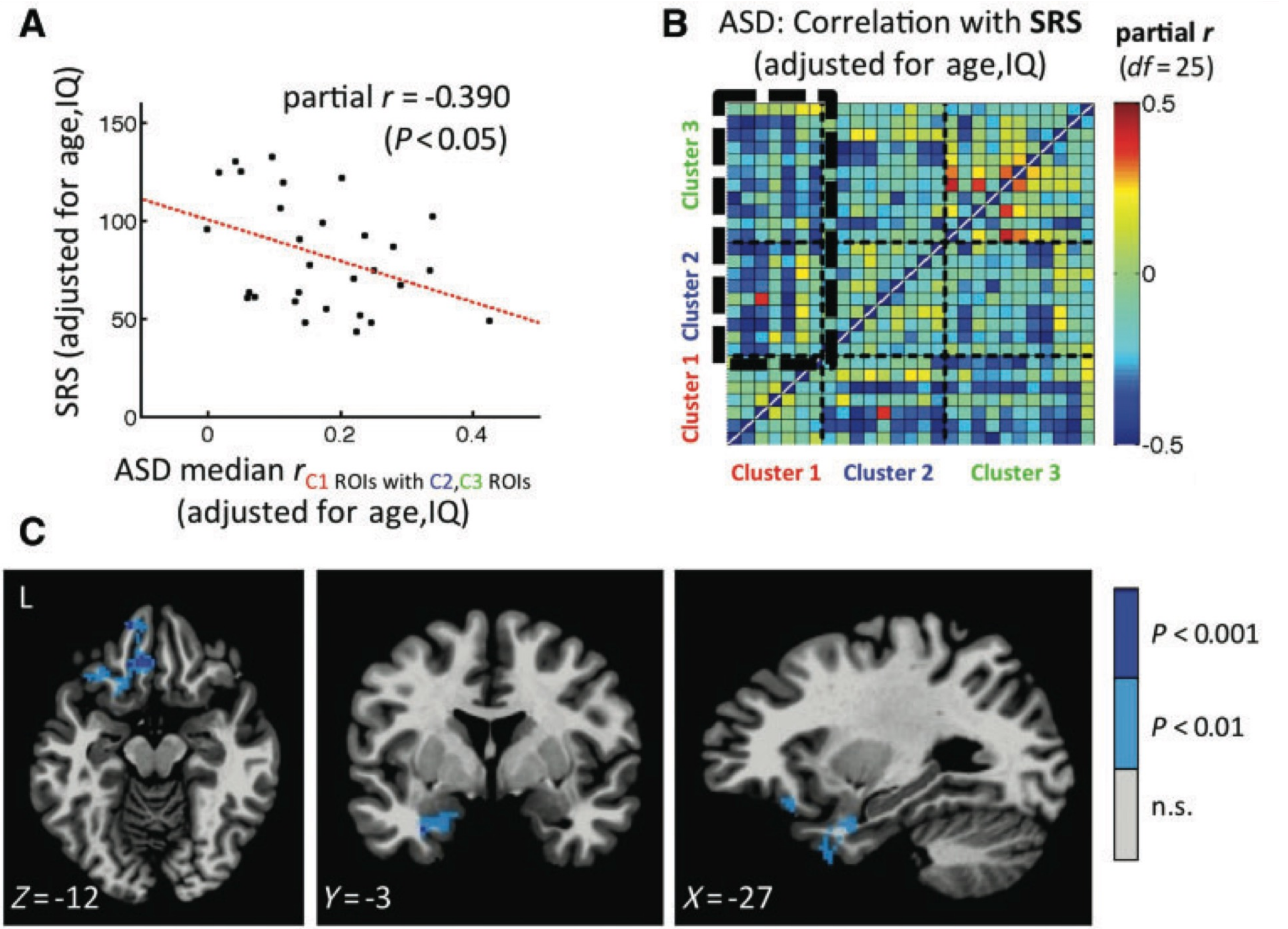
Fractionation of social brain circuits in autism spectrum disorders

Stephen J. Gotts,[1] W. Kyle Simmons,[2] Lydia A. Milbury,[1] Gregory L. Wallace,[1]
Robert W. Cox[3] and Alex Martin[1]

Whole-brain Differences in Functional Connectivity: TD > ASD



Agreement with Social Symptom Correlations (ASD only)



Does Preprocessing Affect Validity?

Does Preprocessing Affect Validity?

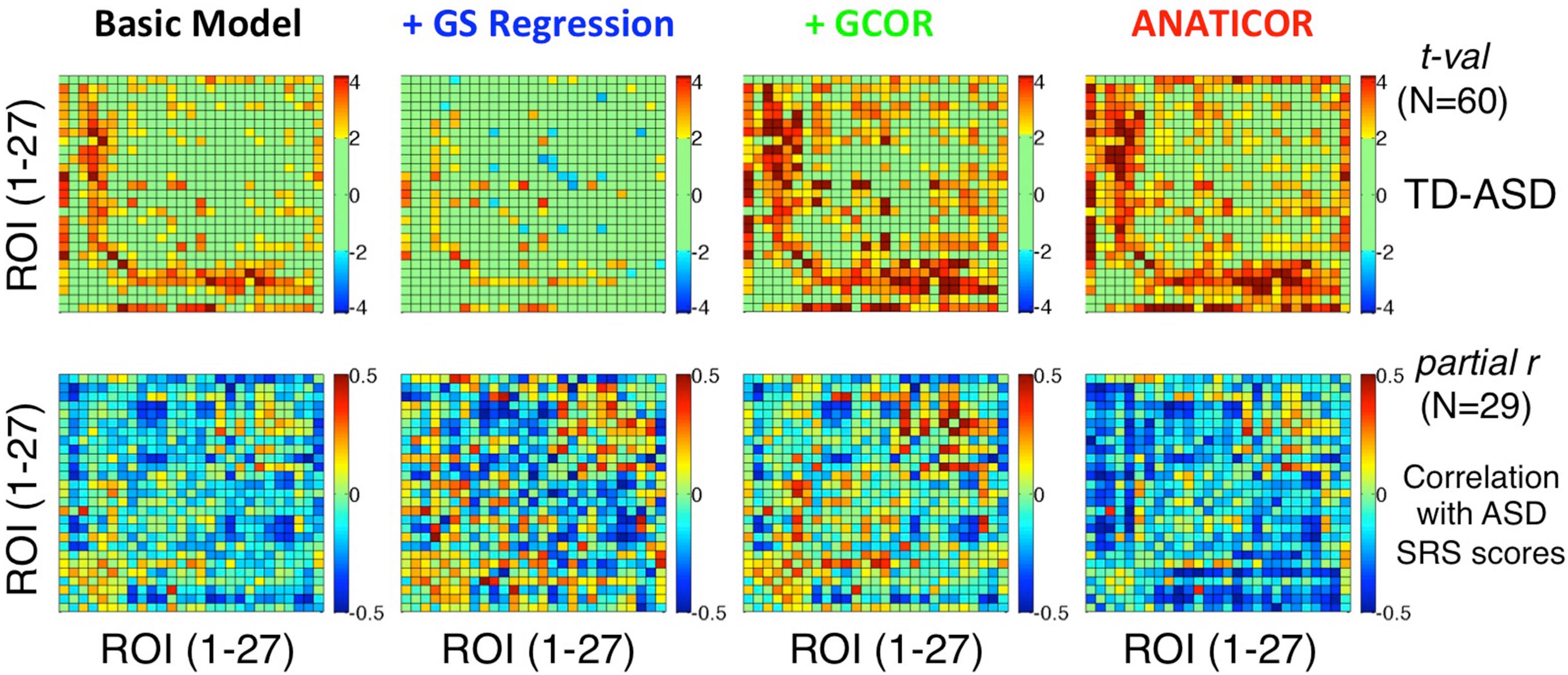
frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 12 July 2013
doi: 10.3389/fnhum.2013.00356

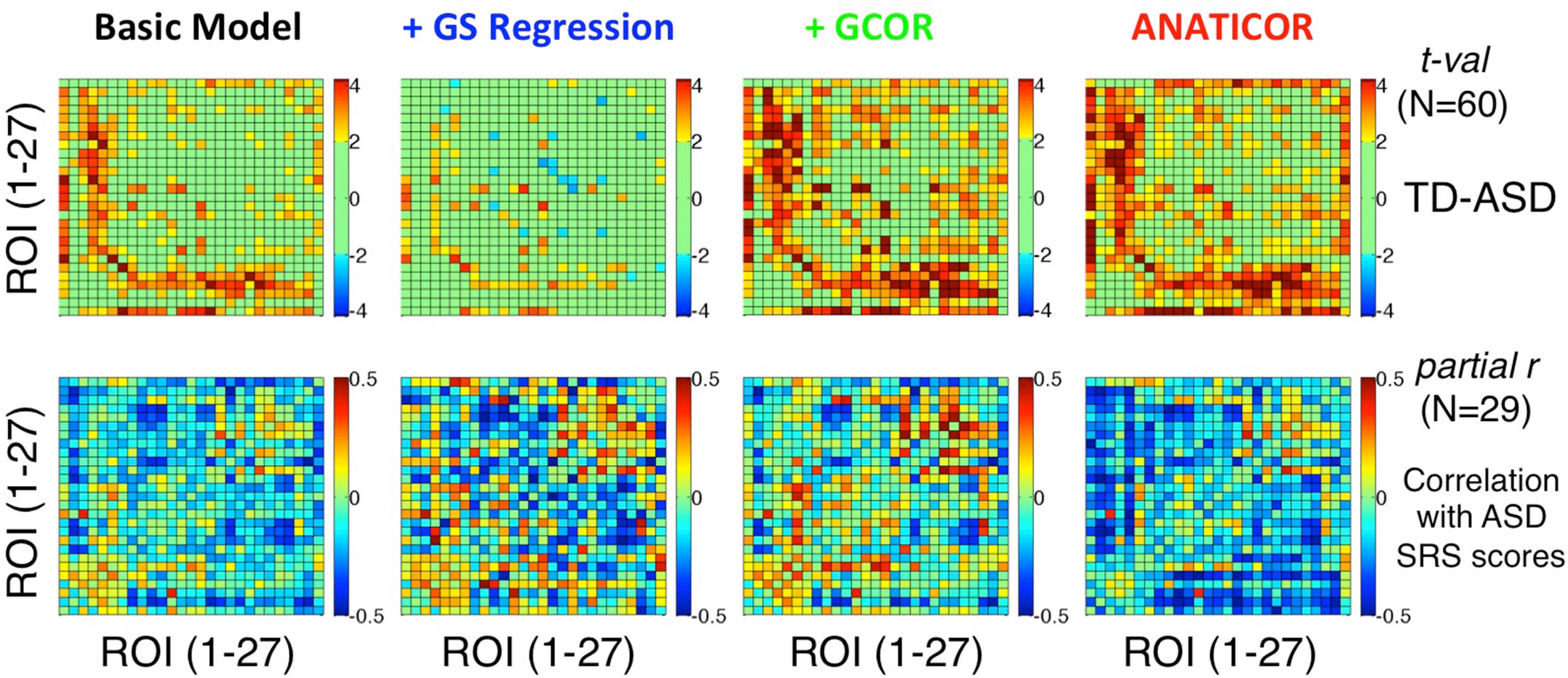
The perils of global signal regression for group comparisons: a case study of Autism Spectrum Disorders

Stephen J. Gotts[1]*, Ziad S. Saad[2], Hang Joon Jo[2], Gregory L. Wallace[1], Robert W. Cox[2] and Alex Martin[1]

[1] *Section on Cognitive Neuropsychology, Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA*
[2] *Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA*

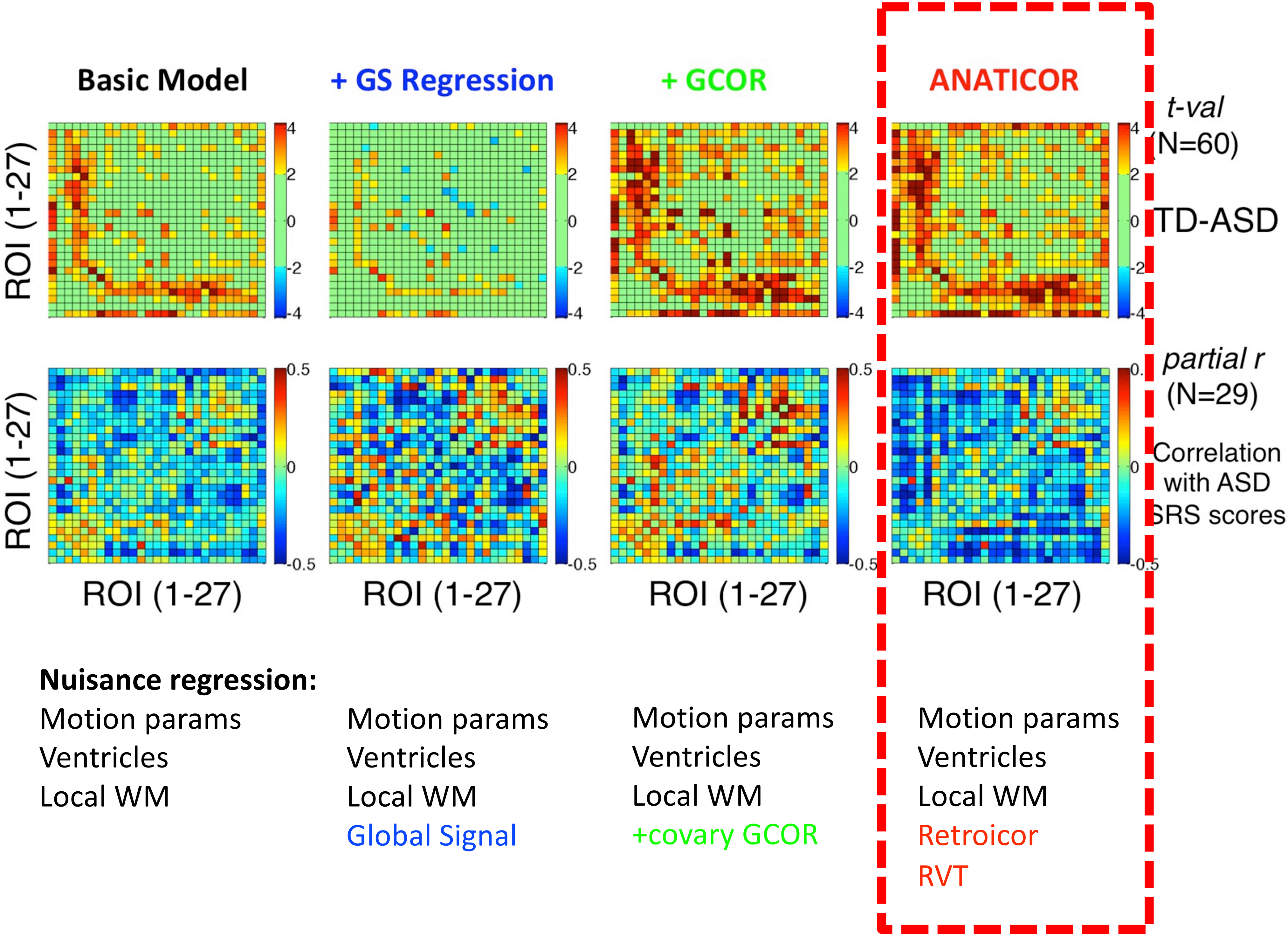


Basic Model
+ GS Regression
+ GCOR
ANATICOR
t-val
(N=60)
TD-ASD
partial r
(N=29)
Correlation with ASD SRS scores
ROI (1-27)
ROI (1-27)
ROI (1-27)
ROI (1-27)
ROI (1-27)
ROI (1-27)



Nuisance regression:

Basic Model	+ GS Regression	+ GCOR	ANATICOR
Motion params	Motion params	Motion params	Motion params
Ventricles	Ventricles	Ventricles	Ventricles
Local WM	Local WM	Local WM	Local WM
	Global Signal	+covary GCOR	Retroicor
			RVT

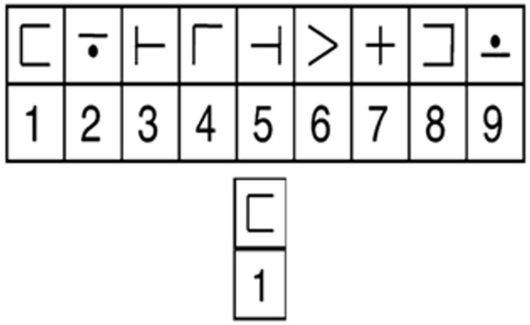


Basic Model
+ GS Regression
+ GCOR
ANATICOR
t-val
(N=60)
TD-ASD
partial r
(N=29)
Correlation
with ASD
SRS scores
ROI (1-27)
ROI (1-27)
ROI (1-27)
ROI (1-27)
ROI (1-27)
ROI (1-27)
4
2
0
-2
-4
0.5
0
-0.5
Nuisance regression:
Motion params
Ventricles
Local WM
Motion params
Ventricles
Local WM
Global Signal
Motion params
Ventricles
Local WM
+covary GCOR
Motion params
Ventricles
Local WM
Retroicor
RVT

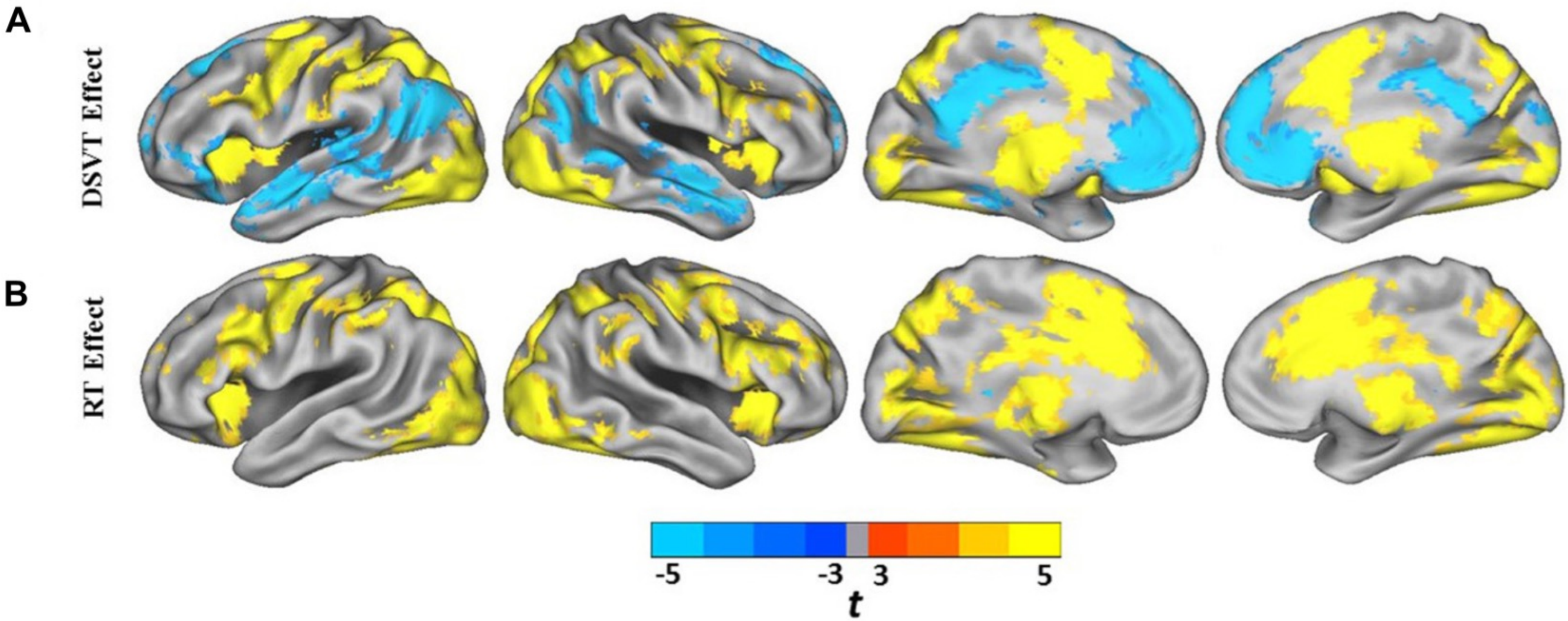
Examining Reliability and Validity Simultaneously: Current Study

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)



Digit Symbol Coding Task



Rao, Motes, & Rypma (2014). Frontiers Hum Neurosci

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)

- slow-event-related fMRI design using overt **picture naming** (and rest)

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)

- slow-event-related fMRI design using overt **picture naming** (and rest)

"tiger"

”car"

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)

- slow-event-related fMRI design using overt **picture naming** (and rest)

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)

- slow-event-related fMRI design using overt **picture naming** (and rest)
- 20 subjects named 100 pictures in two runs (50 per run), with pictures presented every 6.6 to 13.2 seconds

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)

- slow-event-related fMRI design using overt **picture naming** (and rest)
- 20 subjects named 100 pictures in two runs (50 per run), with pictures presented every 6.6 to 13.2 seconds
- response time and accuracy were recorded using a noise-cancelling MRI-compatible microphone

Examining Reliability and Validity Simultaneously: Current Study

Utilizing a well-established task-based phenomenon of trial-to-trial BOLD correlates of response time (uses fluctuations as signal, just like resting-state fMRI): larger BOLD => slower RT (e.g. Yarkoni et al., 2009; Rao, Motes, & Rypma, 2014)

- slow-event-related fMRI design using overt **picture naming** (and rest)
- 20 subjects named 100 pictures in two runs (50 per run), with pictures presented every 6.6 to 13.2 seconds
- response time and accuracy were recorded using a noise-cancelling MRI-compatible microphone
- used a multi-echo fMRI design, with 3 readouts per TR (TR = 2.2 sec, TEs = 12.5 ms, 27.7 ms, 42.9 ms), allowing multi-echo ICA (me-ICA) cleaning

Examining Reliability and Validity Simultaneously: Current Study

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:
1) whole-brain beta weights on run 1 vs run 2

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:

1) whole-brain beta weights on run 1 vs run 2
2) voxelwise task-based FC matrix in run 1 vs run 2

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:
1) whole-brain beta weights on run 1 vs run 2
2) voxelwise task-based FC matrix in run 1 vs run 2

Validity:

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:

1) whole-brain beta weights on run 1 vs run 2
2) voxelwise task-based FC matrix in run 1 vs run 2

Validity:

- Select voxels based on beta weights (t-map of mean trial response vs 0): top 1000, top 2000, etc., to top 10000 voxels

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:
1) whole-brain beta weights on run 1 vs run 2
2) voxelwise task-based FC matrix in run 1 vs run 2

Validity:
- Select voxels based on beta weights (t-map of mean trial response vs 0): top 1000, top 2000, etc., to top 10000 voxels
- Within these voxels, calculate average voxelwise correlation of peak BOLD magnitude per trial (averaging TRs 3 and 4 post-stimulus) and response time (RT) on that trial

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:
1) whole-brain beta weights on run 1 vs run 2
2) voxelwise task-based FC matrix in run 1 vs run 2

Validity:
- Select voxels based on beta weights (t-map of mean trial response vs 0): top 1000, top 2000, etc., to top 10000 voxels
- Within these voxels, calculate average voxelwise correlation of peak BOLD magnitude per trial (averaging TRs 3 and 4 post-stimulus) and response time (RT) on that trial

Main questions:

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:
1) whole-brain beta weights on run 1 vs run 2
2) voxelwise task-based FC matrix in run 1 vs run 2

Validity:
- Select voxels based on beta weights (t-map of mean trial response vs 0): top 1000, top 2000, etc., to top 10000 voxels
- Within these voxels, calculate average voxelwise correlation of peak BOLD magnitude per trial (averaging TRs 3 and 4 post-stimulus) and response time (RT) on that trial

Main questions:
Which preprocessing pipeline gives the best validity?

Examining Reliability and Validity Simultaneously: Current Study

Test-retest reliability:
1) whole-brain beta weights on run 1 vs run 2
2) voxelwise task-based FC matrix in run 1 vs run 2

Validity:
- Select voxels based on beta weights (t-map of mean trial response vs 0): top 1000, top 2000, etc., to top 10000 voxels
- Within these voxels, calculate average voxelwise correlation of peak BOLD magnitude per trial (averaging TRs 3 and 4 post-stimulus) and response time (RT) on that trial

Main questions:
Which preprocessing pipeline gives the best validity?
Does this agree with best test-retest reliability?

Examining Reliability and Validity Simultaneously: Current Study

Examining Reliability and Validity Simultaneously: Current Study

Multi-echo pipelines (all first had 3dTcat, 3dDespike, 3dTshift, 3dvolreg):

1) me-ICA (Kundu et al., 2012)
2) optimally combined (OC) ME data (no nuisance regression)
3) OC + Motion params + Ventricles + local White Matter + GS
4) OC + Motion params + Ventricles + local White Matter
5) OC + ANATICOR (Jo et al., 2010): adding Retroicor and RVT to #4

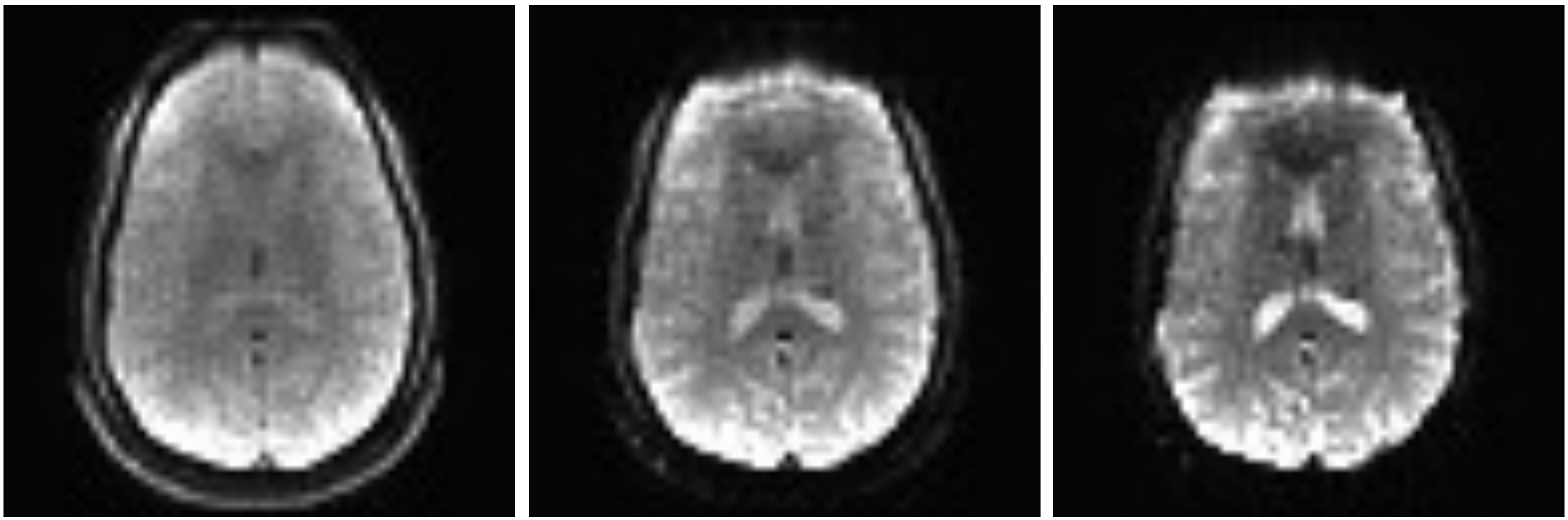
Examining Reliability and Validity Simultaneously: Current Study

Multi-echo pipelines (all first had 3dTcat, 3dDespike, 3dTshift, 3dvolreg):
1) me-ICA (Kundu et al., 2012)
2) optimally combined (OC) ME data (no nuisance regression)
3) OC + Motion params + Ventricles + local White Matter + GS
4) OC + Motion params + Ventricles + local White Matter
5) OC + ANATICOR (Jo et al., 2010): adding Retroicor and RVT to #4

Single-echo pipelines (echo 2):
1) no blurring
2) blur 6mm FWHM
3) blur 6mm + Motion params + Ventricles + local White Matter + GS
4) blur 6mm + Motion params + Ventricles + local White Matter
5) blur 6mm + ANATICOR

Multi-echo EPI data

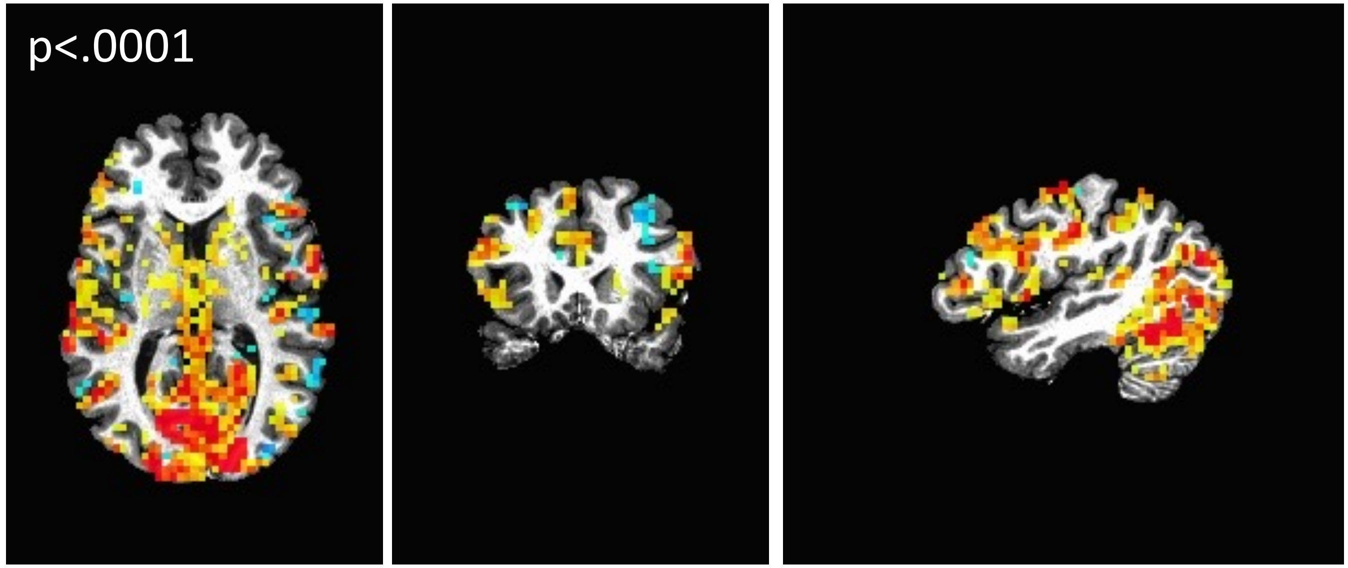
TE = 12.5 ms TE = 27.7 ms TE = 42.9 ms



Beta Weights vs. RT correlations

Beta Weights vs. RT correlations

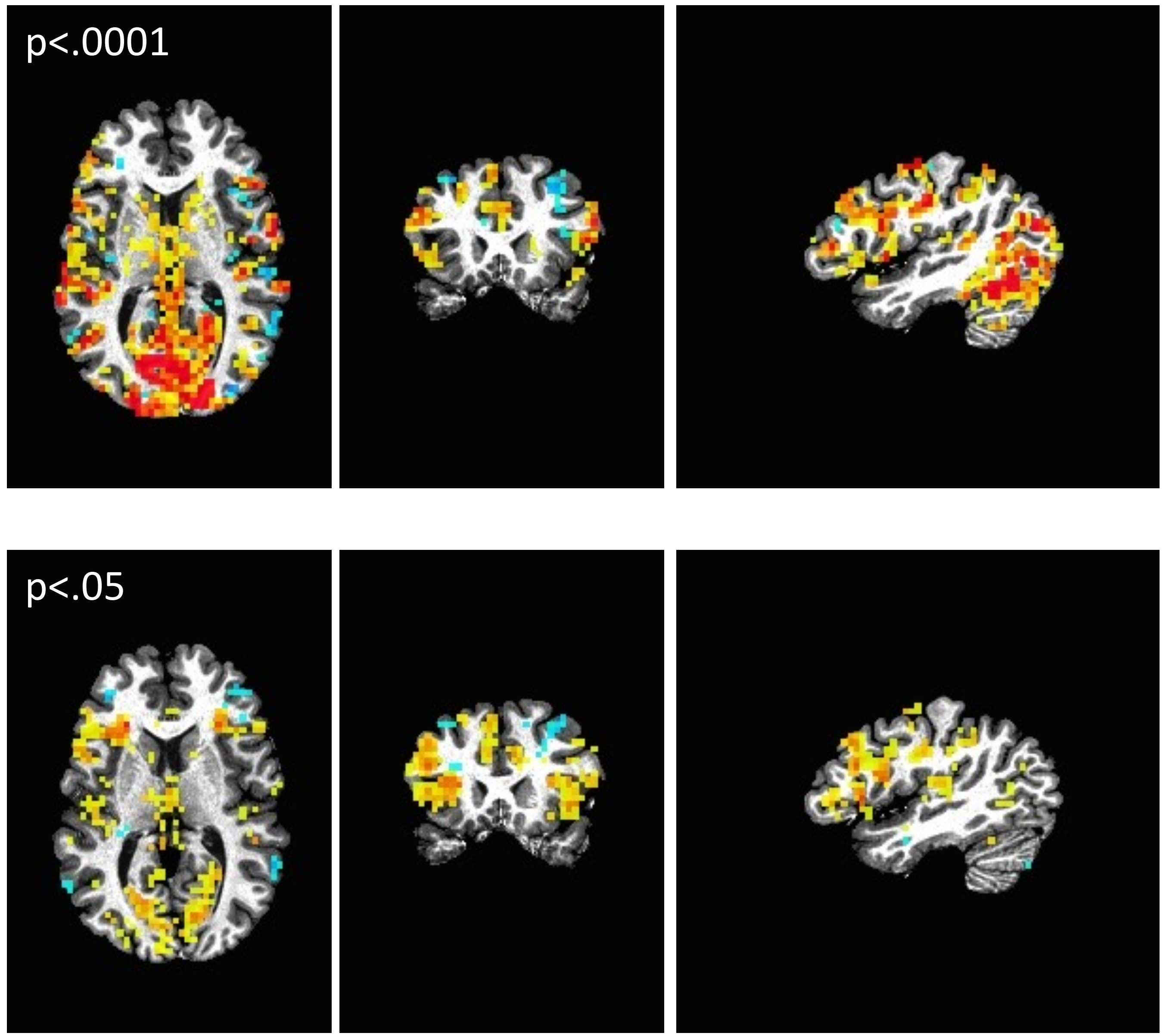
Stimulus vs 0



Beta Weights vs. RT correlations

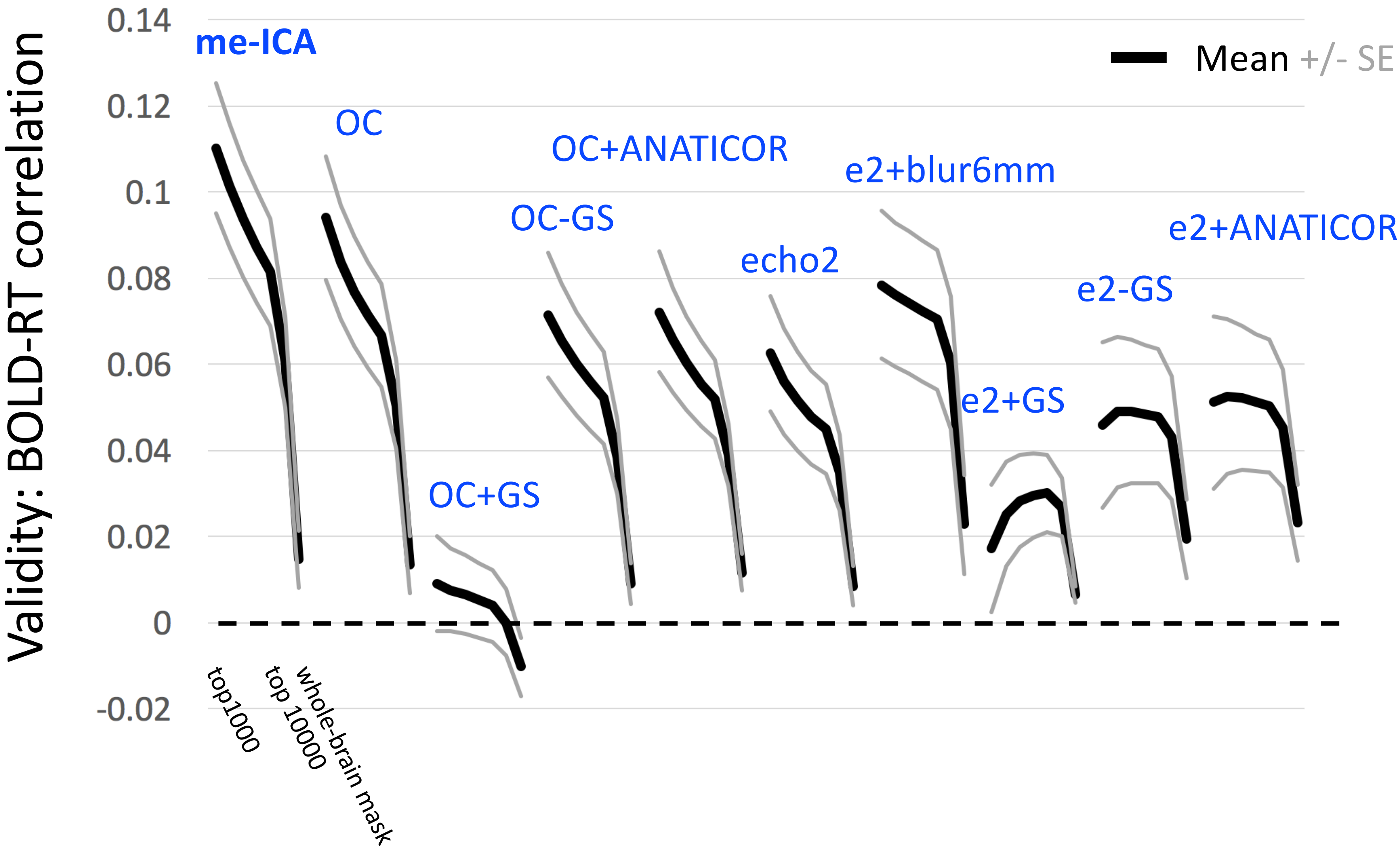
Stimulus vs 0

Correlation of RT and BOLD

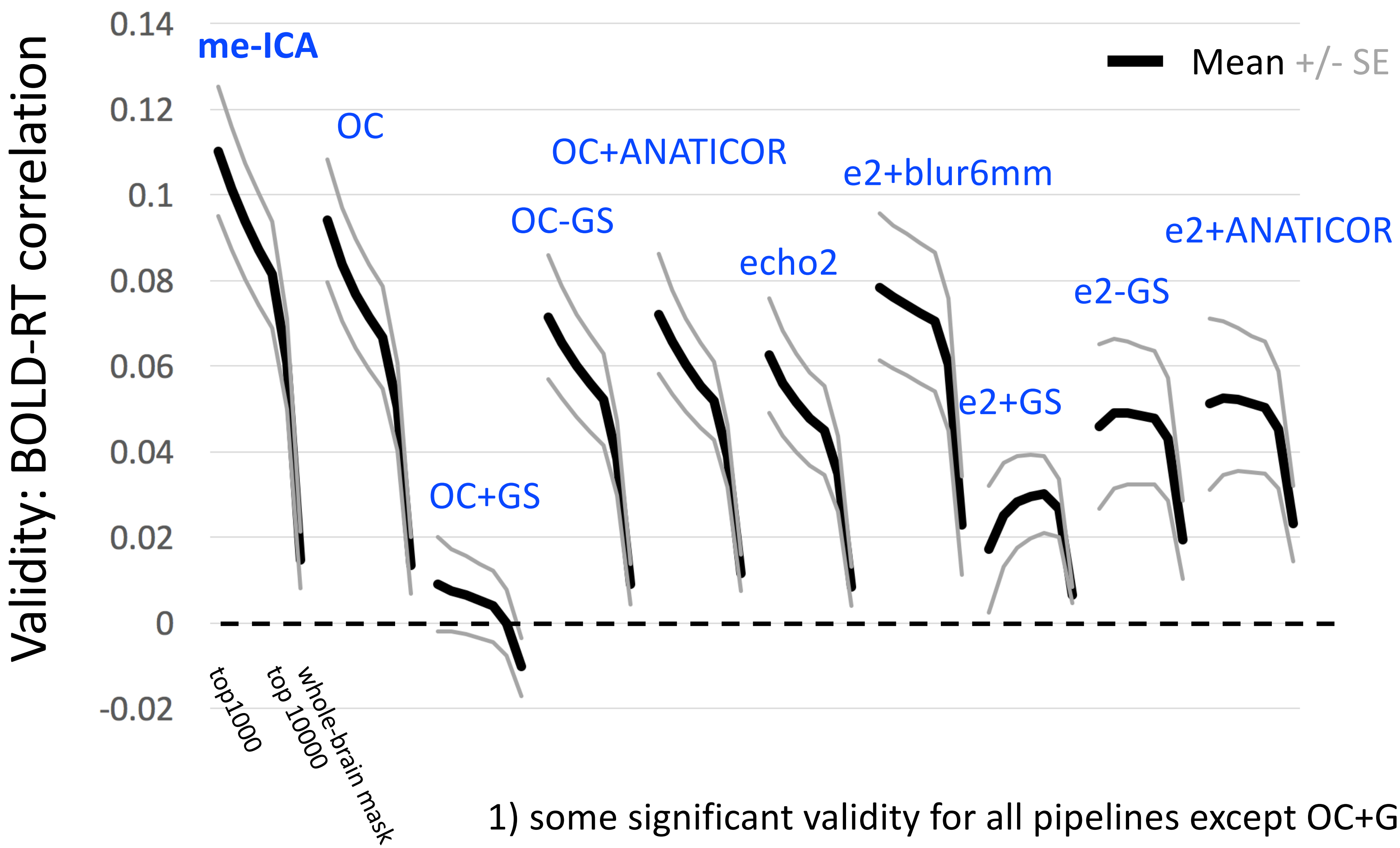


Results: Validity

Results: Validity

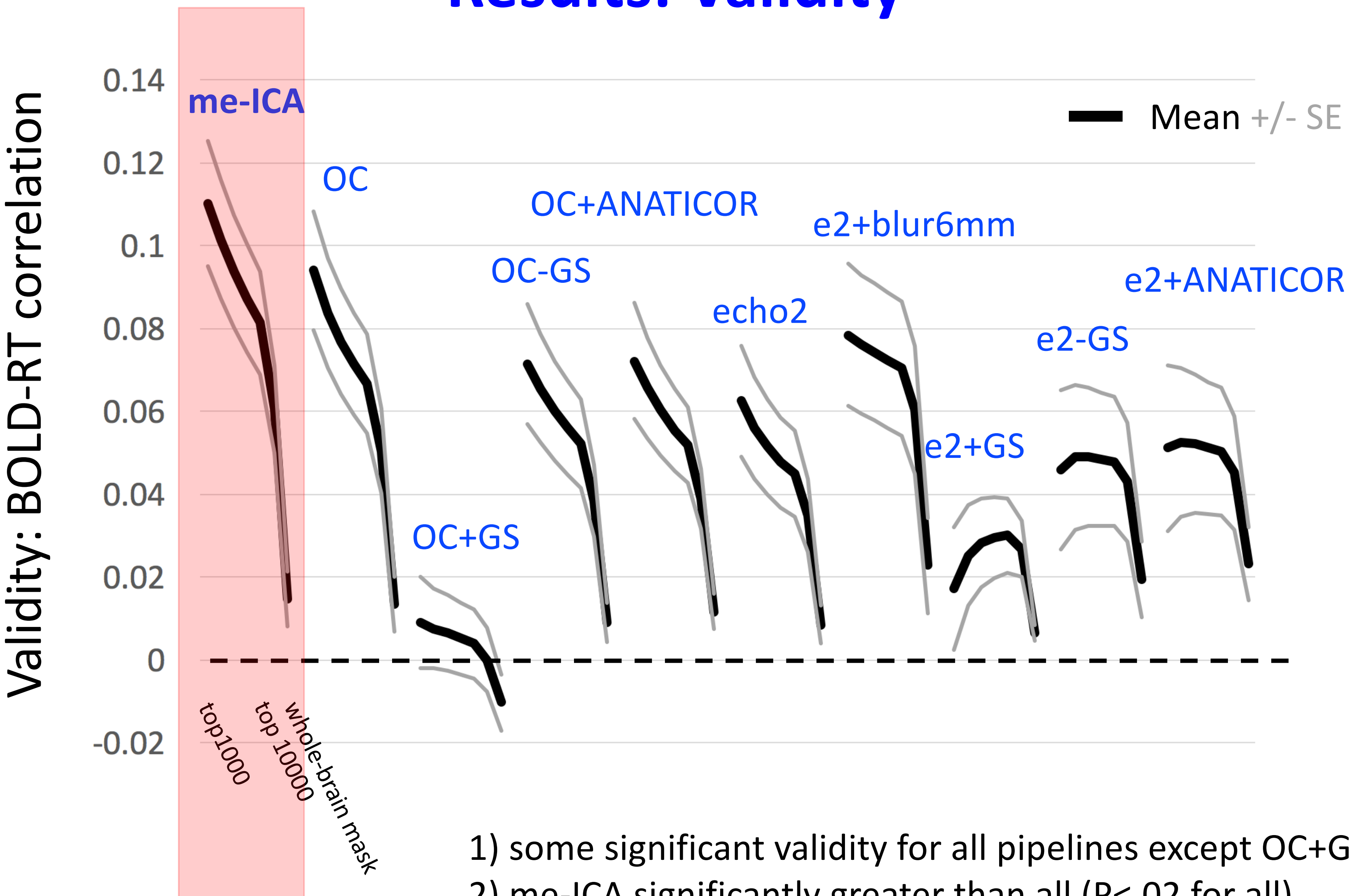


Results: Validity



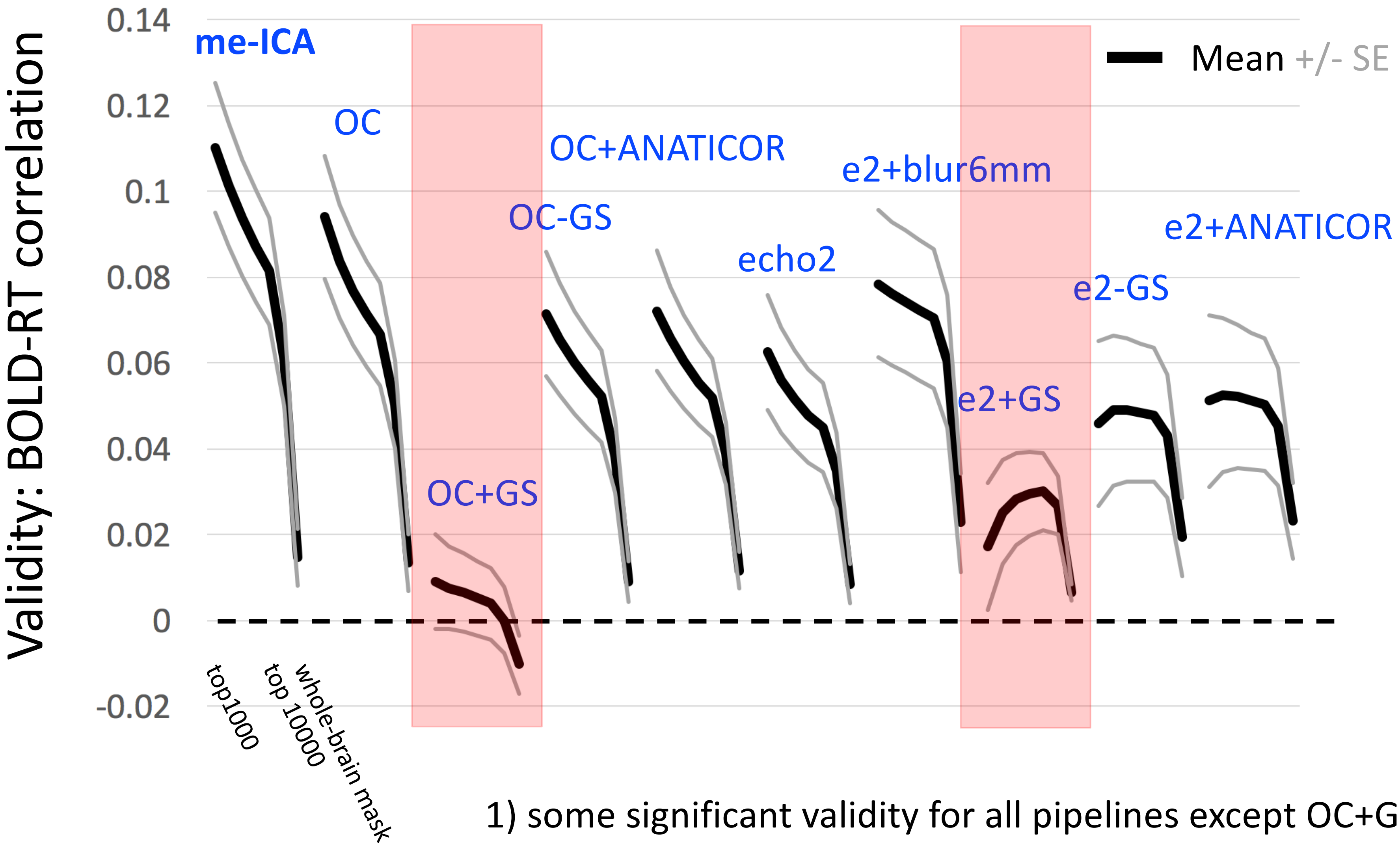
1) some significant validity for all pipelines except OC+GS

Results: Validity



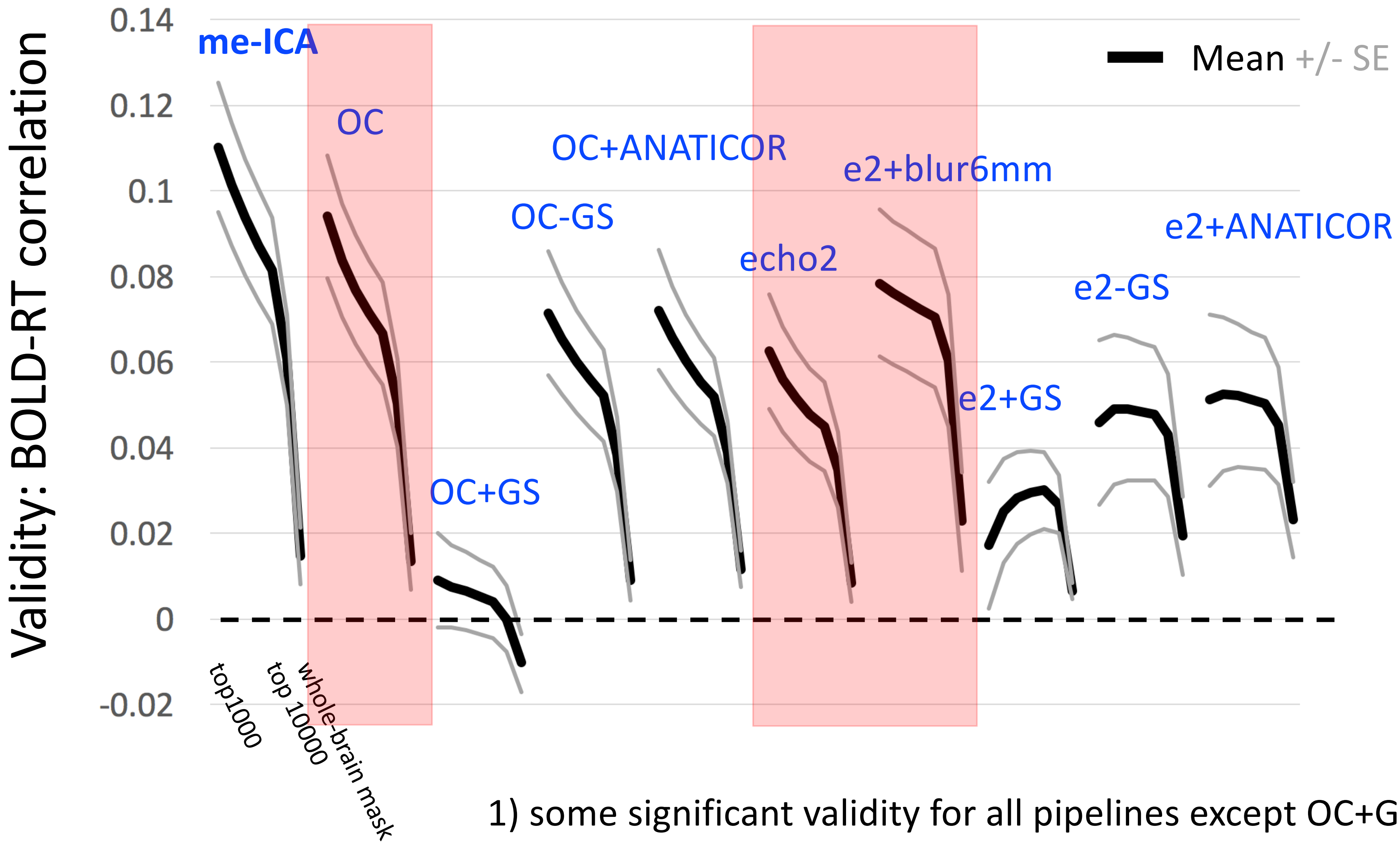
1) some significant validity for all pipelines except OC+GS
2) me-ICA significantly greater than all (P<.02 for all)

Results: Validity



1) some significant validity for all pipelines except OC+GS
2) me-ICA significantly greater than all (P<.02 for all)
3) GSR pipelines (OC+GS, e2+GS) perform the worst

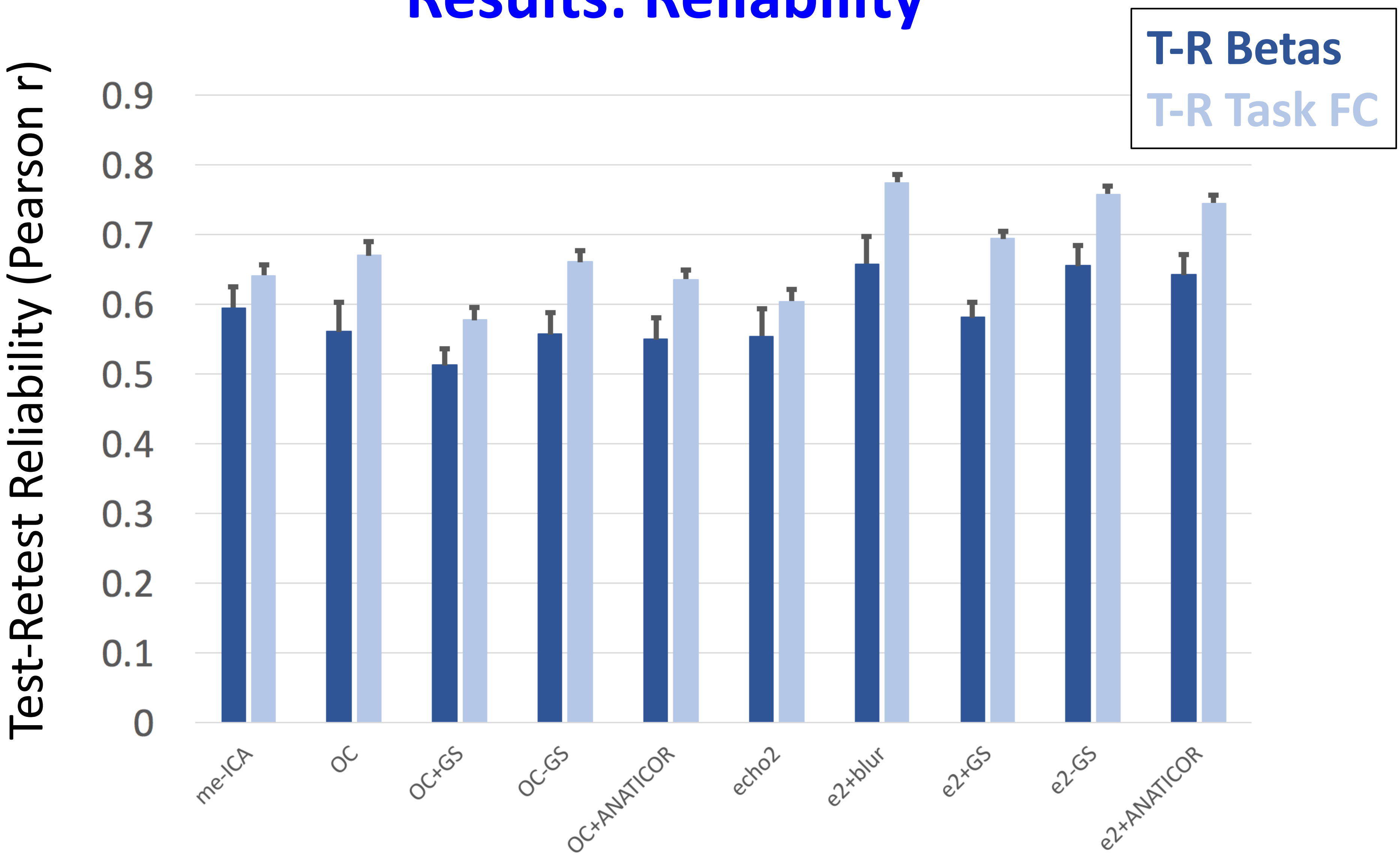
Results: Validity



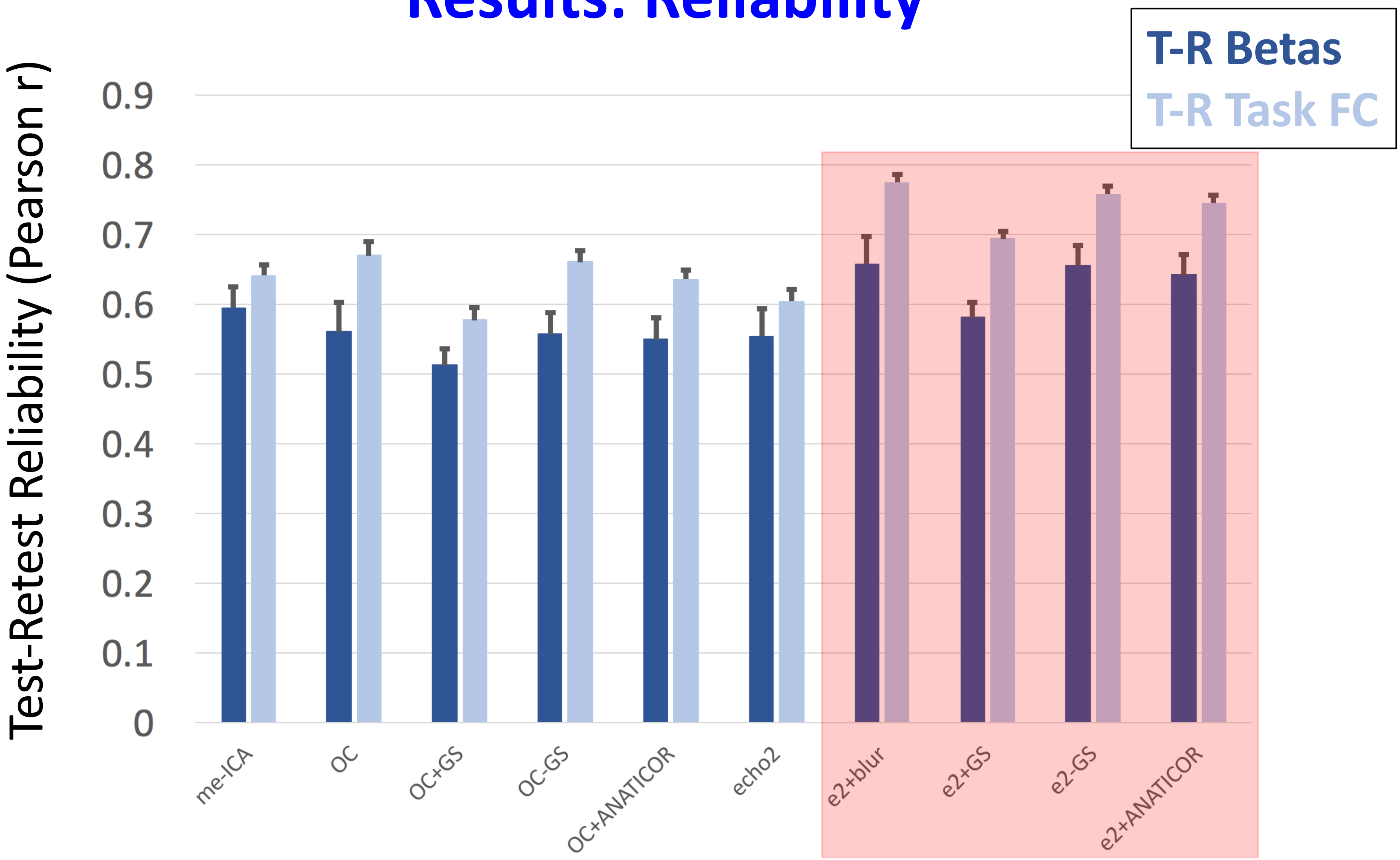
1) some significant validity for all pipelines except OC+GS
2) me-ICA significantly greater than all (P<.02 for all)
3) GSR pipelines (OC+GS, e2+GS) perform the worst
4) nuisance regression is typically worse than no regression

Results: Reliability

Results: Reliability

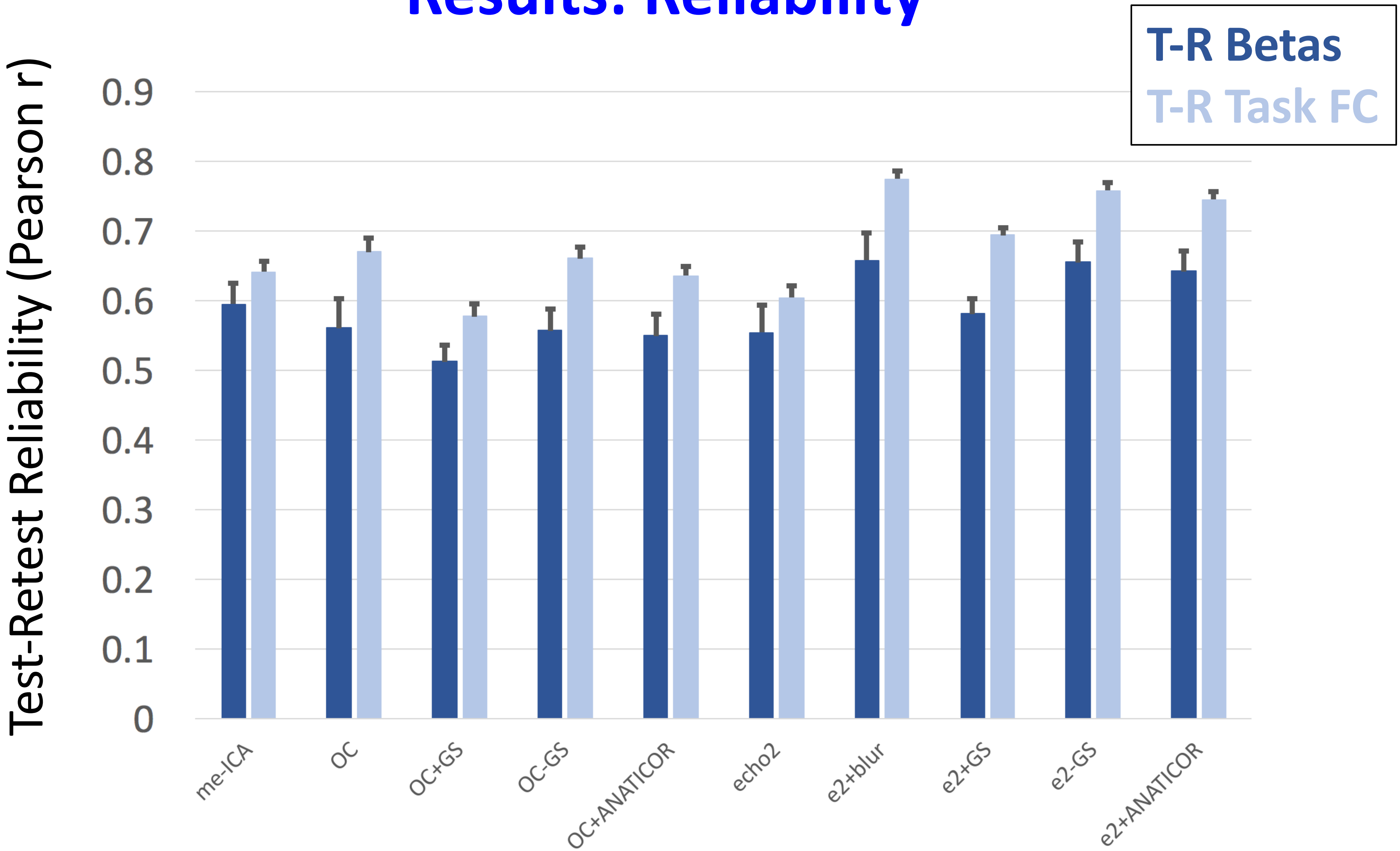


Results: Reliability



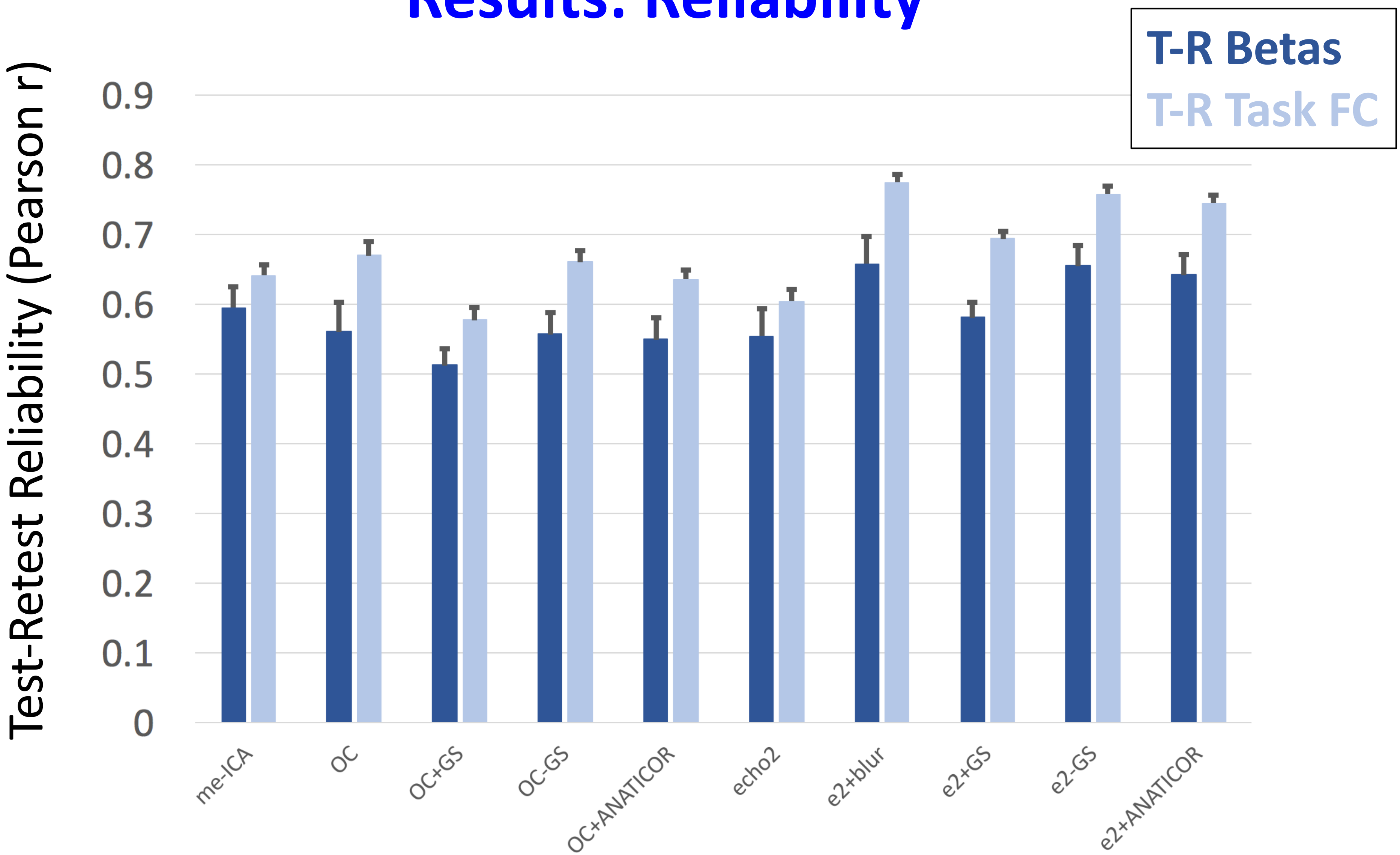
1) Most reliable are echo2 with blurring, not me-ICA

Results: Reliability



1) Most reliable are echo2 with blurring, not me-ICA
2) The two reliability measures are intercorrelated (r = 0.417, p<.0005)

Results: Reliability

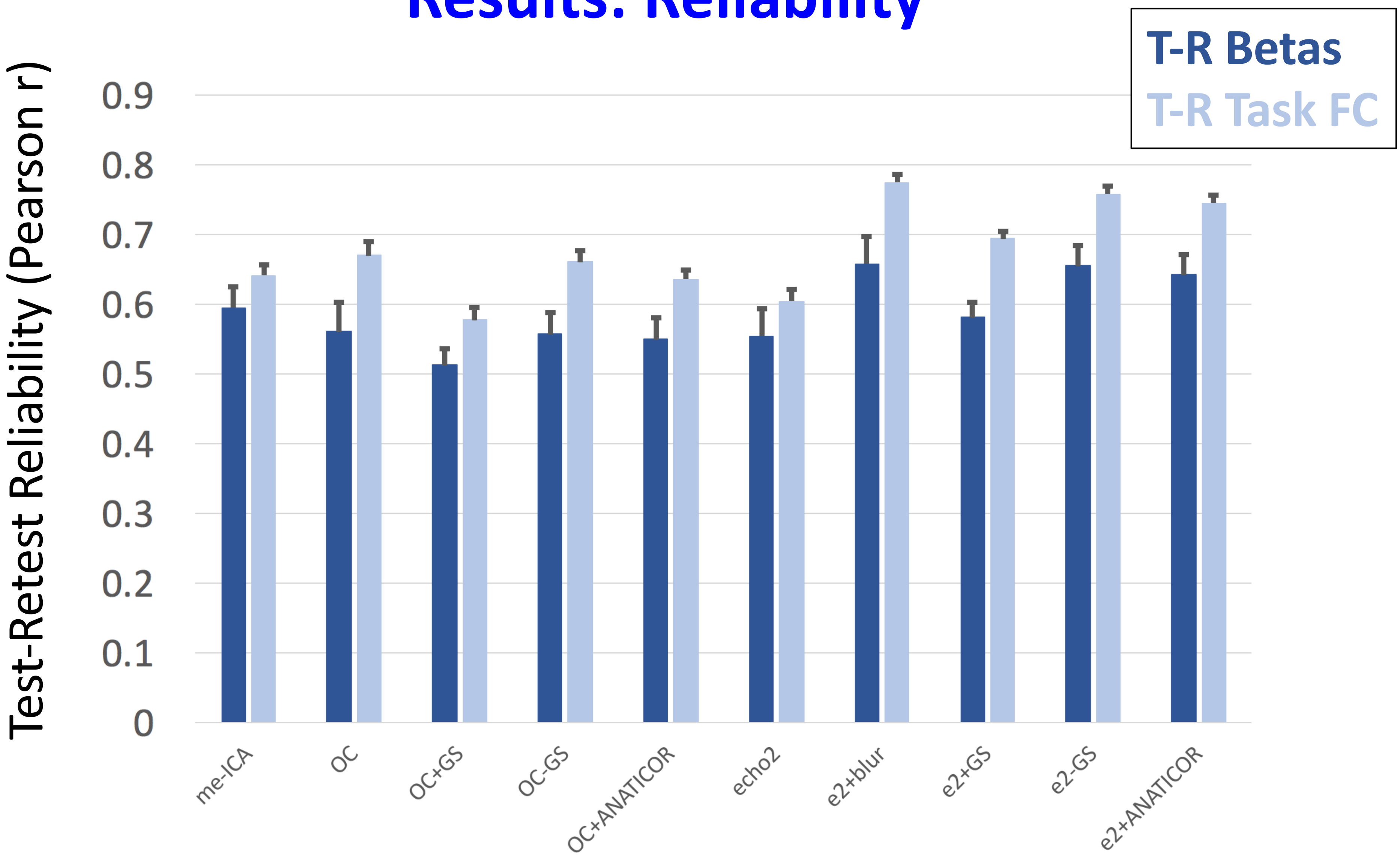


1) Most reliable are echo2 with blurring, not me-ICA

2) The two reliability measures are intercorrelated (r = 0.417, p<.0005)

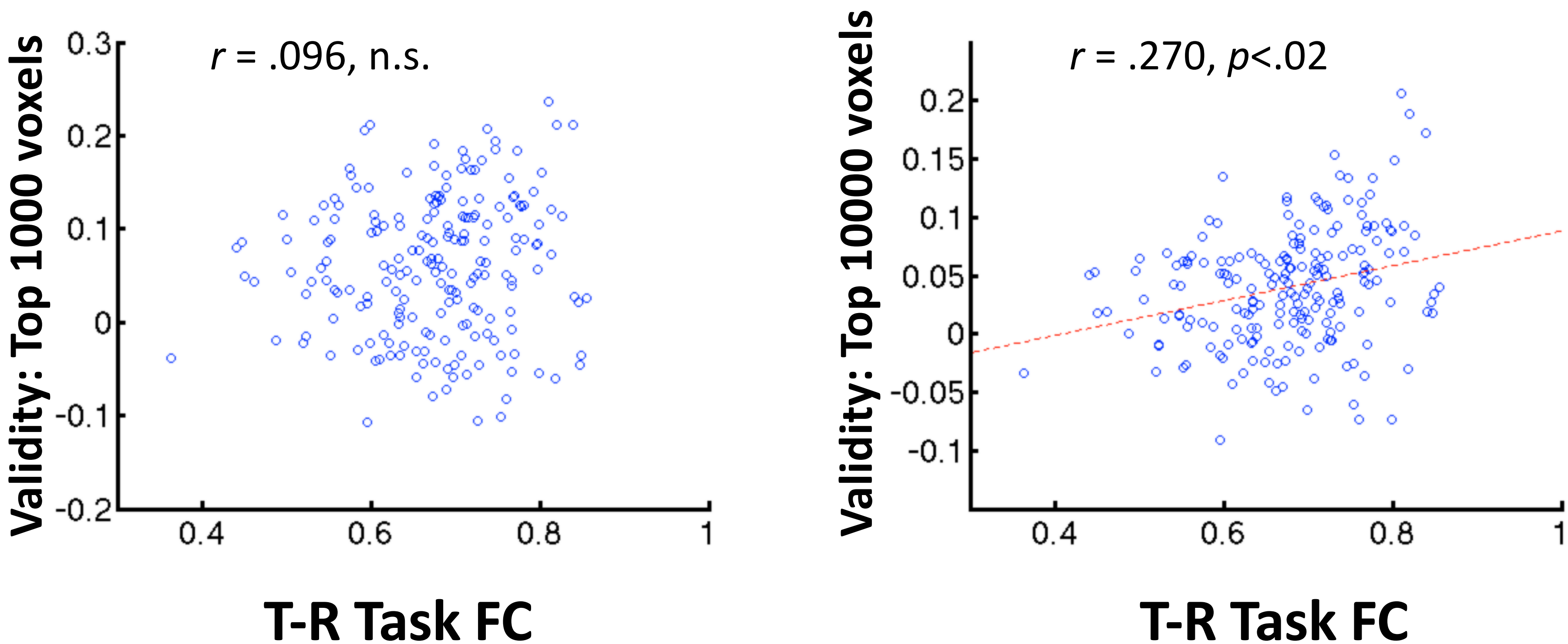
3) Neither reliability measure is related to the best validity (top 1000 voxels): r = .023, .096

Results: Reliability



1) Most reliable are echo2 with blurring, not me-ICA
2) The two reliability measures are intercorrelated (r = 0.417, p<.0005)
3) Neither reliability measure is related to the best validity (top 1000 voxels): r = .023, .096
4) Only task-based FC reliability is related to validity at lower thresholds (r = 0.270, p<.02)

Reliability vs. Validity



Summary

Summary

- Thus far, me-ICA appears to perform the best for validity

Summary

- Thus far, me-ICA appears to perform the best for validity
- While test-retest reliability provides an upper bound on the levels of validity, certain procedures have a pronounced affect on validity that do not appear to be reflected in reliability

Summary

- Thus far, me-ICA appears to perform the best for validity
- While test-retest reliability provides an upper bound on the levels of validity, certain procedures have a pronounced affect on validity that do not appear to be reflected in reliability
- These preliminary results indicate that most of the nuisance regression approaches are removing signal of interest in addition to noise, with a slightly detrimental effect overall

Summary

- Thus far, me-ICA appears to perform the best for validity
- While test-retest reliability provides an upper bound on the levels of validity, certain procedures have a pronounced affect on validity that do not appear to be reflected in reliability
- These preliminary results indicate that most of the nuisance regression approaches are removing signal of interest in addition to noise, with a slightly detrimental effect overall
- We plan on repeating these analyses using rest data for the same participants, correlating average overall response time with resting-state FC (significant correlations exist in me-ICA cleaned data)

Example from Our Lab: Functional Lateralization of Verbal, Visuospatial, and Motor Abilities

Example from Our Lab:
Functional Lateralization of Verbal, Visuospatial, and Motor Abilities

PNAS

PNAS PLUS

Two distinct forms of functional lateralization in the human brain

Stephen J. Gotts[a,1], Hang Joon Jo[b,1,2], Gregory L. Wallace[a], Ziad S. Saad[b], Robert W. Cox[b], and Alex Martin[a]

[a]Section on Cognitive Neuropsychology, Laboratory of Brain and Cognition, and [b]Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892

Edited by Geoffrey K. Aguirre, University of Pennsylvania, Philadelphia, PA, and accepted by the Editorial Board July 25, 2013 (received for review February 8, 2013)

Example from Our Lab:
Functional Lateralization of Verbal, Visuospatial, and Motor Abilities

PNAS

PNAS PLUS

Two distinct forms of functional lateralization in the human brain

Stephen J. Gotts[a,1], Hang Joon Jo[b,1,2], Gregory L. Wallace[a], Ziad S. Saad[b], Robert W. Cox[b], and Alex Martin[a]

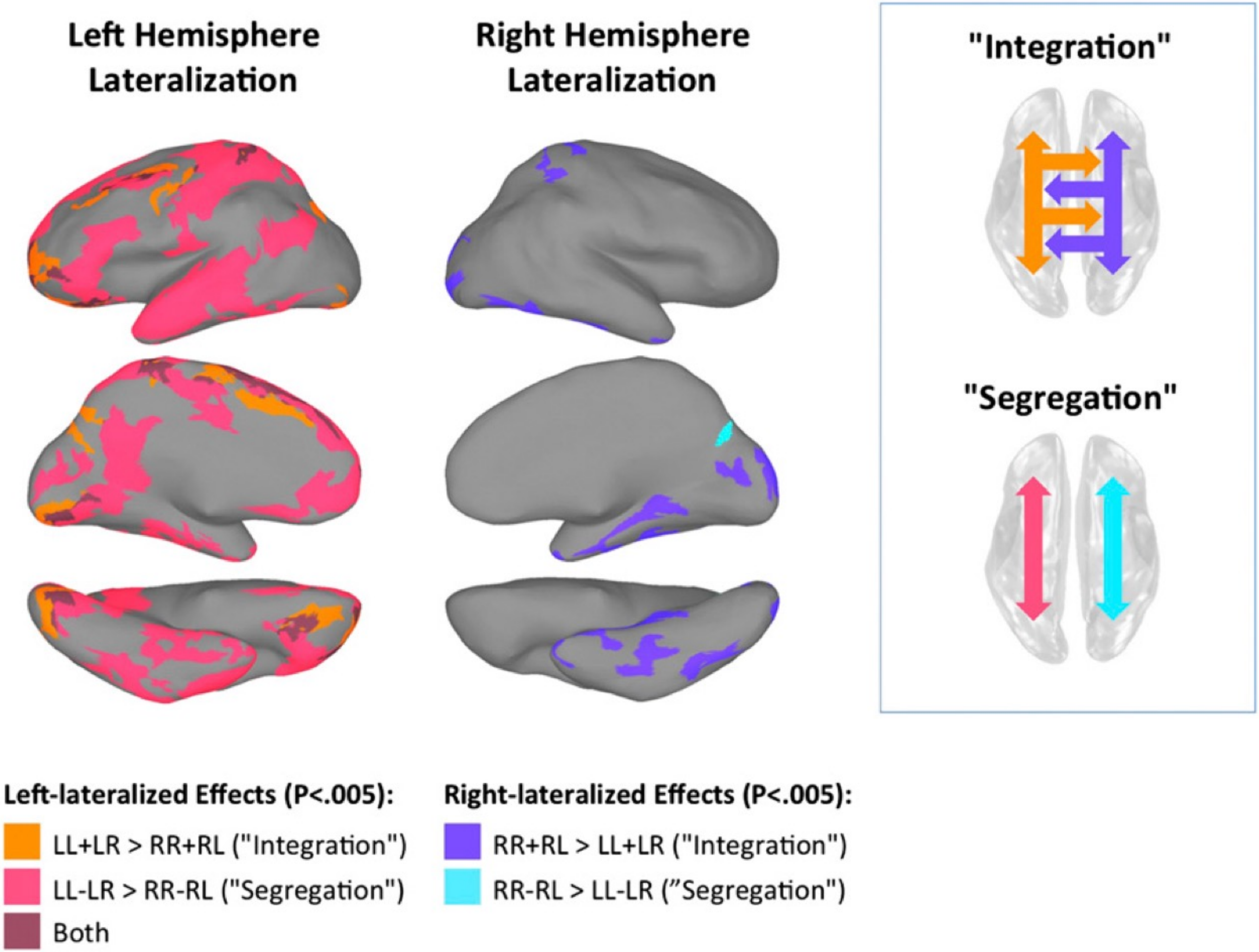
[a]Section on Cognitive Neuropsychology, Laboratory of Brain and Cognition, and [b]Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892

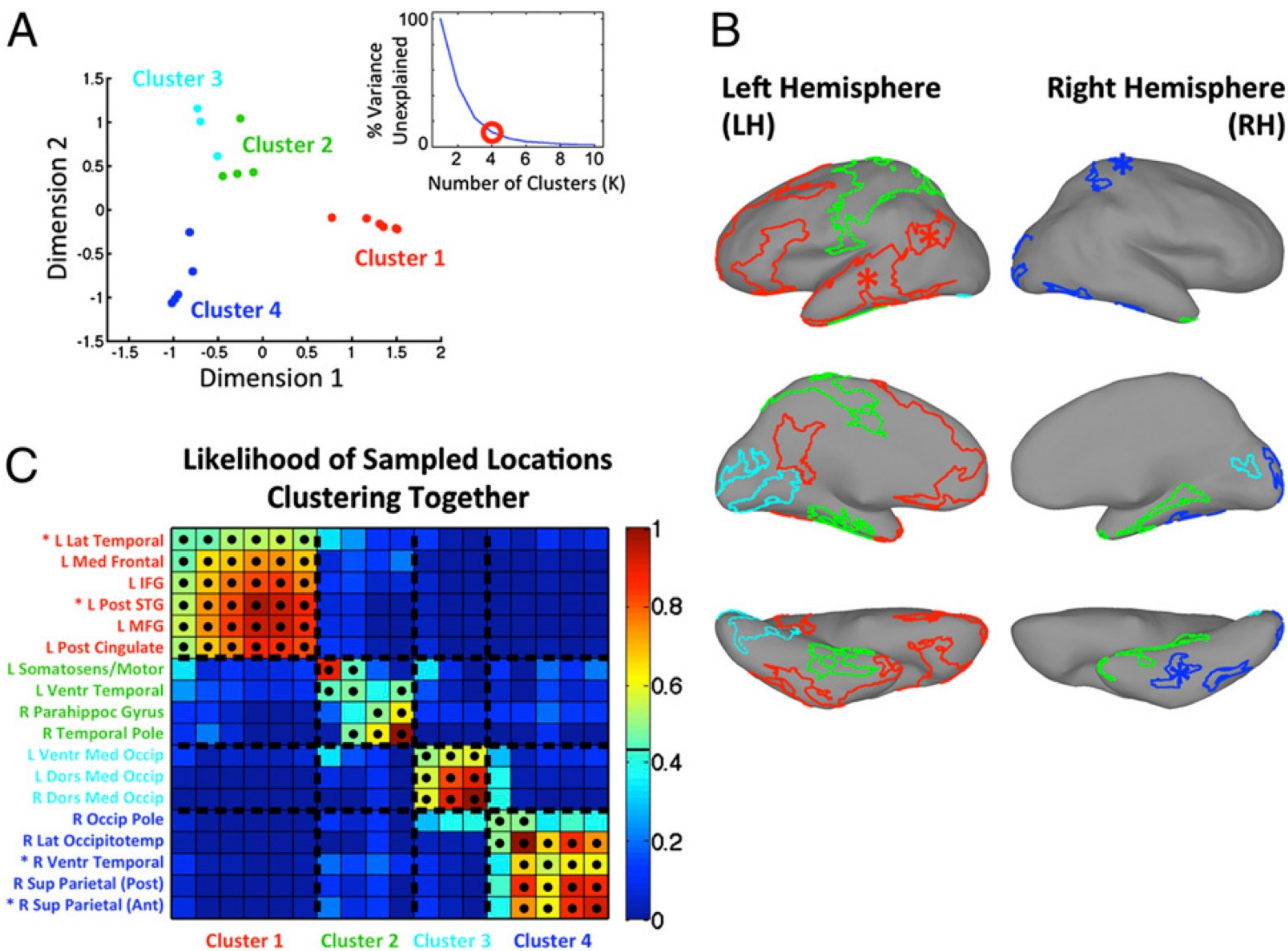
Edited by Geoffrey K. Aguirre, University of Pennsylvania, Philadelphia, PA, and accepted by the Editorial Board July 25, 2013 (received for review February 8, 2013)

Do the hemispheres differ in their within- vs between-hemisphere interactions ?

Does lateralization magnitude predict goodness of function?

Qualitatively Different Forms of Lateralization on Left vs Right




A
B
C
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Dimension 1
Dimension 2
% Variance Unexplained
Number of Clusters (K)
Left Hemisphere (LH)
Right Hemisphere (RH)
Likelihood of Sampled Locations Clustering Together
* L Lat Temporal
L Med Frontal
L IFG
* L Post STG
L MFG
L Post Cingulate
L Somatosens/Motor
L Ventr Temporal
R Parahippoc Gyrus
R Temporal Pole
L Ventr Med Occip
L Dors Med Occip
R Dors Med Occip
R Occip Pole
R Lat Occipitotemp
* R Ventr Temporal
R Sup Parietal (Post)
* R Sup Parietal (Ant)

Lateralization Magnitude Predicts Cognitive Ability