Working memory and reasoning tasks are associated with different modes of large-scale brain dynamics in healthy older adults

Alexander V. Lebedev, Jonna Nilsson, and Martin Lövdén

Aging Research Center, Karolinska Institutet and Stockholm University, Gävlegatan 16, 113 30 Stockholm, Sweden

Address correspondence to:
Alexander V. Lebedev
E-mail: alexander.lebedev@ki.se
Phone: +46 7 659 04 649

Classification: biological sciences

Short title: Brain dynamics of working memory and reasoning

Keywords: reasoning, working memory, modularity, fMRI, cognitive training

Reproducibility statement:
Raw MRI and behavioral data can be requested from the authors and transferred for specific analysis projects that are in line with the original ethical approval. This requires a data use agreement, which effectively transfers the confidentiality obligations of the institution (Karolinska Institutet) at which the original research was conducted to the institution of the recipient of the data. An R-script specifying main analytical and plotting steps is available at: https://github.com/alex-lebedev/RBT1_modularity

Abstract (250 words, actual = 205)

Researchers have proposed that solving complex reasoning problems, a key indicator of fluid intelligence, involves the same cognitive processes as solving working memory tasks. This proposal is supported by an overlap of the functional brain activations associated with the two types of tasks and by high correlations between inter-individual differences in performance. We replicated these findings in fifty-three older subjects but also showed that solving reasoning and working memory problems favors different modes of large-scale neurodynamics and that this dissimilarity increases with higher difficulty load. Specifically, superior performance in a typical working memory paradigm (n-back) was associated with up-regulation of modularity (increased between-network segregation), whereas performance in the reasoning task was associated with effective down-regulation of modularity. We also showed that working memory training promotes task-invariant increases in modularity. Since superior reasoning performance is associated with down-regulation of modular dynamics, training may thus have fostered an inefficient way of solving the reasoning tasks. This could help explain why working memory training does little to promote complex reasoning performance. The study concludes that complex reasoning abilities cannot be reduced to working memory and suggests the need to reconsider the feasibility of using working memory training interventions to attempt to achieve effects that transfer to broader cognition.

Significance statement (120 words, actual = 117)

Reasoning ability (fluid intelligence) is important to many real-life outcomes. Understanding the cognitive and neural processes involved in reasoning is therefore of great interest. Some have suggested that reasoning is structurally similar to working memory (WM). As expected, we found that performance in reasoning and WM tasks was highly correlated and brain activation patterns overlapped. However, complex reasoning performance was better when brain networks were interacting in a more globally integrated way, whereas the opposite was true of WM. WM training increased the level of between-network segregation in ways that did not vary by task, which could be one reason why WM training did little to promote complex reasoning performance. Complex reasoning cannot be reduced to WM.

Fluid intelligence is central to solving complex reasoning tasks; that is, to identifying common patterns and applying logic to extrapolate them to novel problems. The cognitive processes that determine performance in such tasks have been a matter of debate for decades (1-5). Task analyses have led some researchers to propose that reasoning ability depends on and is almost isomorphic with the capacity to maintain and manipulate information in working memory (5, 6). In keeping with this line of reasoning, psychologists have emphasized the demands that both working memory and reasoning tasks place on executive control (7, 8). The high correlation between inter-individual differences in working memory and complex reasoning scores (4, 5) and observations that these tasks elicit similar functional brain activity in the fronto-parietal control network (CN) (8-14) are also consistent with this idea. Yet there are reasons to believe that cognitive processes other than working memory ability are involved in complex reasoning. Among them are exploratory cognitive functions such as perspective shifting and aspects of creativity that may facilitate rule discovery and insight. The latter refers to the spontaneous realization that a difficult problem has a simple solution if perceived from a conceptually different or new perspective. While these processes are much less studied, their importance in solving complex reasoning tasks has been demonstrated in psychometric studies (15, 16).

Functional brain imaging studies of large-scale brain dynamics allow us to disentangle some of the key principles underlying higher cognition. Shaped by evolution to enhance adaptability and efficient information processing, the brain appears to function as a critical system achieving optimal performance near phase transitions from order to disorder (17). This, in turn, requires a hierarchical modular structure balancing segregation and integration of its functional subdivisions (18-20). Modularity has been used as a global measure of such neurodynamics in studies of age-related changes (21, 22), working memory (23, 24), learning (25, 26), reactivity to environmental irregularities (27) and visual awareness (28). While task performance is generally associated with lower

modularity than resting state (29-31), superior performance on cognitive tasks requiring *outward* attentional focus (e.g., working memory tasks) has been linked to relatively more modular (i.e., segregated) brain dynamics at rest (23), whereas tasks requiring *inward* attentional focus, such as odor recognition and episodic memory retrieval, are instead linked to low-modular (i.e., more globally integrated) neurodynamics (32, 33). Relatedly, effective suppression of the default mode network (DMN; functionally integrated group of mostly midline brain structures involved in self-referential processing and other tasks requiring inward focus) is related to performance on problems requiring outward attentional focus (34, 35). An opposite association has been reported for more complex tasks that demand creativity, cognitive flexibility, and, importantly, complex reasoning, in which resting state and task-related connectivity between the DMN and CN have been linked to superior performance (13, 36, 37). Furthermore, the direction of the association between modularity and performance may depend on task complexity: performance on simple tasks is positively associated with modularity, whereas performance on more complex tasks, in contrast, appears to be positively associated with low-modular states of neurodynamics (38).

These disparate findings lead to the hypothesis that, despite a high correlation between performance and convergence of brain activation patterns, reasoning tasks require different, or at least additional, cognitive processes and different modes of large-scale brain dynamics than working memory. Better complex reasoning performance may be linked to a low-modular state of neurodynamics characterized by enhanced communication of the CN with the DMN and other brain networks. This enhanced communication may in turn promote an exploratory mode of cognition (39) – a state of "expected uncertainty" (40) – beneficial to rule discovery and abstraction (41). In contrast, successful solution of working memory tasks may require a more modular network configuration with stabilized dynamics of the fronto-parietal network coupled with

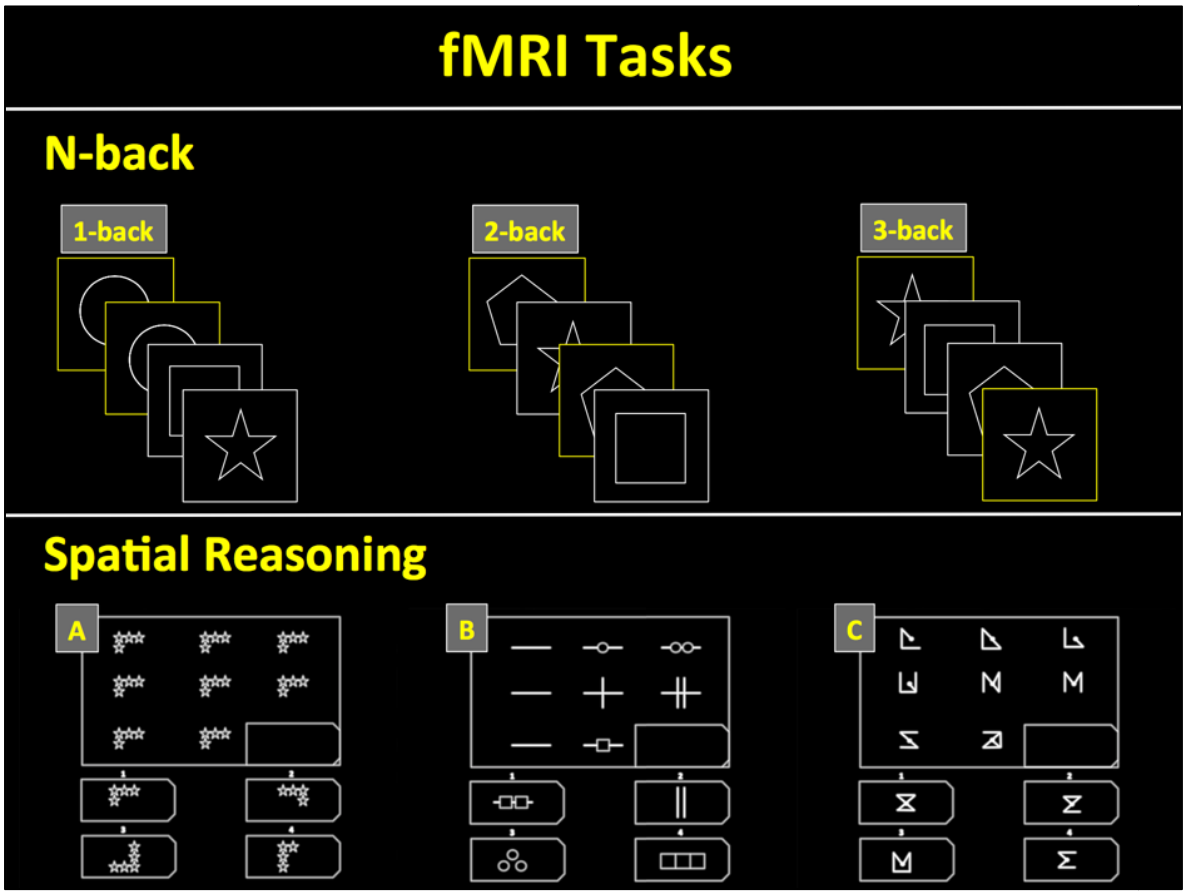
effective suppression of the DMN activity to prevent inward attentional focus and mind-wandering. These divergent associations may be particularly marked at higher difficulty loads, because complex reasoning requires a more exploratory mode of cognition than simpler reasoning, one that benefits pattern discovery. It follows from these hypotheses that improvements from training working memory should lead to more modular brain dynamics during task performance. If such a shift in modularity generalizes to solving the reasoning task then this may only be beneficial for simple reasoning performance but not for solving the complex reasoning tasks that may require a higher degree of between-network integration. We employed large-scale network analysis to investigate these hypotheses. The study used functional magnetic resonance imaging (fMRI) during one reasoning and working memory task. The spatial reasoning task was developed to be similar to the classic matrix reasoning tests and a classical figural updating task (n-back) was used to probe working memory. Both tasks were performed at three different difficulties (**Figure 1**) in a sample of healthy older adults (*n* = 53; age = 65-75 years) before (i.e., at pretest) and after (i.e., at posttest) working memory training (n = 27) and active control training (perceptual matching; n = 26).

Results

Cognitive measures. Initial analyses aimed to probe the reliability and validity of the cognitive tasks. Between-person differences in performance on the in-scanner tasks (**Figure 1**) were relatively stable from before to after cognitive training, with pre-post correlations of r=0.57 (p<0.001) for n-back and r=0.6 (p<0.001) for total reasoning performances averaged across the difficulty levels of the tasks. Corresponding values were equivalent between the groups: $r_{pre\text{-}post}$=0.54/0.47 (z=0.32, p=0.75) in control/active groups for n-back and $r_{pre\text{-}post}$=0.57/0.65 (z=-0.44, p=0.66) in control/active groups for the reasoning task. These results indicate satisfactory lower bounds of reliability.

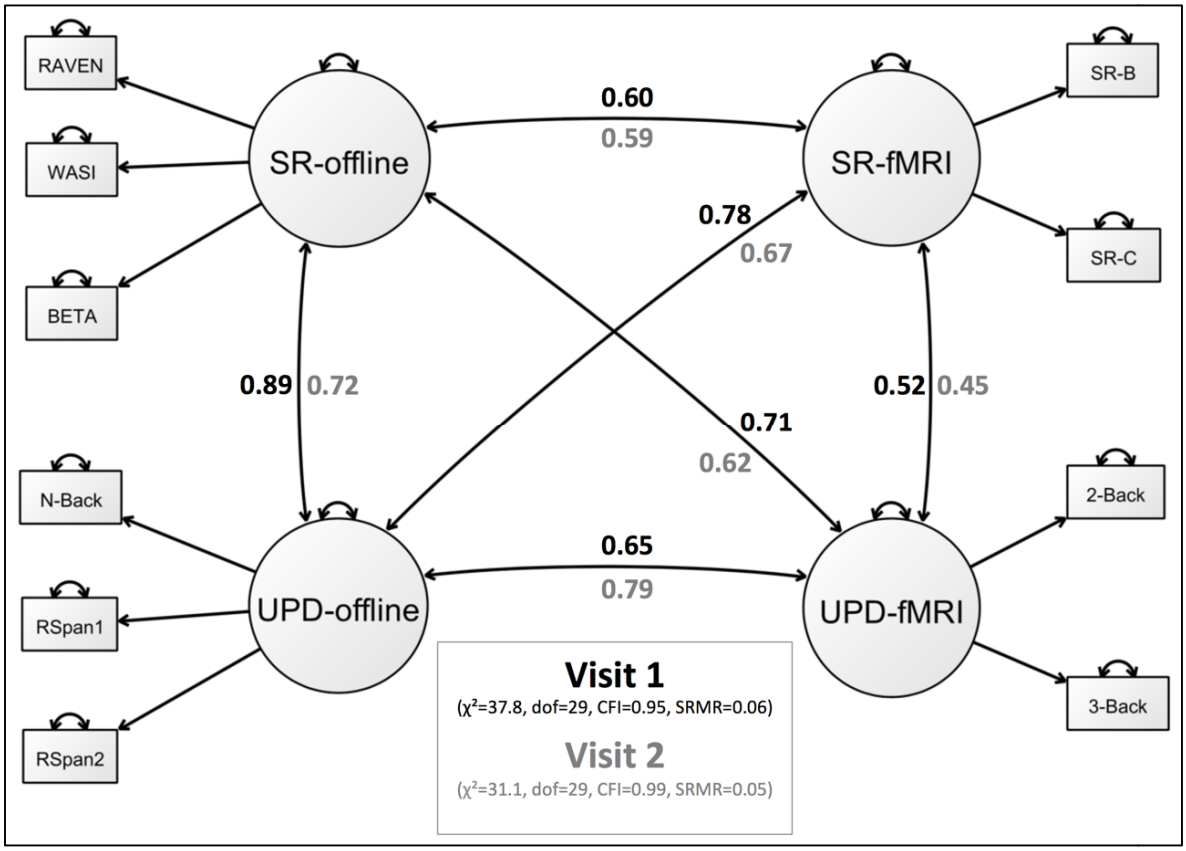
The validity of the in-scanner tasks was evaluated with structural equation modeling (see **Methods**). The n-back task performed in the scanner (UPD-fMRI) correlated strongly (pre: r = 0.65 p<0.001; post: r = 0.79, p<0.001) with a latent factor of working-memory updating (formed by tasks assessed outside the scanner; see **Figure 2** and **Methods** for a more detailed description of the testing battery). Performance on reasoning task administered in the scanner also correlated strongly (pre: r = 0.6, p<0.001; post: r = 59, p<0.001) with a reasoning factor (formed by three tasks assessed outside the scanner). Cross-correlations were equivalent. Performance on the in-scanner n-back correlated in a statistically significant way with the reasoning factor (pre: r = 0.71, p <0.001; post: r = 0.67, p <0.001), and performance on the in-scanner reasoning task, with the working-memory updating factor (pre: r = 0.78, p<0.001; post: r = 0.67, p<0.001). The reasoning and working-memory updating factors were highly correlated (pre: r = 0.89, p<0.001; post: r = 0.72, p<0.001), and so were the in-scanner tasks (pre: r = 0.52, p<0.001; post: r = 0.45, p<0.001). Overall, the reliability as well as discriminant and convergent validity of the tasks used in the scanner were thus satisfactory.

Fig. 1. fMRI tasks.



Two tasks were administered inside the MR camera to evaluate the brain dynamics underlying working memory and reasoning. A figural n-back task assessed working memory performance. Three 1-minute blocks were administered: 1-, 2- and 3-back (each repeated 3 times) with a 4-second fixation screen indicating the rule for the forthcoming level. The second task was developed to evaluate the neurodynamics involved in complex visuospatial reasoning. Subjects were instructed to complete the presented pattern by selecting one of the four solutions shown at the bottom of the screen. One of the four answers was always correct. Similar to the n-back task, this paradigm had three levels of difficulty: A, perceptual matching; B, simple analytical (single rule); and C, difficult analytical (with at least two non-trivial rules) tasks. Each block was repeated 3 times in a pseudo-random order during the scanning session. The reasoning task was self-paced (to better match typical fluid intelligence paradigms such as Raven's progressive matrices) within a fixed (1-minute) block. Thus, different subjects completed different numbers of items per block, and the subjects were explicitly instructed not to guess. For the reasoning paradigm, accuracy was defined as the fraction of correctly solved problems relative to the total number of attempts per block.

Fig. 2. Results of the structural equation modeling analysis.



Circles represent modeled latent variables; rectangles, measured variables (indicators); omnidirectional arrows, dependence; and bidirectional arrows, correlation. SR = spatial reasoning; UPD = updating; SR-B,C = spatial reasoning (fMRI) blocks "B" and "C." Number next to the bidirectional arrows represent standardized covariances (i.e., correlations) between the latent factors at pretest (black) and posttest (grey).

Task-related BOLD activity and modularity at baseline. Both tasks produced similar and reliable load-related activations in the control network (see **Figure 3**) that are typical for working memory and reasoning tasks. Moreover, similarity (voxel-wise Pearson r, see **Methods** for details) of individual activation patterns to the meta-analytical map derived using the Neurosynth package (see **Figure S1**) for "executive control" was statistically significant across all loads (significantly larger than 0 at $p<0.001$) and positively correlated with subjects' performance during n-back (2-back: $r=0.34, p=0.01$; 3-back: $r=0.43, p=0.001$) and simple analytical reasoning tasks (reasoning-B: $r=0.30, p=0.03$; reasoning-C: $r=0.19, p=0.18$). These results indicate that both paradigms reliably engaged the executive control

circuit and that the magnitude of the engagement significantly correlated with successful task solving in all active conditions except for the difficult analytical block ("C") of the reasoning task, in which engagement of the circuit had a limited non-significant effect on performance.

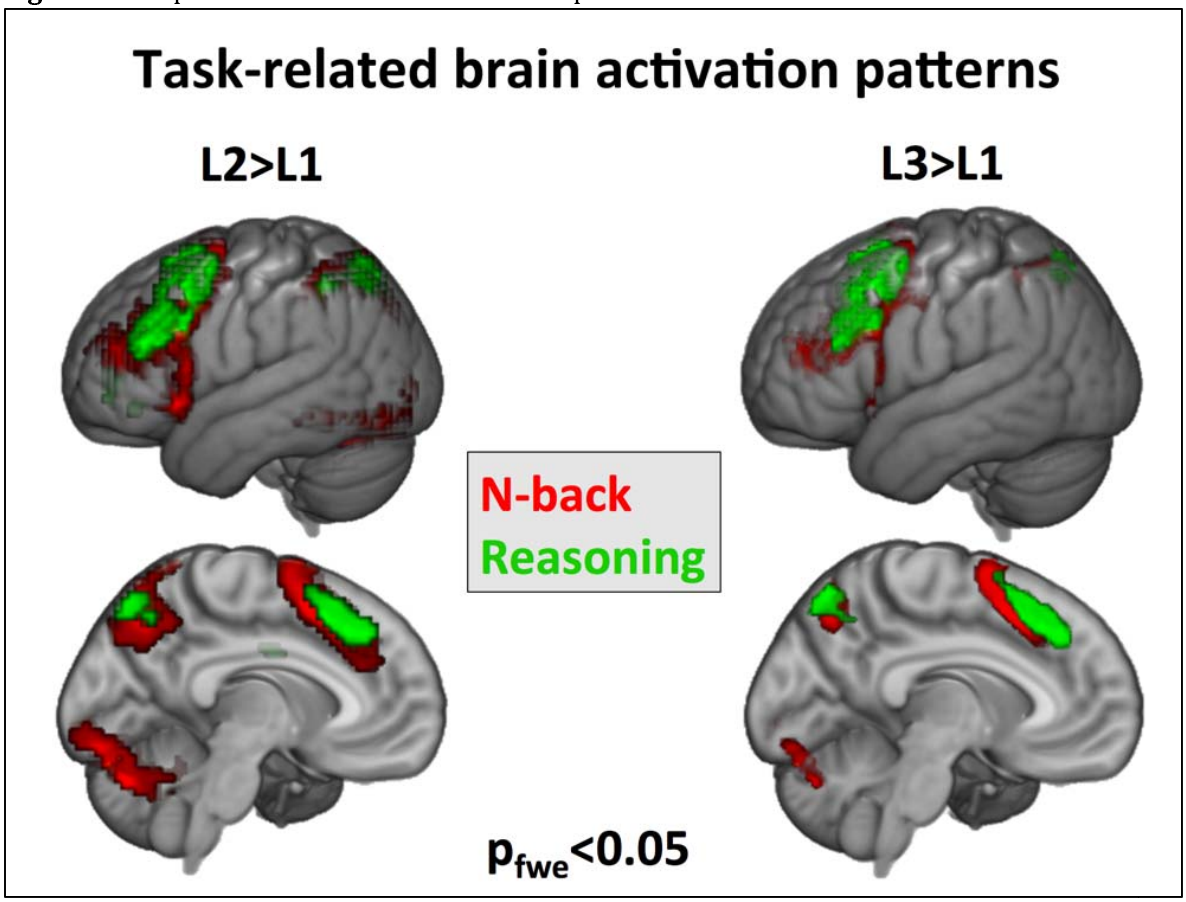
Despite similarities in the activation patterns, we observed a statistically significant effect of load, $t(255)=-2.6$, $p=0.01$ and load-by-task effect, $t(255)=2.67$, $p=0.008$ on modularity (Q). These effects were further qualified by a load-by-task-by-performance effect, $t(255)=2.78$, $p=0.006$. Although performance was implemented as a continuous variable in these analyses, to visualize these effects in an accessible manner we performed a median split on performance (**Figure 4).** The resulting graph shows that individuals who performed well on the n-back task had a tendency to increase modularity (i.e., between-network segregation) with load, whereas individuals with better performance in the more complex reasoning task down-regulated modularity with task complexity**.** In other words, superior performance in the difficult working memory task (n-back) was associated with higher modularity (between-network segregation), whereas superior performance in the complex reasoning task was associated with lower modularity (see also **Figure 5A**). To further assess these effects, we selected a subset of subjects (n=17) who performed high on both tests, defined as above-median performance both in n-back and reasoning. Load-by-task effect was significant in this group, $t(82)=3.28$, $p=0.001$, whereas subjects performing poorly on both tasks (n=18) did not exhibit such pattern, $t(87)=3.28$, $p=0.001$. Thus, individual differences in performance were, as predicted, related to different (segregated vs. integrated) patterns of large-scale dynamics in the working memory and complex reasoning tasks, especially at higher loads.

To identify which of the networks and between-network interactions drive these relations with global modularity, we conducted a number of subsequent tests of raw functional connectivity (42). The results presented a similar pattern of findings, which

suggest that superior n-back performance is linked to increased segregation between the control and the DMN networks, whereas superior performance in the complex reasoning task is linked with more active cross-talk of the control network with the sensory community and the DMN (**Figure 5B**).

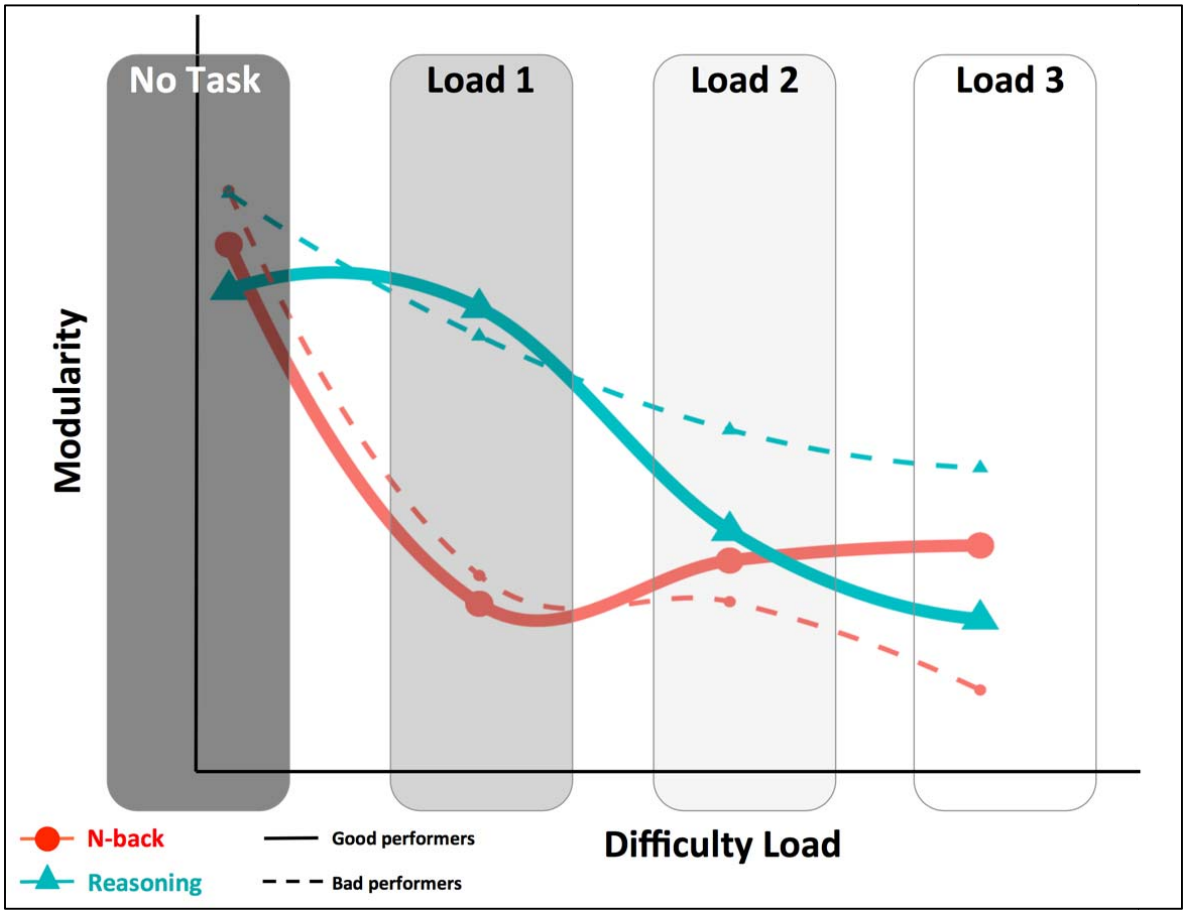
Note that modularity during the fixation condition did not differ in a statistically significant way between the tasks, $t(52)=0.34$, $p=0.736$. Moreover, both tasks induced prominently less modular dynamics than the fixation condition: N-back, $t(52)=6.2$, $p<0.001$; reasoning, $t(52)=7.0$, $p<0.001$ (**Figure 4**). Yet, all difficulty loads were consistently associated with modular neurodynamics: mean Q was significantly and consistently larger than 0 ($p<0.001$), both as a total effect and when taken separately within specific tasks and loads.

Fig. 3. Overlap of the task-related activation patterns



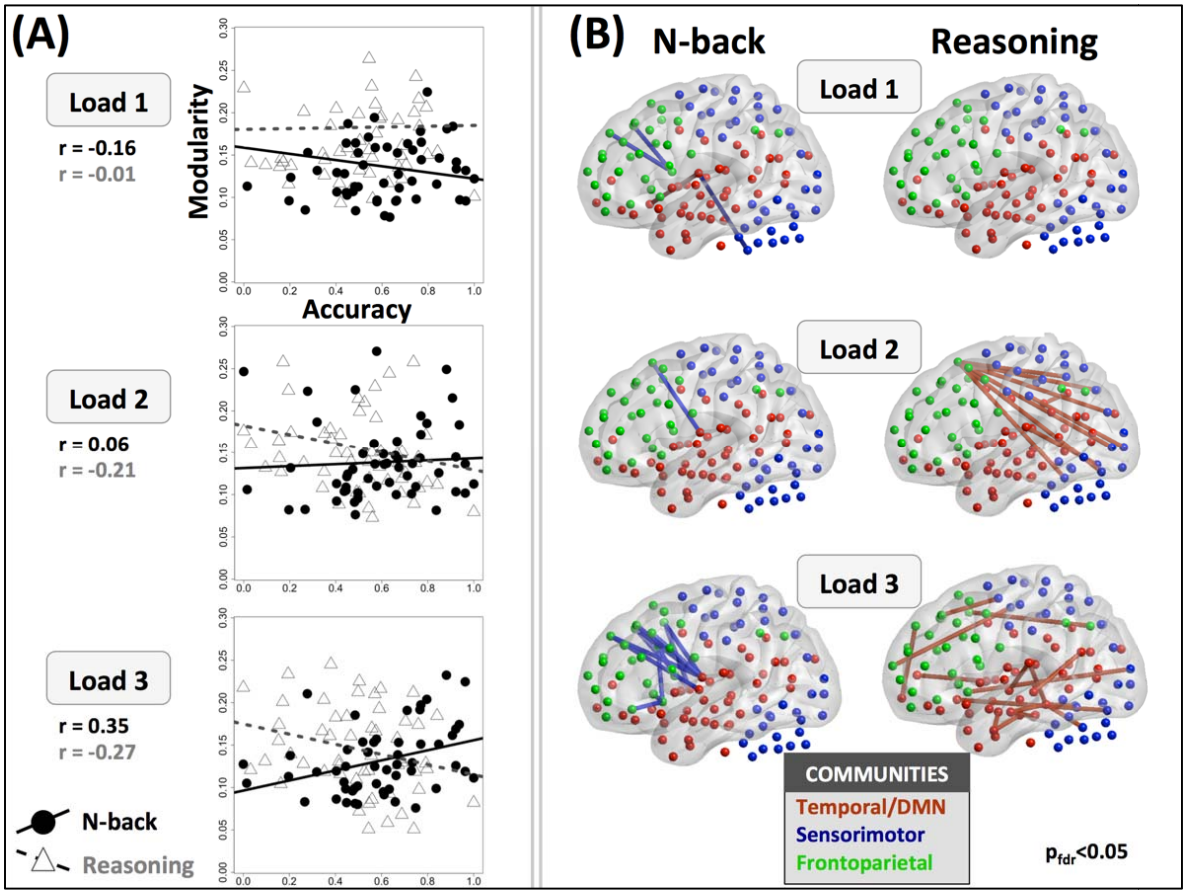
Overlap of the activation patterns in the n-back and spatial reasoning tasks ($p_{FWE}<0.05$). L1 = load 1 (1-back and reasoning-A); L2 = load 2 (2-back and reasoning-B); L3 = load 3 (3-back and reasoning-C).

Fig. 4. Shifts in modular dynamics associated with successfully solving working memory and reasoning tasks.



Different trajectories of modularity (Q, y-axis) change as a function of difficulty load (Loads 1-3, x-axis) in good (thick lines) versus poor (thin lines) performers (median splits);
No task = fixation blocks; Load 1 = 1-back and reasoning-A; Load 2 = 2-back and reasoning-B; Load 3 = 3-back and reasoning-C.

Fig. 5. Different associations between large-scale neurodynamics and performance in n-back and reasoning tasks.



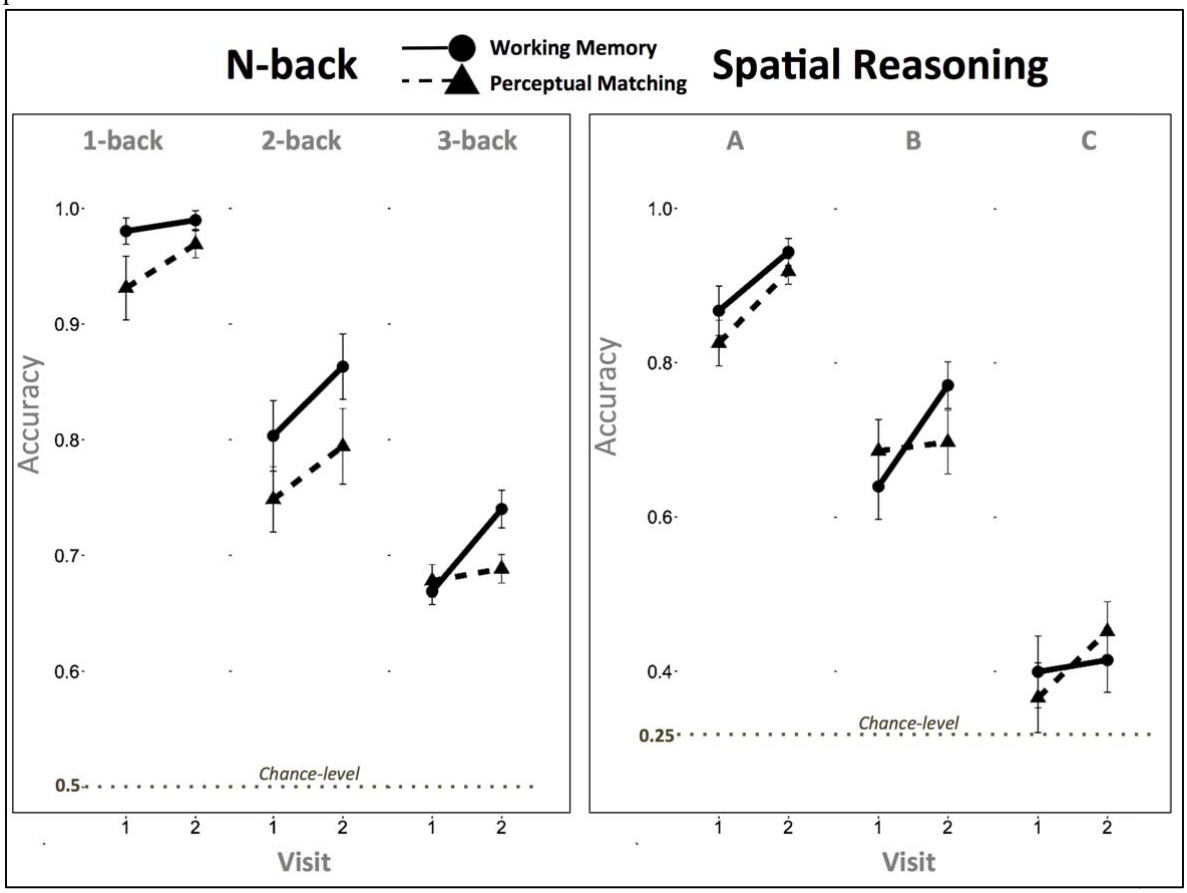
The plot shows associations of subjects' performance with modularity **(A)** and raw functional connectivity **(B)**.
On the plot **B**, blue lines show significant negative associations and red lines indicate positive associations between modularity and performance at p_{FDR}<0.05. The identified pattern suggests that better n-back performance is associated with increased segregation between the control and the DMN communities, whereas superior performance in the complex reasoning task is linked with enhanced connectivity of the control network with the sensory community and the DMN.
Load 1: 1-back and reasoning-A; Load 2: 2-back and reasoning-B; Load 3: 3-back and reasoning-C.

Effects of working memory training on performance and modularity. The main effect of time (pretest vs. posttest) on performance was statistically significant, which indicated general improvements in n-back and reasoning performance (0.57 SDs for n-back, t(259)=4.97, p<0.001; 0.64 SDs for the reasoning task, t(259)=4.02, p<0.001; **Figure 6**). Working memory training led to larger increases in n-back performance at higher loads than perceptual matching training did (load-by-group-by-visit: t(259)=2.33, p=0.02). T-tests comparing gains from working memory training with gains from perceptual matching training at different loads indicated a statistically significant effect at 3-back only. There was a 0.88 SD improvement in the working memory training group and a 0.1 SD

improvement in the control group, t(43.029)=2.92, p=0.003. Neither the group-by-visit nor the load-by-group-by-visit effects on reasoning performance were statistically significant, which confirmed the lack of far transfer effects of working memory training to reasoning published in our previous report of the behavioral data from this study (43). Subsequent t-tests revealed that improvement in block "B" accuracy (simple analytical) tended to be larger in the active (0.695 SD) than the control group (0.13 SD), t(46.331)=-2.01, p=0.026. It is worth noting, however, that this comparison would not survive strict p-threshold correction for multiple testing (0.05/3 tests = 0.016).

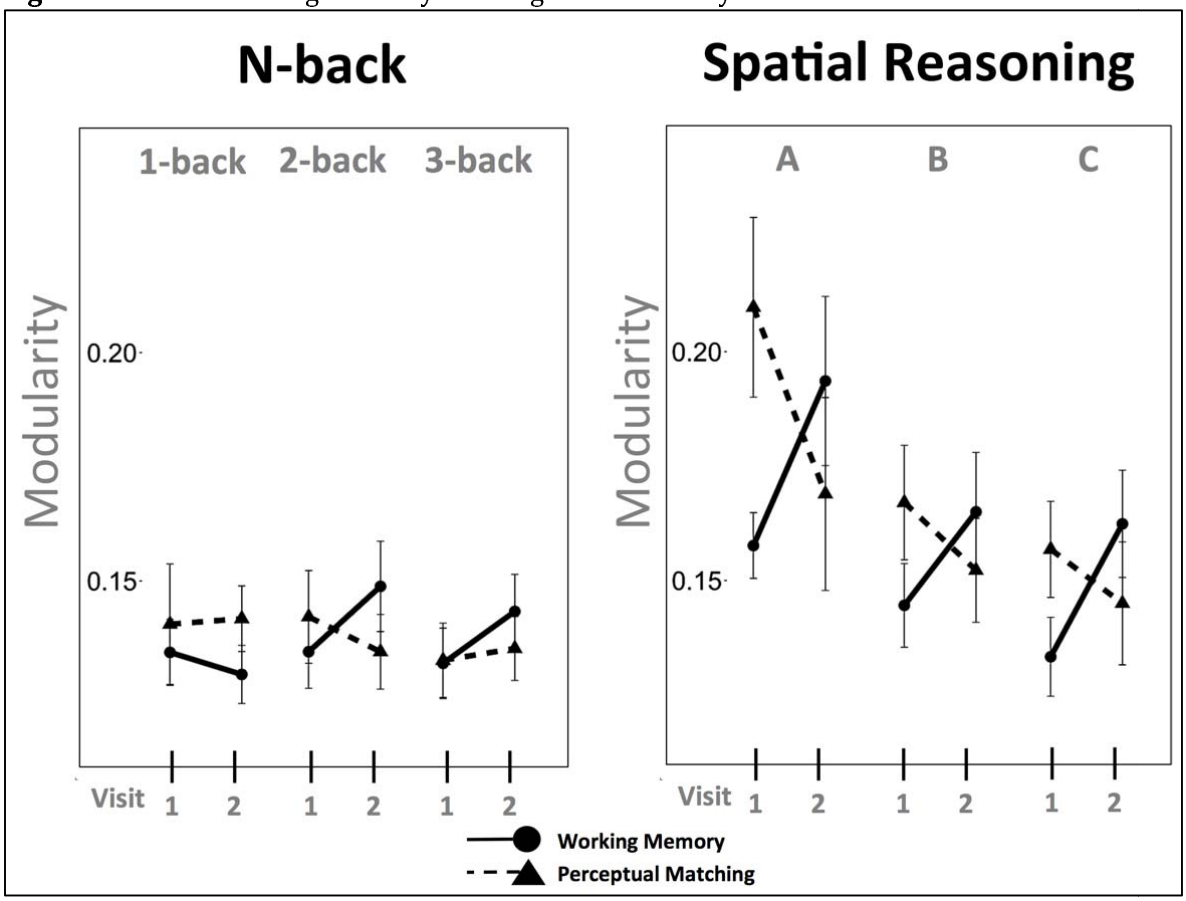
We found that the active and control interventions affected modularity differently, with a statistically significant group by time interaction, t(303)=3.5, p<0.001 (**Figure 7)**. Taken separately, working memory training resulted in a significant increase in modularity, 0.31 SD, t(296)=3.3, p=0.001, whereas perceptual matching training did not result in statistically significant changes in modularity. Instead, it exhibited a trend toward reducing it, -0.2 SD, t(285)=-1.8, p=0.07. The group-by-load-by-time effect, however, was non-significant, which is consistent with a load-invariant effect of working memory training on modularity. The group-by-task-by-time interaction was statistically significant t(569)=2.19, p=0.03, indicating that effects of training on modularity were somewhat larger for the reasoning task.

Fig. 6. Effects of working memory training on in-scanner n-back and reasoning performance.



Both groups showed significant improvements in performance. The active (working memory training) group demonstrated larger improvements on higher difficulty loads in n-back, load-by-group-by-visit: $t(259)=2.33$, $p=0.02$. The effect was mainly driven by improvements in the 3-back task. In addition, the active group demonstrated performance improvements in the "simple analytical" (block "B") part of the reasoning task, $t(46.331)=-2.01$, $p=0.026$.

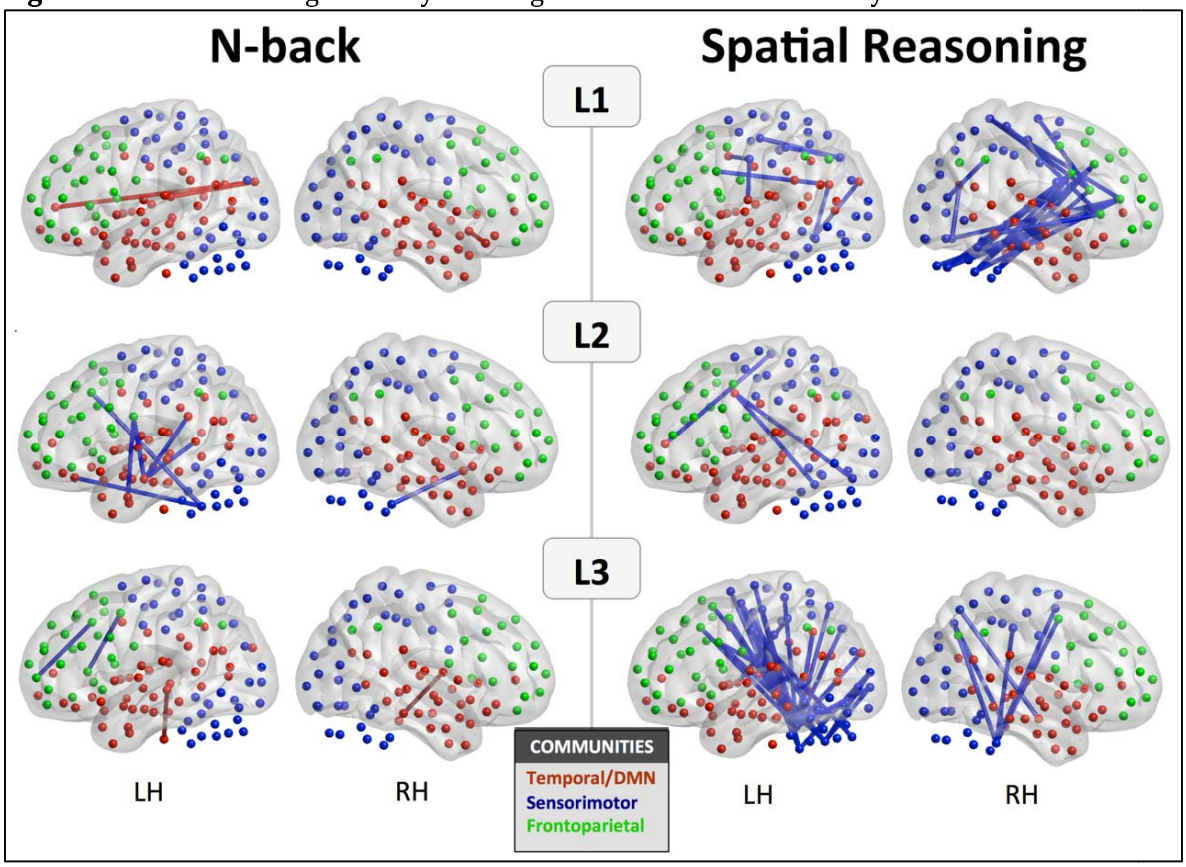
Fig. 7. Effects of working memory training on modularity



The active (working memory training) group produced greater increases in modularity than did the control (perceptual matching training) group, t(303)=3.5, p<0.001. Taken separately, working memory training resulted in a significant increase in modularity, 0.31 SD, t(296)=3.3, p=0.001, whereas perceptual matching training did not result in statistically significant changes in modularity. Instead, it exhibited a trend toward reducing it, -0.2 SD, t(285)=-1.8, p=0.07. The group-by-load effects were non-significant, which suggests a load-invariant influence of working memory training on modularity.

Longitudinal analyses of raw functional connectivity revealed that working memory training reduced connectivity of the frontoparietal network mostly with the sensorimotor community and some components of the default mode / temporal cluster, with more marked effect seen in the reasoning paradigm (Group x Time T-contrast: $p_{fdr}<0.05$, **Figure 8**).

Fig. 8. Effects of working memory training on functional connectivity



Working memory training reduced connectivity of the frontoparietal network with the sensorimotor community and components of the default mode/ temporal cluster; a more marked effect was seen in the reasoning paradigm. Group x Time T-contrast ($p_{fdr}<0.05$), red edges represent significant increases in functional connectivity and blue edges represent training-induced reductions. LH/RH = left/right hemispheres; L1 = load 1 (1-back and reasoning-A); L2 = load 2 (2-back and reasoning-B); L3 = load 3 (3-back and reasoning-C).

Discussion

Employing methods of large-scale network analyses, we found that despite high correlations in the behavioral data and substantial overlap of brain activation patterns, performance in complex reasoning and working memory tasks was associated with different modes of large-scale connectivity. Better performance in a typical working memory task (n-back) was associated with higher modularity (i.e., increased between-network segregation), whereas better performance in a reasoning task was associated with lower modularity (i.e., more interactions between the major brain networks) at higher task-complexity. Furthermore, working memory training modified neurodynamics in both tasks in a similar way; that is, by increasing modularity. Perceptual matching training, on

the other hand, did not have any significant impact on it. The observed effects appeared to be driven by integration within the frontoparietal control network and its segregation from the other communities.

Balanced modular organization of the brain is necessary for the emergence of the critical dynamics important to adaptive information processing and learning (18-20, 44, 45). Previous researchers have found that most cognitive tasks generally tend to perturb these modular dynamics by shifting them to a more integrated state (29-31). We observed a similar pattern of task-general reductions of modularity from rest to task in the present study. However, we also showed that performance on working memory tasks and complex reasoning tasks is associated with modularity in different ways, such that between-network interplay benefits complex reasoning, whereas sustaining a more modular organization benefits working memory performance.

The observed divergent pattern of shifts in connectivity is generally in line with the literature, which suggests that automated tasks requiring outward attentional focus tend to favor segregation of the DMN and the frontoparietal CN (34, 35, 46), whereas tasks requiring novel and creative problem solving benefit from enhanced between-network interplay (13, 36). It may be that lower modularity is associated with better complex reasoning performance because it relates to an exploratory mode of cognition (39) beneficial for rule discovery and abstraction (41). In contrast, such a mode of cognition may hinder working memory performance, which relies on a more modular network configuration with stabilized dynamics of the fronto-parietal network coupled with effective suppression of DMN activity.

These distinct associations were more marked at higher task complexity. The findings suggest that complex reasoning requires a more exploratory mode of cognition than simple reasoning, whereas solving more difficult working memory tasks requires a more focused state of cognition and neurodynamics. This possibility is in line with recent findings that

performance on simple tasks is positively related to resting state modularity, whereas performance on complex tasks is negatively related to resting state modularity (38).

It is worth noting that the circuit described in the influential parieto-frontal integration theory (P-FIT) of intelligence represents a broader community than the fronto-parietal network discussed in the present study and also includes components of the default mode and visual clusters (47). The reported results thus do not contradict P-FIT. Instead, our observation that superior reasoning performance favors enhanced cross-talk of the fronto-parietal regions with occipital and medial prefrontal cortices is nicely in line with the P-FIT model.

Working memory training produced marked increases in global modularity, thus promoting the pattern of large-scale dynamics that was associated with better working memory performance prior to training. Interestingly, the same pattern of modularity changes was observed for the reasoning task although no improvements in total reasoning performance were observed. This suggests that working memory training fosters a mode of solving the reasoning tasks that relies more on working memory and less on an exploratory mode of cognition. Such an interpretation is consistent with the trends for training-related improvements on the "simple analytical" (B) reasoning tasks, which may benefit from superior modularity (because the task demands attention, but the rules for solving it are simple to detect and to abstract) but not on the "complex analytical" (C) tasks, which may benefit from less modular dynamics (because they require complex pattern discovery). It should be acknowledged that the ability to solve the complex analytical blocks is generally more representative of fluid intelligence. Therefore, despite potential trends in training-related differences in performance improvements on the simple analytical part of our reasoning task, our study does not support the thesis that it is possible to improve the spatial aspect of fluid intelligence with working memory training, at least in healthy older people. This conclusion is generally in line with the literature,

which shows that working memory training has either no effects at all (48) or, at best, small effects (49) on reasoning. Our results expand these observations: they suggest that the reason working-memory training does not improve fluid intelligence is because it shifts large-scale brain dynamics in ways that are not beneficial for solving novel and complex reasoning problems.

In summary, although inter-individual differences in reasoning and working memory performance, and most other cognitive tasks too, are positively correlated, the brain may solve these tasks in partly different ways. Rather than indicating strong links between the processes involved in reasoning and working memory, the correlated inter-individual difference could be produced by initially unrelated cognitive domains that support each other in learning and development over the lifespan (50). Indeed, it is clear that higher working memory capacity does not necessarily imply superior intelligence. A particularly illustrative example is the observation that chimpanzees have, in fact, better visual working memory than humans (51), which suggests that the cognitive advancement of the human species was unlikely driven by increased storage capacity. Storage capacity may, in fact, compete with some aspects of complex reasoning, such as insight and deep information processing capabilities (52). Our results are also broadly in line with the recently proposed idea of a trade-off in biological systems, in which high-modular organization provides evolutionary benefits over shorter timescales, and low-modular systems exhibit greater fitness over longer timescales (38, 53). Notably, imposing time constraints on reasoning problems appear to increase the correlation between subjects' performance on such tasks with their performance on working memory tests (4). The trade-off hypothesis (38, 53) and our results are in accordance with the observation that working memory training studies tend to show transfer effects when reasoning tasks are administered under more stringent time constrains (54) (possibly benefiting from high-modular states), whereas

those that administer the tasks in a traditional, time-relaxed, manner (that benefit from low-modular dynamics) do not report such effects (55).

This study had a number of limitations. First, one could argue that the tasks we used inside the scanner do not properly represent corresponding abilities of working memory and fluid intelligence, or at least that the tasks only tap into selected aspects of these abilities. For example, some tasks that are supposed to characterize the same domain of working memory (e.g., complex span and n-back) actually have relatively low correlation and may therefore represent different aspects of working memory (56). Meanwhile, some researchers have questioned the plausibility of this argument, as several studies demonstrate that the performances on the two tasks converge when working memory is modeled as a latent construct (57, 58). However, the present study addressed updating aspects of working memory rather than a broader construct of working memory. Another potential limitation of the study is the absence of counter-balancing of the fMRI tasks, which may have affected the reliability of the results because of carry-over effects of the n-back paradigm on reasoning. However, consistent engagement of the control circuit in both tasks, strong associations between modularity and performance, together with the occurrence of longitudinal changes in the expected direction strongly support plausibility of the proposed interpretations. Finally, the present study was conducted on a sample of older adults, and the results therefore may not be generalizable to younger populations, especially since community structure of the brain undergoes substantial reconfiguration over different periods of the lifespan (21, 22).

Despite these issues, we conclude that reasoning performance and working memory performance in older adults have different associations with large-scale functional brain connectivity at higher task complexity. We also found that working memory training fosters task-invariant increases in functional network modularity and does not improve performance on complex reasoning tasks. Instead, complex reasoning tasks benefit from a

higher degree of between-network integration. Together, these results indicate that complex human reasoning abilities cannot be reduced to working memory and add weight to the conclusion that working memory training interventions will not result in improvements that transfer to broader cognition.

Methods

Subjects. For the present analysis, the sample consisted of 53 participants who had brain scans of acceptable quality from two fMRI scanning sessions, one before and one after four weeks of cognitive training. Imaging data quality criteria were: i) no motion artifacts in the raw data noticeable to the naked eye, ii) max frame-wise displacement below 3mm, iii) less than 10% of missing volumes after the de-spiking procedure. The study was approved by the Regional Ethical Review Board in Stockholm (DNR: 2014/2188-31/1) and was conducted in accordance with the revised declaration of Helsinki [2013]. Informed consent was obtained from all subjects prior to enrollment. Standardized physical and neuropsychological examinations were also carried out at screening. The subjects were randomly assigned to one of the two alternative training programs: (1) working memory (active; n = 28; age = 68.89±3.13 years; education = 15.17±3.25 years; number of women = 14; attendance = 19.07±0.94 session) and (2) perceptual matching (control; n = 25; age = 69.48±2.96 years; education = 15.02±4 years; number of women = 14; Attendance = 19±1.58 session) and were blind to the hypotheses about the two protocols. The randomization was accomplished in an automated manner using R-programming language (59) for each study wave separately with age, gender and fluid intelligence used as stratifies.

The main study targeted 123 participants and imaging data were acquired for a random subsample of 78 subjects, 25 of whom were excluded from the analyses in keeping with the image quality criteria (see above). See (43) for details. As a part of the study, half of the subjects received transcranial direct current stimulation (tDCS), whereas the other half received sham stimulation. Since the main study found no effects of stimulation on any of the collected cognitive measures, we did not estimate corresponding effects of stimulation on the imaging data in the present set of analyses, but accounted for stimulation as an additional nuisance regressor during modeling.

Cognitive training. Between pretest and posttest assessment with cognitive tasks and imaging, a JAVA-based computerized cognitive training battery was administered over 20 sessions (4 weeks, weekends off), each lasting 40 minutes. In the active program, four tasks trained two crucial facets of working memory: updating (running span and n-back tasks) and switching (rule- and task-switching tasks), each with four randomly varying stimuli sets to minimize the use of stimuli-specific strategies. The control intervention was focused on perceptual matching speed, employing four versions of the same task. In the task, participants were asked to assess the similarity of the presented figures and shapes, responding as quickly and accurately as possible. The training batteries were designed to be adaptive, such that the difficulty of the tasks increased systematically with improvements in performance so that the participants always trained at the highest level reached. The batteries were designed and carefully pilot-tested to balance the length and scope of the training, and the working memory training (active) and perceptual matching training (control) groups did not differ in the average level reached by the participants or in overall attendance. The cognitive training took place the week after an extensive 5-day assessment of cognitive functions. See (43) for details and a comparative description of the two batteries.

Magnetic resonance imaging. Structural and fMRI scans were collected with the GE Discovery MR-750 3.0 Tesla scanner with the 32-channel research head coil located at the Karolinska University Hospital, MRI center in Solna, Sweden. Structural images were acquired employing a standardized contrast-enhanced T1 spoiled gradient (SPGR) "BRAVO" sequence with 0.94 mm^3 isotropic voxels, FoV 256 mm (256 x 256 matrix), TR/TI/TE = 5.688/450/2.492 ms, flip angle = 12°. Functional MRI data were collected

using echo-planar T2* imaging to measure the blood-oxygen-level dependent (BOLD) signal (voxel size 3 mm^3 , TE/TR=30/2000 ms, flip angle = 80°).

Functional magnetic resonance imaging tasks. Two tasks were administered inside the MR camera to evaluate brain dynamics during working memory and reasoning. A figural n-back task assessed working memory performance. Three 1-minute blocks were administered: 1-, 2- and 3-back (each repeated 3 times) with a 4-second fixation screen indicating the rule for the forthcoming level. See **Figure 1** for an illustration. For this task, accuracies were calculated as the average value of true positives and true negatives with a chance level of 50%. The second task was developed to evaluate the neurodynamics involved in complex visuospatial reasoning. Subjects were instructed to complete the presented pattern by selecting one of the four solutions shown at the bottom of the screen. One of the four answers was always correct. Like the n-back task, this paradigm had three levels of difficulty: A, perceptual matching; B, simple analytical (single rule); and C, difficult analytical (with at least two non-trivial rules), similar to the implementation proposed by Yamada et al. (60). Each block was repeated 3 times in a pseudo-random order during the scanning session. The reasoning task was self-paced to better match typical fluid intelligence paradigms (such as Raven's progressive matrices) with fixed (1-minute) blocks. The subjects were explicitly instructed not to guess. For the reasoning paradigm, accuracy was defined as fraction of correctly solved problems relative to the total number of attempts per block. In a number of simulations with randomly varying answers and total number of responses, chance-level accuracy was estimated at 25%.

In both tasks, the subjects demonstrated consistent accuracy drop as the difficulty level increased. In all the analyses, performance was measured as total task accuracy, which was defined as the sum of block-specific accuracies. Both tasks were administered twice: before and after the cognitive training.

Cognitive assessment at pretest and posttest. We evaluated the effects of working memory training via an extensive testing battery completed over four sessions, each lasting about 180 minutes. Pre-intervention testing took place 2 weeks before the training started, and post-testing was completed the week after the intervention. The battery was identical on both occasions and included a number of diverse cognitive tests measuring switching (task- and rule-switching), updating (n-back, running span), attention (temporal expectancy, perceptual matching), episodic memory (verbal and spatial recall), and spatial (WASI, Raven's, BETA matrices) and verbal (Syllogisms, BIS Analogies, ETS Kit inference) reasoning. See (43) for detailed description of the cognitive assessment.

Magnetic resonance imaging processing. A standardized SPM12-based (Wellcome Trust Center for Neuroimaging, UCL) pipeline was implemented in the MATLAB R2013a environment (61).

Each subject's T2* imaging data underwent subsequent steps for slice-timing correction, spatial realignment, and registration to standardized Montreal Neurological Institute (MNI) space using a population-specific template produced from the set of T1 images with the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) algorithm (62).

An extended 24-parameter model (63) was employed for head motion correction. The images were also smoothed (FWHM = 6 mm), de-trended, and, for the connectivity part, also band-pass filtered (0.008–0.12 Hz). The resulting output was used to calculate connectivity matrices (one matrix per each task and load level) based on temporal (Pearson r) correlation between the regions parcellated in accordance with the Craddock's 200-ROI scheme (64). After removing negative weights, the adjacency matrices were then used to estimate whole-brain network modularity as implemented in the "Brain

Connectivity Toolbox" (65). The Louvain community detection method generalized for weighted graphs was used in the study as a measure of between-network segregation. The method attempts to solve an optimization problem by maximizing modularity (Q) given observed graph weights:

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j),$$

Where:
A_{ij} - weight (connectivity strength) between nodes i and j
k_i, k_j - sum of the weights attached to nodes I and j
m - sum of all graph weights
c_i, c_j - communities of the nodes
δ - Kronecker delta

When there are few edges between the communities and within-cluster density is high, Louvain modularity is close to 1, but when the number of within-modular connections is comparable to the one from a random graph, modularity usually approaches 0. Practically, it measures level of modular segregation in a graph.

Statistical analyses. After the preprocessing and cleaning, behavioral and imaging measures underwent outlier detection and checks for distribution normality. Between-group comparisons of demographic data used two-sample t-tests for continuous variables and binomial tests for sex ratios.

Next, we conducted a number of analyses to evaluate the reliability and validity of the collected measures. Pre-post Pearson correlation coefficients were calculated for the behavioral data to evaluate reliability. To address the problem of task validity, we analyzed updating and spatial reasoning domains in latent space, employing structural equation modeling (SEM) of the two constructs with *offline* measures and data collected inside the MRI scanner (See **Figure 2**). Each of the *offline* latent variables (LVs) was modeled with 3

tasks (updating: n-back and two running span tasks; spatial reasoning: Raven's progressive matrices, WASI, and BETA). *Online* (fMRI) LVs were formed with two indicators each (updating: 2-back and 3-back; spatial reasoning: "simple analytical" and "difficult analytical"; 1-back and reasoning "A" accuracies were excluded from the SEM due to ceiling effects). The variance of the LVs was constrained to 1 to make path coefficients equivalent to Pearson correlations. SEM was also repeated for block-specific measures of in-scanner performance (2-back and "simple analytical"; 3-back and "difficult analytical").

The next block was focused on evaluating task-related BOLD activity and modularity at baseline. First, we estimated first-level contrasts in both tasks: a) 2-back – 1-back, b) 3-back – 1-back, c) reasoning "B" – reasoning "A," and d) reasoning "C" – reasoning "A." Then we performed a second-level analysis of the activations, which yielded block-specific statistical maps. Once the smoothness of the data had been estimated, adjustment for multiple tests was carried out using Mote Carlo simulations (n=5000) as implemented in the AlphaSim algorithm (66) with an initial cluster-forming threshold of $p<0.005$. In addition, individual block-specific activation patterns were matched with a meta-analytical map derived using Neurosynth software (http://www.neurosynth.org/) (67). This procedure employs voxel-wise Pearson correlation as a measure of similarity (68) between the statistical map and the meta-analytical image generated for a specific feature of interest. In the present analysis, the search was completed for the terms "working memory" (n=901 studies) and "reasoning" (n=147). The search was also completed for "executive control" (n=157), which was used as a third independent term to evaluate validity of the activation patterns.

Next, we embarked on the analyses of large-scale dynamics, evaluating the differences in such dynamics between the tasks and between-task correlation of modularity across different loads. Subsequently, we conducted a number of mass-univariate tests of raw functional connectivity to identify which of the networks and

between-network interactions were involved in the reported global effects. This was accomplished by fitting a linear model that evaluated correlations between block-specific functional connectivity and total task performance.

The last analytical block addressed effects of working memory training on performance and large-scale neurodynamics. We started by evaluating the main effects of time on task performance and then estimated load x group x time effects in each of the two tasks using linear mixed-effects models (random effect: "subjects"). Standardized effect sizes were also calculated and reported as the mean difference in performance divided by the pre-test standard deviation.

Modularity was analyzed as overall pre/post change, as well as Time x Group interactions. At this stage, a full-interaction model was fitted to evaluate the effects of working memory training in two tasks (modularity ~ load x task x time x group x WM performance, random effect: "subjects"). Finally, a subsequent analysis of raw functional connectivity was conducted to localize the reported global effects byestimating Group x Time effects on connectivity for each block.

Statistical software and toolboxes. All statistical analyses of behavioral data and modularity were performed with the R programming language, version 3.2.2 (69) and "nlme" toolbox (70). In all models, age, sex, and first 3 principal components from the motion parameters were introduced as nuisance covariates.

Mass-univariate analyses of BOLD activation patterns were carried-out in SPM12 (Wellcome Trust Center for Neuroimaging, UCL).

The Network-Based Statistics Toolbox (42) was used to analyze raw functional connectivity.

Structural equation modeling was conducted using Ωnyx software (71).

Data availability. Raw MRI and behavioral data can be requested from the authors and transferred for specific analysis projects that are in line with the original ethical approval. This requires a data use agreement, which effectively transfers the confidentiality obligations of the institution (Karolinska Institutet) at which the original research was conducted to the institution of the recipient of the data. An R-script specifying main analytical and plotting steps is available at: https://github.com/alex-lebedev/RBT1_modularity

Acknowledgements. We thank Anders Rydström, Linda Lidborg, Jakob Norgren, Helena Franzén, and Simon Peyda for help with data collection and Marie Helsing for help with recruitment and organization. Thanks also to Alireza Salami, Goran Papenberg and Lars Nyberg for comments on an earlier draft of this article. This research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement # 617280 - REBOOT. Martin Lövdén was also supported by a "distinguished younger researcher" grant from the Swedish Research Council (446-2013-7189).

Conflicts of interest. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Conway AR, Kane MJ, & Engle RW (2003) Working memory capacity and its relation to general intelligence. *Trends in cognitive sciences* 7(12):547-552.
2. Colom R, Abad FJ, Quiroga MA, Shih PC, & Flores-Mendoza C (2008) Working memory and intelligence are highly related constructs, but why? *Intelligence* 36(6):584–606.
3. Oberauer K, Süβ H-M, Wilhelm O, & Wittmann WW (2008) Which working memory functions predict intelligence? *Intelligence* 36(6):641–652.
4. Chuderski A (2013) When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence* 41(4):244–262.
5. Little DR, Lewandowsky S, & Craig S (2014) Working memory capacity and fluid abilities: the more difficult the item, the more more is better. *Frontiers in psychology* 5:239.
6. Carpenter PA, Just MA, & Shell P (1990) What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychol Rev* 97(3):404-431.
7. Wiley J, Jarosz AF, Cushen PJ, & Colflesh GJ (2011) New rule use drives the relation between working memory capacity and Raven's Advanced Progressive Matrices. *Journal of experimental psychology. Learning, memory, and cognition* 37(1):256-263.
8. Gray JR, Chabris CF, & Braver TS (2003) Neural mechanisms of general fluid intelligence. *Nat Neurosci* 6(3):316-322.
9. Duncan J, *et al.* (2000) A neural basis for general intelligence. *Science* 289(5478):457-460.
10. Lee KH, *et al.* (2006) Neural correlates of superior intelligence: stronger recruitment of posterior parietal cortex. *Neuroimage* 29(2):578-586.
11. Colom R, Karama S, Jung RE, & Haier RJ (2010) Human intelligence and brain networks. *Dialogues in clinical neuroscience* 12(4):489-501.
12. Krawczyk DC (2012) The cognition and neuroscience of relational reasoning. *Brain research* 1428:13-23.
13. Hearne LJ, Mattingley JB, & Cocchi L (2016) Functional brain networks related to individual differences in human intelligence at rest. *Scientific reports* 6:32328.
14. Clark CM, Lawlor-Savage L, & Goghari VM (2017) Comparing brain activations associated with working memory and fluid intelligence. *Intelligence* 63:66-77.
15. Benedek M, Jauk E, Sommer M, Arendasy M, & Neubauer AC (2014) Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence* 46:73-83.
16. Necka E, Zak P, & Gruszka A (2016) Insightful Imagery is Related to Working Memory Updating. *Frontiers in psychology* 7:137.
17. Chialvo DR (2010) Emergent complex neural dynamics. *Nature Physics* 6:744–750.
18. Wang SJ & Zhou C (2012) Hierarchical modular structure enhances the robustness of self-organized criticality in neural networks. . *New. J. Phys* 14.
19. Moretti P & Munoz MA (2013) Griffiths phases and the stretching of criticality in brain networks. *Nature communications* 4:2521.
20. Hilgetag CC & Hutt MT (2014) Hierarchical modular brain connectivity is a stretch for criticality. *Trends in cognitive sciences* 18(3):114-115.
21. Onoda K & Yamaguchi S (2013) Small-worldness and modularity of the resting-state functional brain network decrease with aging. *Neurosci Lett* 556:104-108.
22. Geerligs L, Renken RJ, Saliasi E, Maurits NM, & Lorist MM (2015) A Brain-Wide Study of Age-Related Changes in Functional Connectivity. *Cereb Cortex* 25(7):1987-1999.

23. Stevens AA, Tappon SC, Garg A, & Fair DA (2012) Functional brain network modularity captures inter- and intra-individual variation in working memory capacity. *PLoS One* 7(1):e30468.
24. Stanley ML, Dagenbach D, Lyday RG, Burdette JH, & Laurienti PJ (2014) Changes in global and regional modularity associated with increasing working memory load. *Frontiers in human neuroscience* 8:954.
25. Bassett DS, *et al.* (2011) Dynamic reconfiguration of human brain networks during learning. *Proc Natl Acad Sci U S A* 108(18):7641-7646.
26. Gallen CL, *et al.* (2016) Modular Brain Network Organization Predicts Response to Cognitive Training in Older Adults. *PLoS One* 11(12):e0169015.
27. Andric M & Hasson U (2015) Global features of functional brain networks change with contextual disorder. *Neuroimage* 117:103-113.
28. Godwin D, Barry RL, & Marois R (2015) Breakdown of the brain's functional network modularity with awareness. *Proc Natl Acad Sci U S A* 112(12):3799-3804.
29. Kitzbichler MG, Henson RN, Smith ML, Nathan PJ, & Bullmore ET (2011) Cognitive effort drives workspace configuration of human brain functional networks. *J Neurosci* 31(22):8259-8270.
30. Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, & Stamatakis EA (2015) Default Mode Dynamics for Global Functional Integration. *J Neurosci* 35(46):15254-15262.
31. Liang X, Zou Q, He Y, & Yang Y (2016) Topologically Reorganized Connectivity Architecture of Default-Mode, Executive-Control, and Salience Networks across Working Memory Task Loads. *Cereb Cortex* 26(4):1501-1511.
32. Meunier D, *et al.* (2014) Modular structure of functional networks in olfactory memory. *Neuroimage* 95:264-275.
33. Westphal AJ, Wang S, & Rissman J (2017) Episodic Memory Retrieval Benefits from a Less Modular Brain Network Organization. *J Neurosci* 37(13):3523-3531.
34. Hampson M, Driesen N, Roth JK, Gore JC, & Constable RT (2010) Functional connectivity between task-positive and task-negative brain areas and its relation to working memory performance. *Magnetic resonance imaging* 28(8):1051-1057.
35. Piccoli T, *et al.* (2015) The default mode network and the working memory network are not anti-correlated during all phases of a working memory task. *PLoS One* 10(4):e0123354.
36. Beaty RE, *et al.* (2014) Creativity and the default network: A functional connectivity analysis of the creative brain at rest. *Neuropsychologia* 64:92-98.
37. Douw L, Wakeman DG, Tanaka N, Liu H, & Stufflebeam SM (2016) State-dependent variability of dynamic functional connectivity between frontoparietal and default networks relates to cognitive flexibility. *Neuroscience* 339:12-21.
38. Yue Q, *et al.* (2017) Brain Modularity Mediates the Relation between Task Complexity and Performance. *Journal of cognitive neuroscience* 29(9):1532-1546.
39. Senden M, Goebel R, & Deco G (2012) Structural connectivity allows for multi-threading during rest: the structure of the cortex leads to efficient alternation between resting state exploratory behavior and default mode processing. *Neuroimage* 60(4):2274-2284.
40. Friston K, Breakspear M, & Deco G (2012) Perception and self-organized instability. *Frontiers in computational neuroscience* 6:44.
41. Cao B, Li W, Li F, & Li H (2016) Dissociable roles of medial and lateral PFC in rule learning. *Brain and behavior* 6(11):e00551.
42. Zalesky A, Fornito A, & Bullmore ET (2010) Network-based statistic: identifying differences in brain networks. *Neuroimage* 53(4):1197-1207.

43. Nilsson J, Lebedev AV, Rydström A, & Lövdén M (2017) Direct-current stimulation does little to improve the outcome of working-memory training in older adults. *Psychological Science (in press)*.
44. de Arcangelis L & Herrmann HJ (2010) Learning as a phenomenon occurring in a critical state. *Proc Natl Acad Sci U S A* 107(9):3977-3981.
45. Cocchi L, Gollo LL, Zalesky A, & Breakspear M (2017) Criticality in the brain: A synthesis of neurobiology, models and cognition. *Prog Neurobiol*.
46. Jolles DD, van Buchem MA, Crone EA, & Rombouts SA (2013) Functional brain connectivity at rest changes after working memory training. *Hum Brain Mapp* 34(2):396-406.
47. Jung RE & Haier RJ (2007) The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *The Behavioral and brain sciences* 30(2):135-154; discussion 154-187.
48. Melby-Lervag M, Redick TS, & Hulme C (2016) Working Memory Training Does Not Improve Performance on Measures of Intelligence or Other Measures of "Far Transfer": Evidence From a Meta-Analytic Review. *Perspectives on psychological science : a journal of the Association for Psychological Science* 11(4):512-534.
49. Au J, *et al.* (2015) Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic bulletin & review* 22(2):366-377.
50. van der Maas HL, *et al.* (2006) A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychol Rev* 113(4):842-861.
51. Inoue S & Matsuzawa T (2007) Working memory of numerals in chimpanzees. *Current biology : CB* 17(23):R1004-1005.
52. DeCaro MS, Van Stockum CA, Jr., & Wieth MB (2016) When higher working memory capacity hinders insight. *Journal of experimental psychology. Learning, memory, and cognition* 42(1):39-49.
53. Chen M & Deem MW (2015) Development of modularity in the neural activity of children's brains. *Physical biology* 12(1):016009.
54. Jaeggi SM, Buschkuehl M, Jonides J, & Perrig WJ (2008) Improving fluid intelligence with training on working memory. *Proc Natl Acad Sci U S A* 105(19):6829-6833.
55. Colom R, *et al.* (2013) Adaptive n-back training does not improve fluid intelligence at the construct level: gains on individual tests suggest that training may enhance visuospatial processing. . *Intelligence* 41:712–727.
56. Redick TS & Lindsey DR (2013) Complex span and n-back measures of working memory: a meta-analysis. *Psychonomic bulletin & review* 20(6):1102-1113.
57. Schmiedek F, Hildebrandt A, Lovden M, Lindenberger U, & Wilhelm O (2009) Complex span versus updating tasks of working memory: the gap is not that deep. *Journal of experimental psychology. Learning, memory, and cognition* 35(4):1089-1096.
58. Schmiedek F, Lovden M, & Lindenberger U (2014) A task is a task is a task: putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in psychology* 5:1475.
59. Anonymous (R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN: 3-900051-07-0. URL: http://www.R-project.org/.
60. Yamada T, *et al.* (2012) Functional alterations in neural substrates of geometric reasoning in adults with high-functioning autism. *PLoS One* 7(8):e43220.
61. Anonymous (MATLAB version 8.1. Natick, Massachusetts: The MathWorks Inc., 2013.
62. Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* 38(1):95-113.

63. Friston KJ, Williams S, Howard R, Frackowiak RS, & Turner R (1996) Movement-related effects in fMRI time-series. *Magn Reson Med* 35(3):346-355.
64. Craddock RC, James GA, Holtzheimer PE, 3rd, Hu XP, & Mayberg HS (2012) A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum Brain Mapp* 33(8):1914-1928.
65. Rubinov M & Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52(3):1059-1069.
66. Ward BD (2000) Simultaneous inference for fMRI data. Biophysics Research Institute, Medical College of Wisconsin. http://afni.nimh.nih.gov/pub/dist/doc/manual/AlphaSim.pdf
67. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, & Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* 8(8):665-670.
68. Smith SM, *et al.* (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A* 106(31):13040-13045.
69. R Development Core Team (2015) R: A language and environment for statistical computing. in *R Foundation for Statistical Computing.*
70. Pinheiro J, Bates D, DebRoy S, Sarkar D, & and R Core Team (2016) nlme: Linear and Nonlinear Mixed Effects Models. http://CRAN.R-project.org/package=nlme.).
71. von Oertzen T, Brandmaier AM, & Tsang S (2015) Structural Equation Modeling with Ωnyx. *Structural Equation Modeling: A Multidisciplinary Journal,* 22(1):148-161.

Figure legends

Fig. 1. fMRI tasks.

Two tasks were administered inside the MR camera to evaluate the brain dynamics underlying working memory and reasoning. A figural n-back task assessed working memory performance. Three 1-minute blocks were administered: 1-, 2- and 3-back (each repeated 3 times) with a 4-second fixation screen indicating the rule for the forthcoming level. The second task was developed to evaluate the neurodynamics involved in complex visuospatial reasoning. Subjects were instructed to complete the presented pattern by selecting one of the four solutions shown at the bottom of the screen. One of the four answers was always correct. Similar to the n-back task, this paradigm had three levels of difficulty: A, perceptual matching; B, simple analytical (single rule); and C, difficult analytical (with at least two non-trivial rules) tasks. Each block was repeated 3 times in a pseudo-random order during the scanning session. The reasoning task was self-paced (to better match typical fluid intelligence paradigms such as Raven's progressive matrices) within a fixed (1-minute) block. Thus, different subjects completed different numbers of items per block, and the subjects were explicitly instructed not to guess. For the reasoning paradigm, accuracy was defined as the fraction of correctly solved problems relative to the total number of attempts per block.

Fig. 2. Results of the structural equation modeling analysis.

Circles represent modeled latent variables; rectangles, measured variables (indicators); omnidirectional arrows, dependence; and bidirectional arrows, correlation. SR = spatial reasoning; UPD = updating; SR-B,C = spatial reasoning (fMRI) blocks "B" and "C." Number next to the bidirectional arrows represent standardized covariances (i.e., correlations) between the latent factors at pretest (black) and posttest (grey).

Fig. 3. Overlap of the task-related activation patterns

Overlap of the activation patterns in the n-back and spatial reasoning tasks (pFWE<0.05). L1 = load 1 (1-back and reasoning-A); L2 = load 2 (2-back and reasoning-B); L3 = load 3 (3-back and reasoning-C).

Fig. 4. Shifts in modular dynamics associated with successfully solving working memory and reasoning tasks.

Different trajectories of modularity (Q, y-axis) change as a function of difficulty load (Loads 1-3, x-axis) in good (thick lines) versus poor (thin lines) performers (median splits);
No task = fixation blocks; Load 1 = 1-back and reasoning-A; Load 2 = 2-back and reasoning-B; Load 3 = 3-back and reasoning-C.

Fig. 5. Different associations between large-scale neurodynamics and performance in n-back and reasoning tasks.

The plot shows associations of subjects' performance with modularity (A) and raw functional connectivity (B).
On the plot B, blue lines show significant negative associations and red lines indicate positive associations between modularity and performance at pFDR<0.05. The identified pattern suggests that better n-back performance is associated with increased segregation between the control and the DMN communities, whereas superior performance in the complex reasoning task is linked with enhanced connectivity of the control network with the sensory community and the DMN.
Load 1: 1-back and reasoning-A; Load 2: 2-back and reasoning-B; Load 3: 3-back and reasoning-C.

Fig. 6. Effects of working memory training on in-scanner n-back and reasoning performance.

Both groups showed significant improvements in performance. The active (working memory training) group demonstrated larger improvements on higher difficulty loads in n-back, load-by-group-by-visit: t(259)=2.33, p=0.02. The effect was mainly driven by improvements in the 3-back task. In addition, the active group demonstrated performance improvements in the "simple analytical" (block "B") part of the reasoning task, t(46.331)=-2.01, p=0.026.

Fig. 7. Effects of working memory training on modularity.

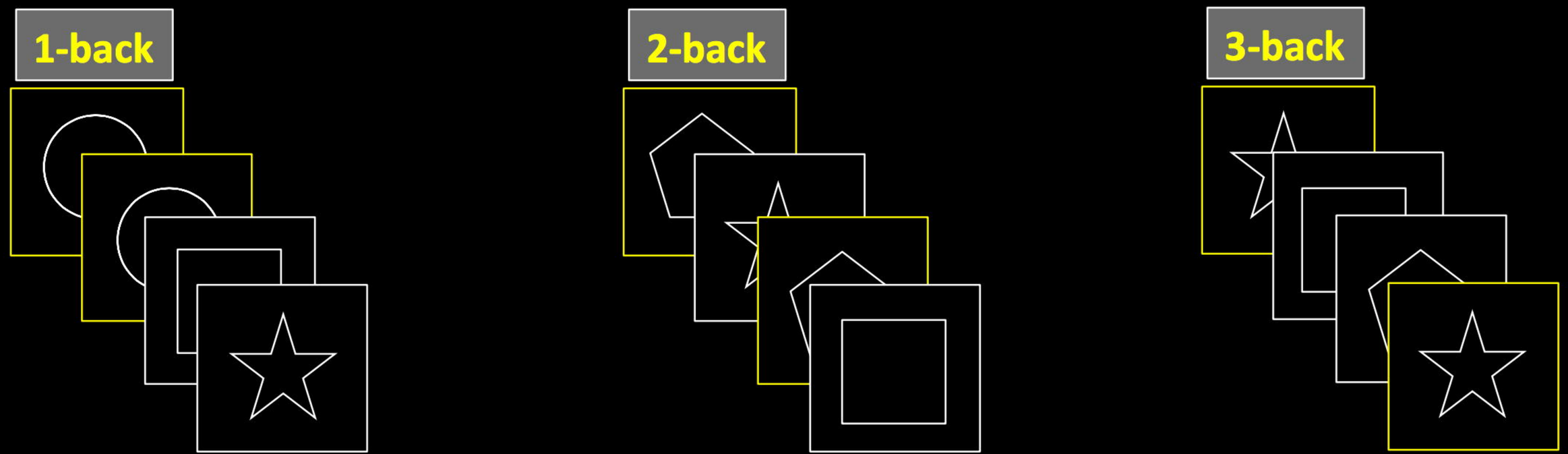
The active (working memory training) group produced greater increases in modularity than did the control (perceptual matching training) group, t(303)=3.5, p<0.001. Taken separately, working memory training resulted in a significant increase in modularity, 0.31 SD, t(296)=3.3, p=0.001, whereas perceptual matching training did not result in statistically significant changes in modularity. Instead, it exhibited a trend toward reducing it, -0.2 SD, t(285)=-1.8, p=0.07. The group-by-load effects were non-significant, which suggests a load-invariant influence of working memory training on modularity.

Fig. 8. Effects of working memory training on functional connectivity

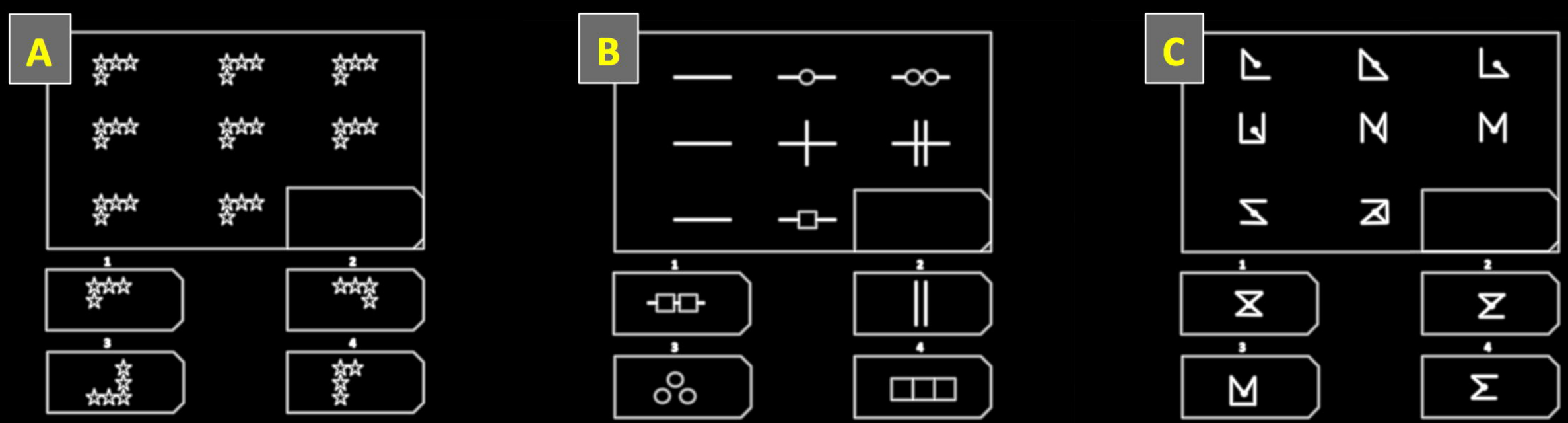
Working memory training reduced connectivity of the frontoparietal network with the sensorimotor community and components of the default mode/ temporal cluster; a more marked effect was seen in the reasoning paradigm. Group x Time T-contrast (pfdr<0.05), red edges represent significant increases in functional connectivity and blue edges represent training-induced reductions. LH/RH = left/right hemispheres; L1 = load 1 (1-back and reasoning-A); L2 = load 2 (2-back and reasoning-B); L3 = load 3 (3-back and reasoning-C).

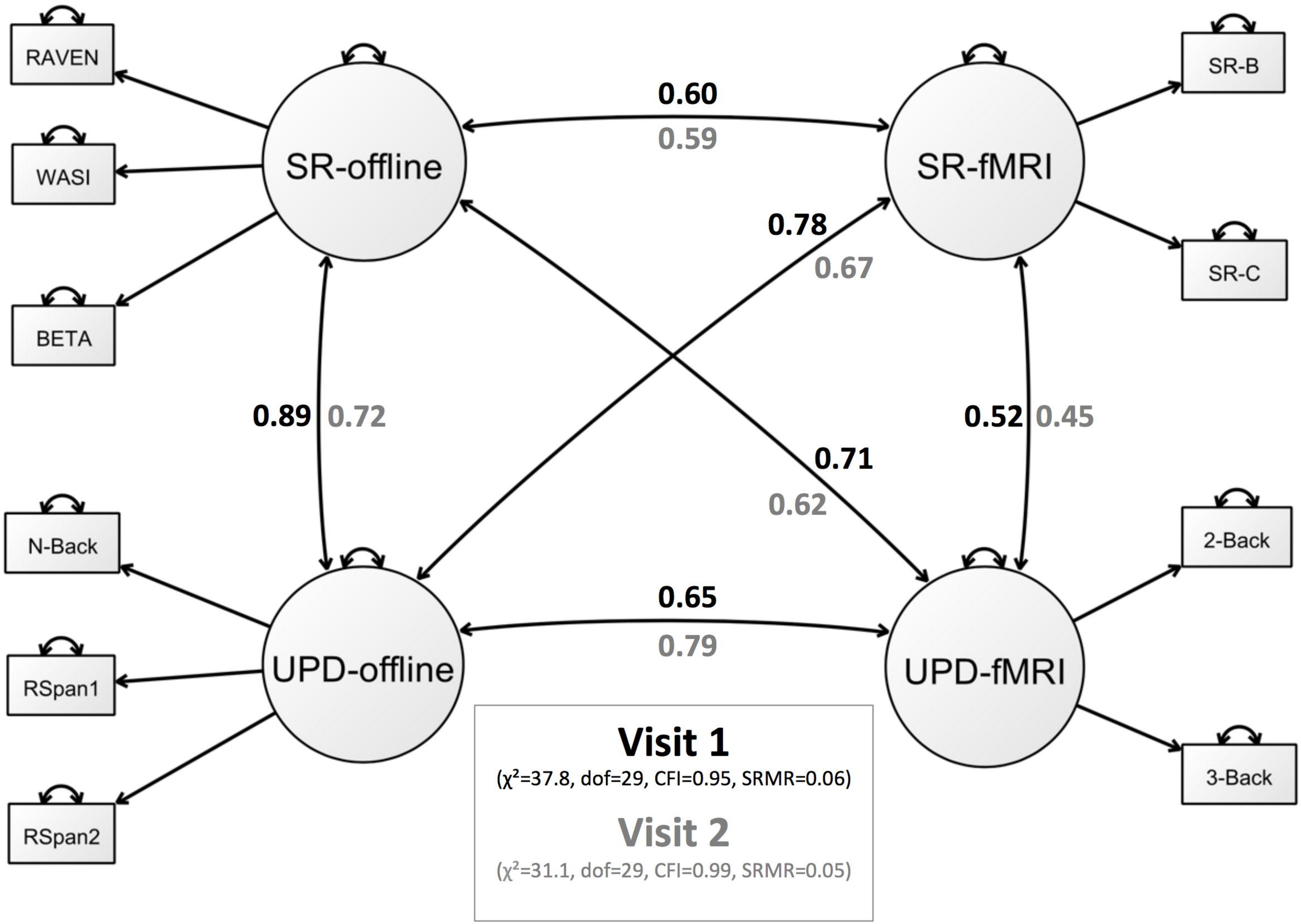
fMRI Tasks

N-back

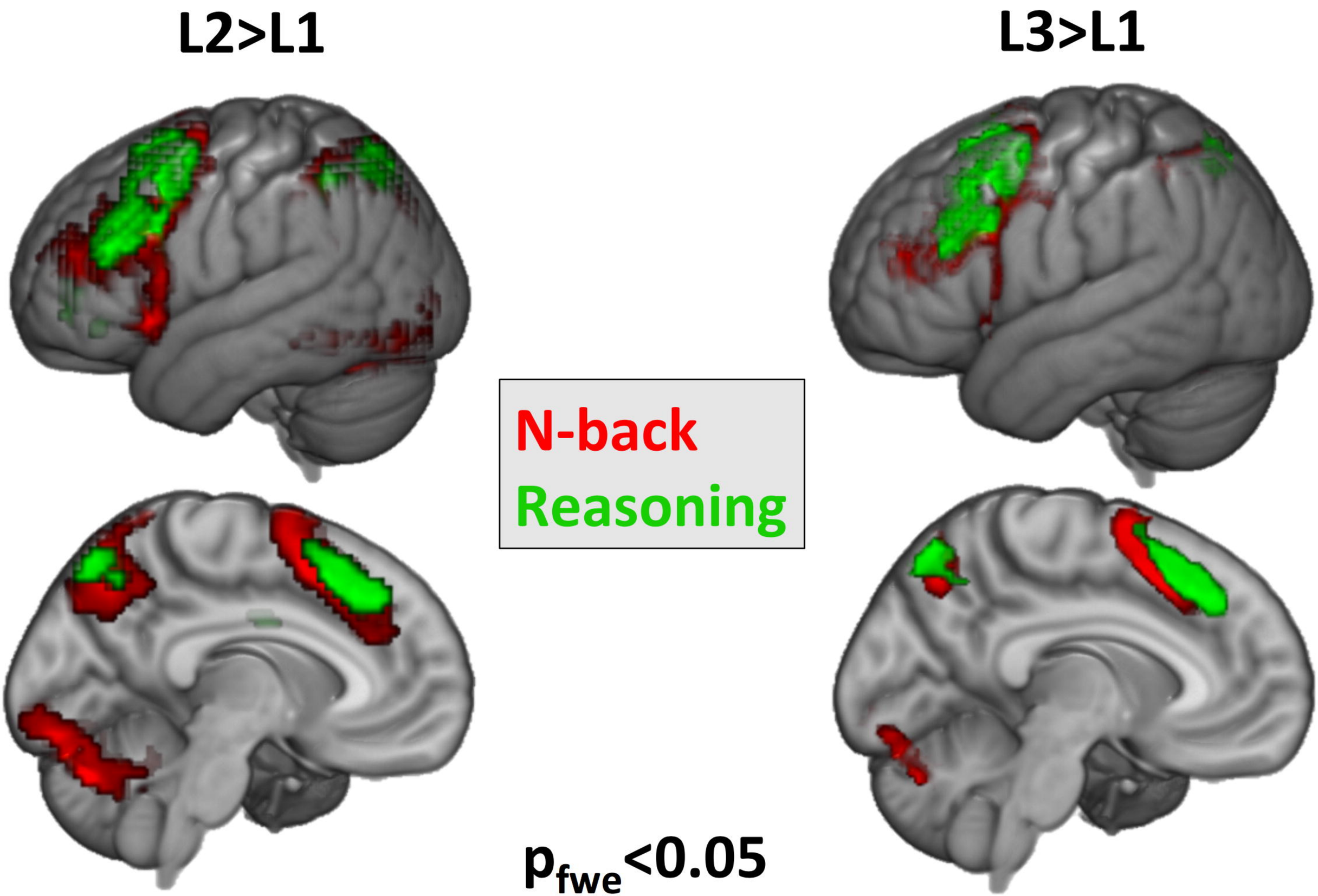


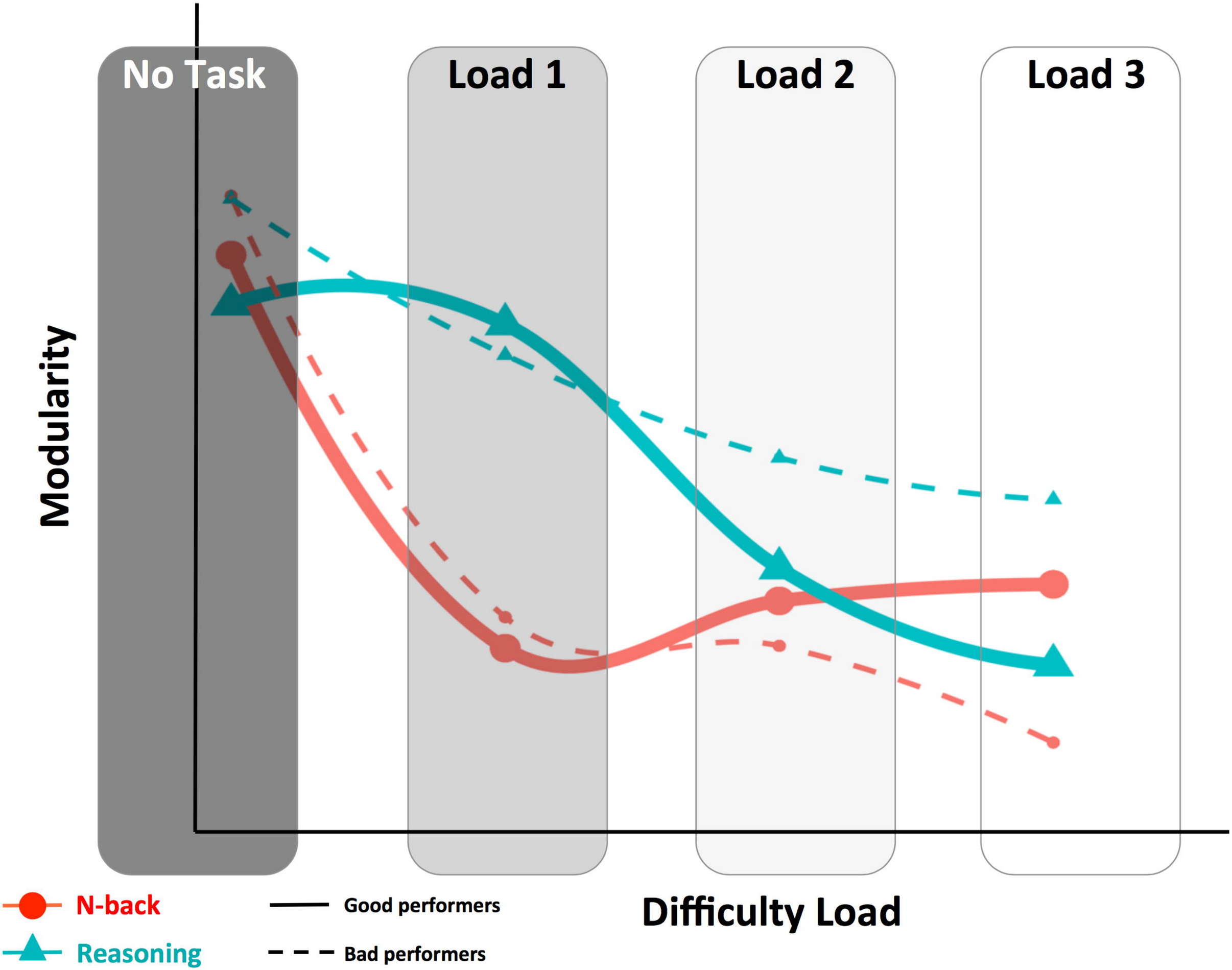
Spatial Reasoning

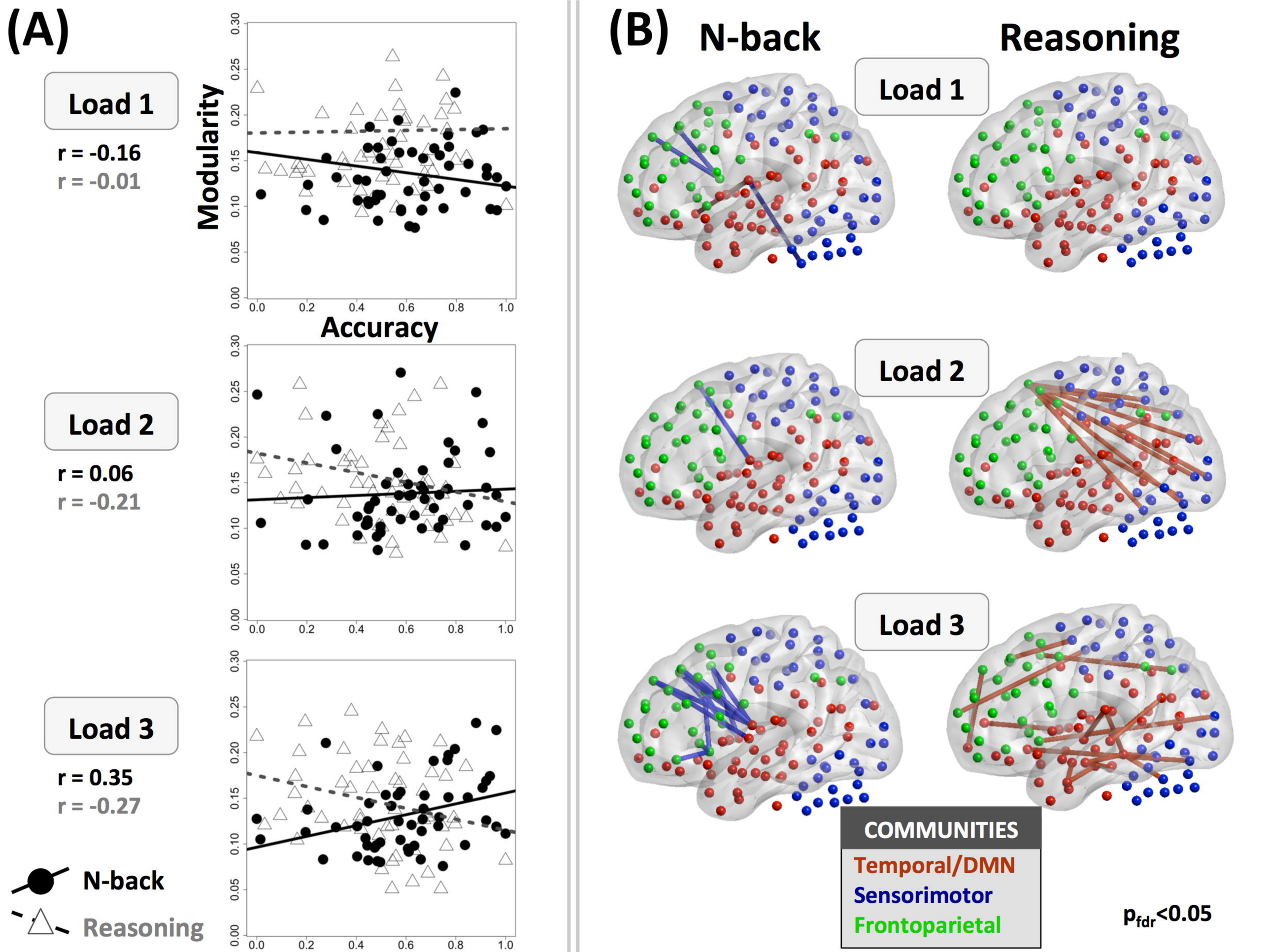


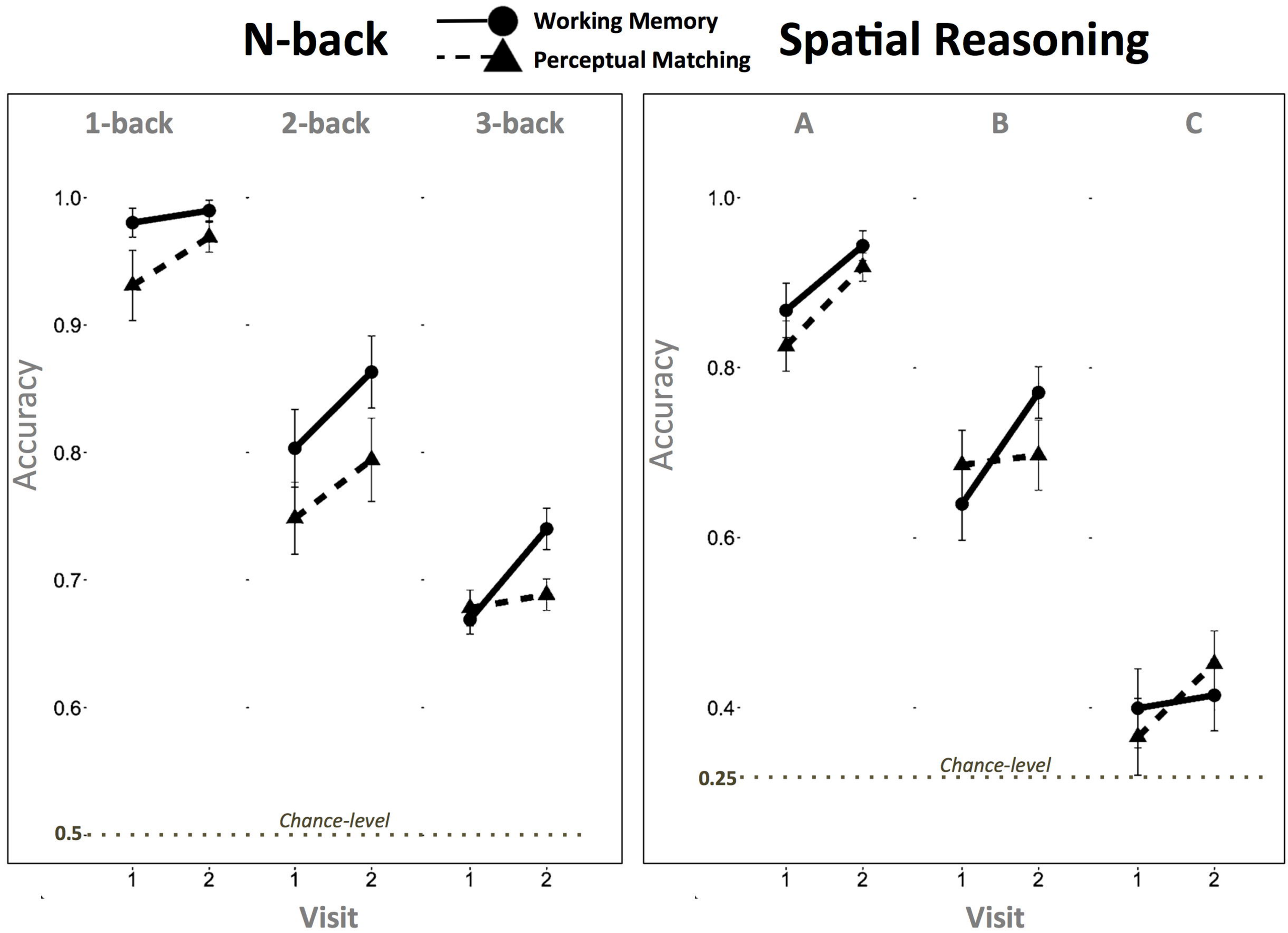

RAVEN
WASI
BETA
SR-offline
0.60
0.59
SR-fMRI
SR-B
SR-C
0.78
0.67
0.89
0.72
0.52
0.45
0.71
0.62
N-Back
RSpan1
RSpan2
UPD-offline
0.65
0.79
UPD-fMRI
2-Back
3-Back
Visit 1
(χ²=37.8, dof=29, CFI=0.95, SRMR=0.06)
Visit 2
(χ²=31.1, dof=29, CFI=0.99, SRMR=0.05)

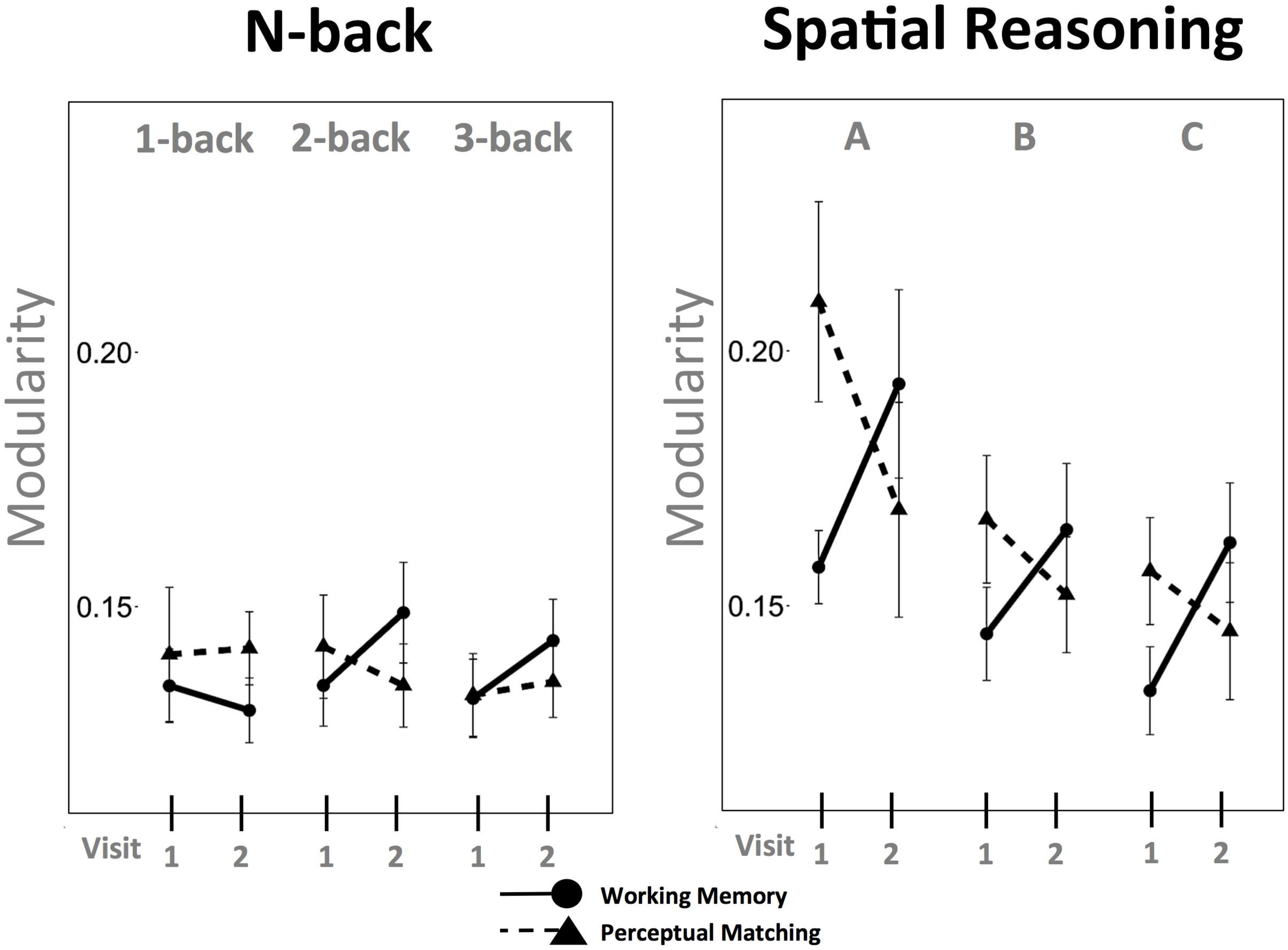
Task-related brain activation patterns

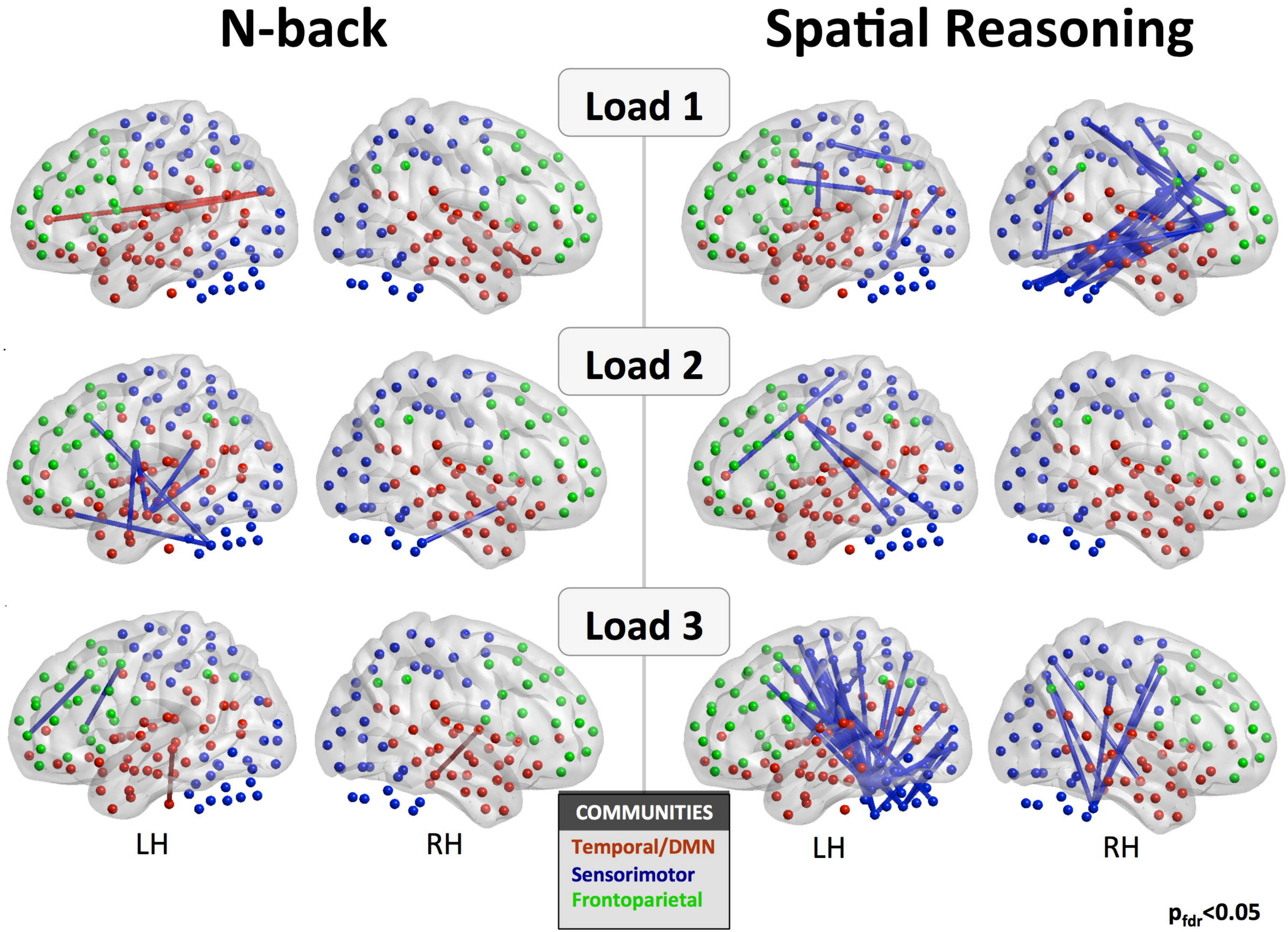



No Task
Load 1
Load 2
Load 3
Modularity
Difficulty Load
N-back
Reasoning
Good performers
Bad performers


(A)
Load 1
r = -0.16
r = -0.01
Load 2
r = 0.06
r = -0.21
Load 3
r = 0.35
r = -0.27
Modularity
Accuracy
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.0
0.2
0.4
0.6
0.8
1.0
N-back
Reasoning
(B)
N-back
Reasoning
Load 1
Load 2
Load 3
COMMUNITIES
Temporal/DMN
Sensorimotor
Frontoparietal
$p_{fdr}<0.05$


N-back
Working Memory
Perceptual Matching
Spatial Reasoning
1-back
2-back
3-back
A
B
C
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0.25
Chance-level
1
2
Visit


N-back
Spatial Reasoning
1-back
2-back
3-back
A
B
C
Modularity
0.20
0.15
Visit
1
2
Working Memory
Perceptual Matching


N-back
Spatial Reasoning
Load 1
Load 2
Load 3
LH
RH
LH
RH
COMMUNITIES
Temporal/DMN
Sensorimotor
Frontoparietal
$p_{fdr}<0.05$