NeuroImage 277 (2023) 120262

Contents lists available at ScienceDirect

NeuroImage

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

Instructional load induces functional connectivity changes linked to task automaticity and mnemonic preference

Alexander W. Baumann[a,*], Theo A.J. Schäfer[b], Hannes Ruge[a]

[a] Faculty of Psychology, Technische Universität Dresden, Dresden, Germany
[b] Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

ARTICLE INFO

Keywords:
Instruction-based learning
Functional connectivity
Automaticity
Load
Episodic memory
Working memory

ABSTRACT

Learning new rules rapidly and effectively via instructions is ubiquitous in our daily lives, yet the underlying cognitive and neural mechanisms are complex. Using functional magnetic resonance imaging we examined the effects of different instructional load conditions (4 vs. 10 stimulus-response rules) on functional couplings during rule implementation (always 4 rules). Focusing on connections of lateral prefrontal cortex (LPFC) regions, the results emphasized an opposing trend of load-related changes in LPFC-seeded couplings. On the one hand, during the low-load condition LPFC regions were more strongly coupled with cortical areas mostly assigned to networks such as the fronto-parietal network and the dorsal attention network. On the other hand, during the high-load condition, the same LPFC areas were more strongly coupled with default mode network areas. These results suggest differences in automated processing evoked by features of the instruction and an enduring response conflict mediated by lingering episodic long-term memory traces when instructional load exceeds working memory capacity limits. The ventrolateral prefrontal cortex (VLPFC) exhibited hemispherical differences regarding whole-brain coupling and practice-related dynamics. Left VLPFC connections showed a persistent load-related effect independent of practice and were associated with 'objective' learning success in overt behavioral performance, consistent with a role in mediating the enduring influence of the initially instructed task rules. Right VLPFC's connections, in turn, were more susceptible to practice-related effects, suggesting a more flexible role possibly related to ongoing rule updating processes throughout rule implementation.

1. Introduction

Learning through symbolic instructions is one of the most remarkable abilities within the human cognitive flexibility skillset (Cole et al., 2017; Liefooghe et al., 2018; Meiran et al., 2017). Especially the rapid timescale at which newly instructed rules can be encoded and executed accurately even at the very first attempt is impressive (Cole et al., 2013a). While in our daily lives we usually take this ability for granted, the underlying cognitive and neuronal processes have received increasing attention in the cognitive neuroscience domain over the past decade (Greve et al., 2017; Hartstra et al., 2011; Lee et al., 2015; Muhle-Karbe et al., 2017; Ruge and Wolfensteller, 2010).

A central requirement for such successful instruction-based learning (IBL; Ruge and Wolfensteller, 2010) or 'rapid instructed task learning' (RITL; Cole et al., 2013a) is the transformation of abstract, symbolic rule representations into pragmatic ones. This so-called 'symbolic-pragmatic transfer' (Ruge and Wolfensteller, 2010) or 'proceduralisation' (Brass et al., 2017) is not a trivial process as is illustrated by the phenomenon of goal-neglect (Bhandari and Duncan, 2014; Duncan et al., 2008) in which an instruction can be accurately perceived, encoded and even verbally repeated (i.e., memorized) – yet it cannot be implemented. Thus, a declarative rule representation that merely reflects the instructed stimulus-response (S-R) association without incorporation of a precise action plan seems to be insufficient for IBL to happen. This claim is supported by a variety of imaging studies that focused on the configuration and temporal evolution of rule representations following novel task instruction (González-García et al., 2020; Muhle-Karbe et al., 2017; Ruge et al., 2019; Sobrado et al., 2021).

Using multivariate pattern analysis (MVPA) techniques that are sensitive to subtle changes in activity patterns it has been suggested that in the lateral prefrontal cortex (LPFC) rule representations essential to the execution of newly instructed S-R associations can be decoded (Muhle-Karbe et al., 2017; Palenciano et al., 2019; Woolgar et al., 2015). This seems to be the case already at the first implementation trial and on the level of individual S-R bindings with a particular emphasis on the ventrolateral part (VLPFC) of the prefrontal cortex (Ruge et al., 2019). Interestingly however, these identity-specific activity patterns were not affected by differences in instruction load despite marked differences in

* Corresponding author.
E-mail address: alexander_willy.baumann@tu-dresden.de (A.W. Baumann).

https://doi.org/10.1016/j.neuroimage.2023.120262.
Received 10 February 2023; Received in revised form 5 June 2023; Accepted 29 June 2023
Available online 30 June 2023.

behavioral performance accuracy (Ruge et al., 2019). This finding contrasted with the original hypothesis – partly derived from earlier insights on rule integrity in error trials (Cole et al., 2016; Rigotti et al., 2013) – suggesting that higher instruction load would result in a weaker or incomplete prefrontal rule representation (and thus in a reduced identity-specific MVPA effect). This raises the question of whether the detected prefrontal rule representations actually contribute to the implementation of newly instructed S-R associations and if so in which manner.

Specifically, the degree to which instructed S-R associations are successfully implemented might depend not (only) on the integrity of prefrontal rule representations but also on how mnemonic processes related to information maintenance over time and retrieval interact with these representations. In turn, instructional load might affect the way prefrontal rule representations are accessed during instruction implementation rather than representational integrity itself (cf. Bartsch and Oberauer, 2022). From this perspective, a more adequate analytical approach is to ask how prefrontal rule representations might be differentially integrated with other brain networks that are related to storage and retrieval processes depending on instructional load. To address this question, the present paper specifically focused on differential functional connectivity patterns of IBL-critical prefrontal regions as a function of instructional load.

Connectivity-based approaches in the context of IBL and RITL have thus far focused mainly on global brain connectivity and large-scale networks (Cole et al., 2013b; Mohr et al., 2016, 2018). They have emphasized the importance of so-called control networks like the fronto-parietal network (FPN) and the dorsal attention network (DAN). FPN-regions are supposed to provide the basis of the flexible behavior essential to IBL by their tendency to act as 'flexible hubs' (Cole et al., 2013b, 2017). Depending on the current task demands they flexibly reconfigure their global functional connectivity (FC) patterns and seem to have a coordinating function (Cocuzza et al., 2020). It looks like such an involvement of the FPN is especially important during the very first implementation trials after novel instructions as evidenced by an increasing disengagement of the FPN – regarding activation as well as connectivity measures (Hampshire et al., 2019; Mohr et al., 2016). Furthermore, the rapidly evolving familiarity with newly instructed S-R rules has been linked to an increase in FC between prefrontal and striatal areas (Ruge and Wolfensteller, 2013, 2015).

While such studies have focused mostly on the temporal dynamics in IBL little is known about how interactions between brain regions are affected by changes of instructional load. So far, effects of load manipulation on FC have been studied extensively using standard working memory paradigms (e.g., Finc et al., 2017; Nour et al., 2019; Vatansever et al., 2017a). The concept of working memory (i.e., short-term maintenance and manipulation of currently relevant information e.g., D'Esposito and Postle, 2015) is relevant to rapid forms of IBL, both from a theoretical – e.g., due to the temporal proximity between instruction and implementation – and from an empirical perspective (Formica et al., 2020; Pereg et al., 2019; Pereg and Meiran, 2019). Yet, the exact mechanisms at work are still unclear (Monsell and Graham, 2021). Therefore, a connectivity-based approach examining load effects specifically in the context of newly instructed S-R rules (in contrast to 'pure' working memory tasks relying on continuous updating and/or manipulating of information) might further clarify the contribution of working memory to IBL.

Besides a potential involvement of working memory, other mnemonic processes reaching beyond the temporal and conceptual scope of WM have been discussed regarding their contribution to successful IBL. Monsell and Graham (2021) demonstrated that effective representations of S-R associations can persist for a considerable amount of time after only very few implementation trials following a novel instruction, and concluded that such representations might be stored in long-term memory (LTM). Another line of research focusing on 'automatic effects of instructions' has linked those exact effects to episodic memory retrieval (Meiran et al., 2017). From a neuroimaging perspective, a set of brain regions within a so-called 'core recollection network' has been implicated in such episodic memory retrieval processes (Rugg and Vilberg, 2013) and overlaps (at least partially) with another, well investigated large-scale network, the default mode network (DMN; e.g., Andrews-Hanna et al., 2014; Buckner and DiNicola, 2019).

The DMN, or more precisely, the way it interacts with other networks during task performance, has also been related to automatic processing (e.g., Vatansever et al., 2017b). In this regard, greater DMN-decoupling from other large-scale brain networks and increasing within-DMN coupling have been observed once task rules were fully established and could be applied in a relative 'automatic' fashion (Mohr et al., 2016, 2018; Vatansever et al., 2017b). Although involving little practice, and thus seemingly at odds with traditional views that emphasize the role of extensive practice (Logan, 1988; Schneider and Shiffrin, 1977), automaticity is a cornerstone in theories of IBL/RITL as well. Two aspects are frequently mentioned in this context: First, instructions are supposed to trigger 'automatic' (i.e., unintended) responses that are evident already directly after getting instructed (e.g., Liefooghe et al., 2012; Meiran et al., 2015) and seem to be quite resistant to cancelling (Abrahamse et al., 2022). Second, rules conveyed via explicit instructions seem to need only very little time to be established – as evidenced by rapid changes in network configuration (Mohr et al., 2016, 2018) as well as by plateauing behavioral effects after few implementation trials (e.g., Monsell and Graham, 2021; Ruge and Wolfensteller, 2010) – indicating rapidly evolving automaticity. Both aspects – the former referred to as 'intention-based reflexivity' (Meiran et al., 2012) and the latter as 'short-term automatization' (Chein and Schneider, 2005; Mohr et al., 2016) – seemingly result in similar (i.e., automatic) behavior, yet their differences are also emphasized (cf. Cole et al., 2017).

Working memory and long-term (episodic) memory are hypothesized to contribute to IBL, yet it remains relatively unclear how these systems are engaged and balanced in the service of successful learning. The rationale behind the work presented in this paper was to conceptualize instructional load as a variable that strains the capacity limits of one (here: working memory) of the involved systems which in turn might boost the engagement of the other (here: episodic memory) system (Bartsch and Oberauer, 2022). As argued further above, such load-dependent re-balancing might be reflected by altered functional integration of the lateral PFC into different functional networks supporting these two different memory systems. In addition, instructional load might also affect automaticity-related effects linked to IBL and should potentially be reflected in typical functional connectivity patterns. To this end, we revisited the data from experiment 1 originally reported in Ruge et al. (2019) and applied a seed-based connectivity approach. Our goal was to reconcile insights from pervious works focusing on prefrontal rule representation in IBL with knowledge about load-induced functional connectivity changes. Specifically, we focused on the functional coupling of 4 prefrontal seed regions which had been shown to play a distinctive role in cortical coding of newly instructed task rules.

2. Methods

2.1. Experiment and design

2.1.1. Sample

The original sample consisted of 68 participants of which 3 had to be excluded due to incomplete data collection. This resulted in a final sample of 65 participants (mean age: 24.2 years, age range: 19–33 years; 32 females, 33 males).

All participants were right-handed, neurologically healthy and had normal or corrected-to-normal vision. The experimental protocol was approved by the Ethics Committee of the Technische Universität Dresden and conformed to the World Medical Association's Declaration of Helsinki. The participants gave written consent before taking part in the experiment and were paid 10 Euro or compensated with course credit, respectively, for their participation.

2.1.2. Task and experimental procedure

Over the course of the experiment the participants had to complete multiple blocks of a rapid instruction-based learning paradigm. In each block there were two main phases: an instruction phase followed by an implementation phase. During the instruction phase of each block participants were presented with a set of novel stimulus-response (S-R) links (either 4 or 10). These mappings consisted of disyllabic German nouns (the stimuli) and the correct response button was instructed by spatially congruent cues. Since the pairing of each word with response buttons was unique (i.e., the same word was not paired with another button press over the course of the experiment) these pairings are referred to as 'rule identities'. To minimize potential reliance on 'shortcut strategies' the number of response options was varied between the blocks with either 2 or 3 responses being possible. Hence, the chance to infer a correct response from another mapping (e.g., 'if response A is not required here the other response has to be correct') was minimized – the memorization of the rule identity itself (e.g., 'this stimulus requires response B') was encouraged (cf. Liefooghe and Houwer, 2018). To prevent unintended interactions with the different experimental conditions the blocks requiring either 2 or 3 responses were equally distributed across the conditions. Furthermore, in blocks involving 2 different responses, an equal number of 2 stimuli (low load) or 5 stimuli (high load) was assigned to each response during the instruction phase. For blocks involving 3 responses, either 2 (low load) or 4 (high load) stimuli were assigned to one response and either 1 (low load) or 3 (high load) stimuli were assigned to each of the remaining two responses during the instruction phase. For the subsequent implementation phase the scheme of to-be-implemented rules in the high-load condition was identical to that determined by the low-load condition. Specifically, either two responses were each required twice (in blocks with 2 instructed responses) or one response was required twice and the other two responses were each required once (in blocks with 3 instructed responses). The behavioral implications of this assignment focusing on ambiguities at the response-selection level are reported in *supplementary analyses 1*. Each of the 3 response options (right, middle and left) was randomly selected to be assigned to 2 stimuli in the implementation phase in one third of the 18 blocks that involved 3 different responses. Per subject, words were randomly drawn from a wordlist and assigned to the different conditions.

Each instruction phase started after a variable delay of 2–4 s counting from the beginning of each measurement run or the end of a preceding implementation phase. At the start of each instruction phase the German word for 'memorize' ('Einprägen') was displayed in red color for 2 s on the screen. After this announcement novel nouns were shown on the screen in rapid succession (2000 ms per noun). These nouns were framed by two vertical bars which indicated the instructed response: if a noun was equally close to the two vertical bars a middle finger response was indicated, if it was closer to the right bar a ring finger response was indicated and if it was closer to the left bar an index finger response was indicated. Each noun was presented only once during the instruction phase and no responses were supposed to be executed within that time span. Instructed S-R rules were presented sequentially such that all S-R rules requiring a left (index finger) response were presented before all stimuli requiring a middle (finger) response which, in turn, were presented before all stimuli requiring a right (ring finger response). This chunked sequential presentation was chosen to keep the direction (left-to-right), number, and variability of response switches during the instruction phase relatively constant across both load conditions. Specifically, we thereby ensured that the frequency of response switches was kept within a small range between 1 and 2 equally for both load conditions. This compares to a range between either 1 and 3 (low-load) or 1 and 9 (high-load) for randomly presented S-R rules.

The implementation phase followed immediately after the instruction phase, announced by the German expression for 'implement' ('Ausführen') that was displayed on screen in green color for 2 s. The stimuli from the preceding instruction phase were presented in pseudo-random order without vertical bars framing them. Each stimulus was presented 4 times. In contrast to the instruction phase in which no responses were required, the participants were supposed to react to the stimuli by pressing the button that corresponded to the instructed response. The subjects were asked to use right hand fingers for responding and to avoid interference with handedness only right-handed subjects were recruited. No performance feedback was provided after single trials (to rule out interference with reinforcement learning), only at the end of each block the mean performance accuracy across all trials from that block was shown on screen. The stimulus onset asynchrony (SOA) interval varied between 2 and 4 s in steps of 0.5 s. It was placed before the start of a new trial. The performance feedback for each block was presented for 2 s after a variable delay of 2–4 s counting from the end of the block's last trial which indicated the end of the implementation phase.

The central experimental manipulation concerned instruction load and involved a 'low-load' condition and a 'high-load' condition: The low-load condition comprised 4 instructed task-rules and the high-load condition comprised 10 instructed task-rules. Importantly, the two conditions did not differ regarding the number of to be executed task-rules in the implementation phase which was always 4. While the implementation phase for the two conditions was identical, the instruction phase differed regarding number of stimuli and encoding time: In the low-load condition 4 nouns, each displayed for 2 s, were instructed, whereas 10 nouns, each displayed for 1 second, were instructed in the high-load condition (for an overview, see Fig. 1). By reducing the encoding time in the high-load condition the total instruction phase duration was approximately the same across the two conditions. Over the course of the experiment, 18 blocks of each load condition were completed, adding up to a total of 36 blocks. In each of the 3 consecutive scanning runs 6 blocks of each load condition were performed and each SOA level occurred equally often in the two load conditions.

Three consecutive scanning runs each lasting approximately 13 min were carried out adding up to a total duration of about 40 min. Before scanning, participants were familiarized with the task by completing a short practice session that included two task blocks (one for each experimental condition) outside the scanner. The task blocks from the practice session were not part of the final experimental task set.

2.1.3. Additional behavioral measures

After scanning was finished, participants were asked to perform a computerized simple digit span task to obtain a measure of individual short-term memory span (Wechsler, 1997): Random sequences of digits were displayed on a screen with each individual digit being shown for one second and appearing only once in a sequence. The presentation of a sequence was followed by the display of as many question marks as there were digits in the preceding sequence and subjects were prompted to reproduce the sequence either in forward or backward order. If a subject reproduced a sequence correctly it advanced to the next, longer sequence. The first sequence consisted of 3 digits and the number of digits was increased by one with each sequence displayed (up to 10). If a subject made an error a new sequence of the same length was displayed, and if this new sequence was reproduced incorrectly again the test stopped. The length of the last successfully reproduced sequence (i.e., maximal number of correctly remembered digits) constituted the final score. At last, participants performed a short computerized version of the standard progressive matrices intelligence test (Raven, 2003) in which only the two most difficult matrix sets out of all five sets were used (D and E). Each set consisted of 12 matrices that were presented in progressively difficult order. The sum of all correctly solved matrices constituted the non-standardized intelligence score.

2.2. fMRI-data acquisition and preprocessing

2.2.1. Acquisition

A Siemens 3T whole body Trio System (Erlangen, Germany) with a 32 channel head coil was used for data acquisition. Ear plugs were

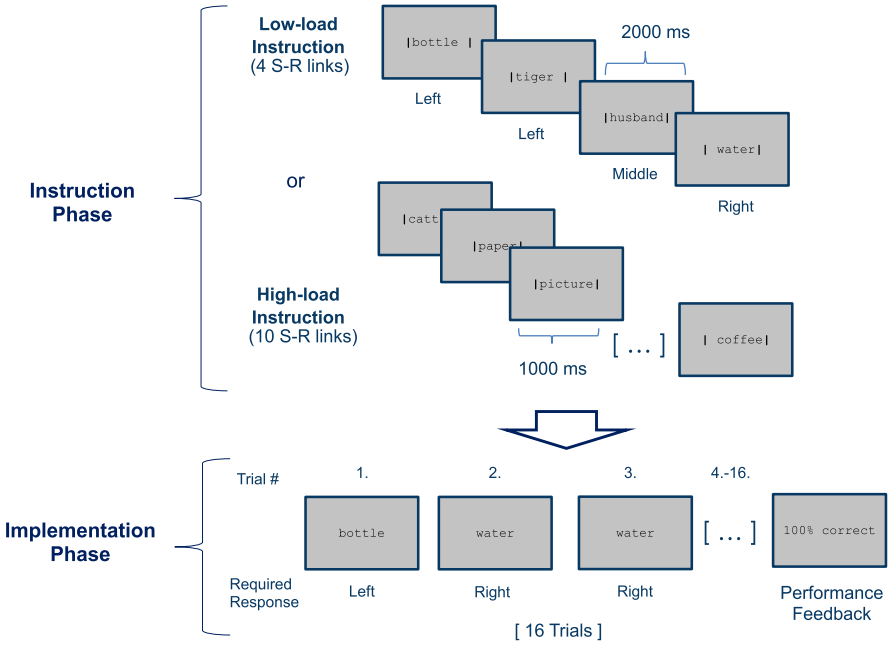


Fig. 1. Overview of the experimental paradigm. Each learning block comprised an instruction phase and an implementation phase. Novel S-R links were introduced in the instruction phase with vertical bars next to the verbal stimuli indicating either an index finger (left), a middle finger (middle) or a ring finger (right) response. The low-load condition comprised 4 novel S-R links (each presented for 2000 ms) whereas the high-load condition comprised 10 novel S-R links (each for 1000 ms). The newly instructed S-R mappings were not to be executed during the instruction phase. Actual response execution was required exclusively in the implementation phase and regardless of the instructional load condition, always only 4 S-R links had to be implemented four time each per learning block. No vertical bars next to the verbal stimuli were displayed during the implementation phase. Summary performance feedback was shown only at the end of each learning block.

used to dampen the scanner noise. Structural images were acquired after the experimental session using a T1-weighted sequence (TR = 1900 ms, TE = 2.26 ms, TI = 900 ms, flip = 9) with a resolution of 1 mm x 1 mm x 1 mm. For the acquisition of functional images a gradient echo planar sequence (TR = 2000 ms, TE = 30 ms, flip angle = 80°) was used. Each volume contained 32 slices that were measured in ascending order. The voxel size was 4 mm x 4 mm x 4 mm (gap: 20%). Additionally, field maps with the same spatial resolution as the functional images were acquired to correct for inhomogeneity of the static magnetic field (TR = 352 ms, short TE = 5.32 ms, long TE = 7.78 ms, flip angle = 40°). The experiment ran on E-Prime 2.0.

2.2.2. Preprocessing

After data acquisition the fMRI data were preprocessed using SPM12 running on MATLAB R2016a. The functional images were slice-time corrected at first, then spatially realigned and unwarped using the acquired field maps. The structural image from each participant was co-registered to the mean functional image and segmented. Afterwards, spatial normalization was performed by applying the deformation fields that were generated in the segmentation process to the functional images (resolution: 3 mm x 3 mm x 3 mm). Except for the activation analyses, images were not smoothed before general linear model (GLM) estimation. Subjects' whole-brain connectivity maps were smoothed with 6 mm FWHM before being submitted to group level analysis.

2.3. Task-related fMRI analysis

Our primary goal was to assess load-dependent coupling of the LPFC during instruction implementation. To this end, we employed the beta-series correlation approach (Rissman et al., 2004). Since this approach requires single-trial modeling we will first describe how single-trial GLMs were estimated. After that, the beta-series correlation procedure itself will be introduced. At last, we will describe the classical mean activation analysis that was conducted as a complementary analysis.

2.3.1. Single-trial bold estimation

To estimate voxel-wise bold activation for each single trial the General Linear Model approach implemented within the SPM12 framework was employed. A first-order auto-regressive model was used and a high-pass filter of 1/128 Hz was included to account for slow signal drifts.

Single-trial GLMs were used to model the BOLD activations during the implementation phase. The least-squares-separate (LSS) modeling approach (Mumford et al., 2011, 2014) was chosen in which one regressor is modeling the specific trial and additional regressors are added to model the mean activity for all other trials according to condition. Specifically, for each model, in addition to the single trial regressor, two regressors were included to model correct trials separately for low-load trials and high-load trials. Similarly, two regressors were added to model error trials again separately for the two load conditions. Additional regressors were included to model activity associated with performance feedback presented at the end of each implementation block separately for the two load conditions. The regressors were created using stick functions convolved with SPM12's default canonical hemodynamic response function that were synchronized to the onsets of each implementation trial or the onset of the feedback display. To model the activity associated with the instruction phase, Fourier basis set regressors time-locked to the start of the instruction phase (including 20 different sinewave regressors spanning 44 s) were used to appropriately account for BOLD activation during the instruction phase that lasted either 12 (low-load) or 14 (high-load) seconds.

In this fashion, as many independent LSS-GLMs were estimated as there were trials, adding up to a total of 576 individual GLMs (36 blocks times 16 trials per block).

2.3.2. Beta-series correlation

The single-trial estimates computed as described above were used for the beta-series correlation analysis (Abdulrahman and Henson, 2016; Rissman et al., 2004) that was employed to assess connectivity between various seed regions and all other voxels in the brain. A trial-based connectivity measure was used as we intended to (a) mainly investigate processes directly related to instruction implementation of individual S-R rules and (b) effectively compare load-dependent effects specifically related to correct trials. This favourably compares to alternative block-related connectivity measures which would confound low load vs. high load conditions with fewer errors vs. more errors, respectively. The basic idea of this approach is to use the beta estimates from the single-trial GLMs to form a series of beta estimates for each voxel. These estimates can then be sorted along the different task conditions (here, this could be *instructional load* or *stimulus repetition*) such that there is a beta-series for each of the relevant conditions. The correlation between the beta-

series of a given voxel (or the average beta-series of all voxels within a ROI) and (all) other voxels (or, again, the average time series within other ROIs) can be viewed as a measure of the functional connectivity between specific voxels/ROIs. This results in a connectivity map where the value assigned to each voxel represents its connectivity with the seed region or voxel during a specific task condition. Differential connectivity between two task conditions can then be computed by contrasting the respective connectivity maps.

Here, the mean beta-series for the seed ROIs (see below) were extracted for each task condition and the Fisher-z (Fisher, 1921) transformed Pearson correlation was computed between the seed region's beta-series and the beta-series of each brain voxel. Connectivity maps for each condition, each seed region, and each subject were obtained in this way. Only the correct trials (for error-coding, see Section 2.3.6) were included in the analyses.

The single subject's connectivity maps were then entered into a second level GLM and were analyzed at the group level via one-sample t-tests. Whole-brain analyses using SPM12 running on MATLAB R2018a were performed to identify the regions exhibiting differential functional connectivity with the seed regions between the respective conditions (e.g., high-load/low-load). As one explicit goal of this work was to follow up on earlier MVPA findings, the four primary seed ROIs were the same as those used in Ruge et al. (2019): the left/right VLPFC as well as the left/right DLPFC. Those ROIs were adapted from the automatic anatomical labeling (aal) atlas (Tzourio-Mazoyer et al., 2002): aal-regions 'inferior frontal gyrus pars opercularis' and 'inferior frontal gyrus pars triangularis' of each hemisphere were combined for the VLPFC-seed, aal-region 'middle frontal gyrus' was selected for the DLPFC seed. Unless indicated differently all whole brain analyses were corrected for multiple comparisons using family-wise error correction at the cluster-level, with pFWE < 0.05 and an initial threshold of $p_{uncorr.} < 0.001$.

2.3.3. Standard activation analysis

Although we focus mainly on connectivity in this work, a standard whole-brain activation analysis was conducted as well. Since the relationship between activation and connectivity is not necessarily straightforward (Di and Biswal, 2019) their complementary evaluation can potentially reveal interesting insights (Gerchen and Kirsch, 2017).

The activation analysis was performed on the preprocessed and smoothed (6 mm FWHM) fMRI-data. In contrast to the single-trial GLMs that constituted the basis for our connectivity analyses a conventional modeling approach was used here. The 8 condition-specific regressors defined by the two conditions 'instructional load' (low-load and high-load) and 'stimulus repetition' (1 to 4) were created by convolving stick functions synchronized to trial onsets (correct ones only) with SPM12's default canonical HRF. Regressors of no-interest – error-trials, very first trial per block and feedback after each block – were created in the same way. Just like for the single-trial GLMs the instruction phase was modelled using Fourier basis set regressors. To capture the constant activity across each implementation phase additional regressors were created by convolving a boxcar function synchronized to onset and duration of the implementation blocks with the canonical HRF

Since the study design effectively equals a mixed block/event-related design (Dosenbach et al., 2006; Visscher et al., 2003) we were interested in both the constant activation across each implementation block and the activation related to the specific trials. Therefore, contrast images for the instructional load conditions were computed at the first level using either the betas from the implementation block regressors or from the trial-specific regressors (aggregated across stimulus repetitions). Both kinds of contrast images were then evaluated separately at second level using one-sample t-tests.

2.3.4. Follow-up connectivity analyses

To put the findings regarding the load-dependent connectivity profiles of our selected LPFC seeds into perspective we decided to conduct follow-up whole-brain connectivity analyses. To this end, we used some of the regions identified via the previous LPFC-seeded whole-brain connectivity analyses themselves as new seeds for further whole-brain BSC analyses. Since there was a considerable overlap between the connectivity profiles evoked by the four primary LPFC seeds we accounted for this overlap by creating 'combined maps': Clusters of increased connectivity that survived the whole-brain correction were extracted such that binary whole-brain maps for each seed and each contrast direction (i.e., low-load>high-load and high-load>low-load) were obtained (1=voxel within significantly connected clusters, 0=otherwise). The resulting four binary maps per contrast direction were combined (logical or), resulting in one 'combined connectivity map' (i.e., each voxel had been part of significantly coupled cluster for at least one of the 4 primary seeds) for each contrast. Clusters of contiguous voxels were then extracted from those maps, resulting in specific 'combined clusters' for each contrast direction that served as seeds for our follow-up whole-brain connectivity analyses.

2.3.5. Correlations with behavioral variables

To assess the behavioral relevance of the connectivity patterns we computed the correlation between the prefrontal connections detected in the preceding analysis steps and the behavioral measures obtained during or after the experiment.

We chose an ROI-based approach in which we used the 'connections' (between seeds and significantly coupled regions) as defined in Section 3.3.1. Average differential beta-series correlation values for each connection were extracted from subjects' contrast maps. Afterwards, we computed the Pearson product-moment correlation as well as the Spearman rank-correlation coefficient between each extracted connection and trait-like behavioral measures (total scores of Raven's progressive matrices, forward digit span and backward digit span) as well as first trial accuracy (low-load, high-load, and low-load minus high-load). Correlations were corrected for multiple testing by controlling the false discovery rate (FDR; Benjamini and Hochberg, 1995). The amount of correlations the FDR was controlled for was given by the number of connections detected per seed (variable) times the number of behavioral variables taken into account for analysis (always 6). FDR correction was therefore implemented individually for each of the 4 seeds. Only correlations that survived this correction procedure are reported.

2.3.6. Definition of correct and erroneous responses

Due to the absence of response feedback during individual implementation trials, participants could not know whether or not they had correctly retrieved the instructed response for a certain stimulus. Hence, it seems a reasonable strategy to stick with the response that was executed for the same stimulus in a previous implementation trial assuming it was the originally instructed one. Accordingly, we classified a response as 'correct' when it matched the response given in the preceding trial for the same stimulus. In the special case of the first implementation trial, a response was classified as 'correct' when it matched the originally instructed response for a given stimulus, thus an 'objective' measure of successful instruction implementation was obtained in this case. Only trials coded as 'correct' in that manner were used for the fMRI data analyses. For a depiction of accuracy rates based on an alternative accuracy definition (i.e., accuracy as measured by the instructed response at all 4 stimulus repetitions), please see supplementary figure 1.

2.3.7. Conflict trials

The absence of direct performance feedback also creates a specific type of potential conflict: If the implemented S-R rule does not match the instructed S-R rule (e.g., the stimulus 'picture' requires a middle finger response but an index finger response is executed) this just implemented S-R rule likely becomes the new 'self-generated' correct rule in the upcoming trial with the same stimulus ('picture' would now require an index finger response to be classified as correct). Since this

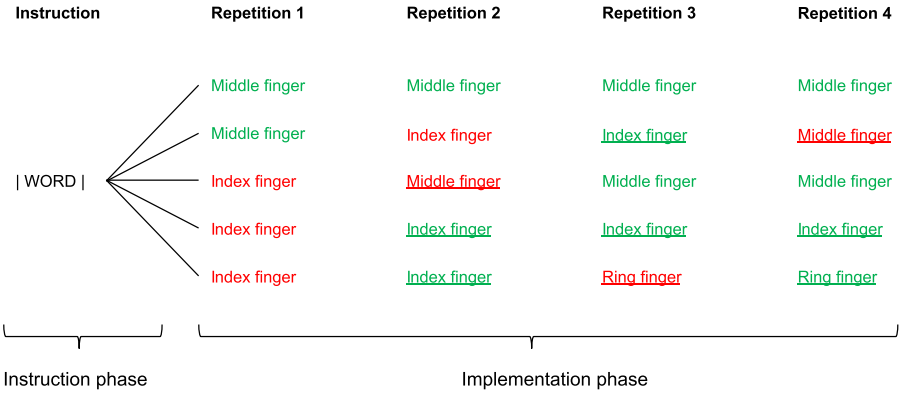


Fig. 2. Illustration of several possible response trajectories across the implementation phase. Correct responses (see Section 2.3.6 for error definition) are printed in green and incorrect responses are printed in red. Conflict trials (i.e., trials in which instructed and previous response to the same stimulus do not match) are underlined. The illustration is a non-exhaustive selection of a total of 3^4 = 81 possible trajectories.

self-generated S-R rule and the originally instructed S-R rule ('picture' still requires a middle finger response) are conflicting, this upcoming trial is labelled as *conflict trial* (see Fig. 2 for a schematic illustration). Behavioral performance in conflict trials provides an indication of the extent to which the original instruction is still active even if it had not been implemented correctly at some point in the implementation phase. Please note, that a conflict trial can only occur at stimulus repetitions 2 to 4 as in stimulus repetition 1 there is no previously executed S-R rule.

3. Results

3.1. Behavioral performance

3.1.1. Mean response times and accuracy

Response accuracies and reaction times (RTs) were already reported in Ruge et al. (2019). Since they are at least partly relevant for the analyses conducted here, the results will be briefly summarized in this section.

Repeated measures ANOVAs were computed with the independent variables *stimulus repetition* and *instructional load*. Accuracies (see Fig. 3a) were found to be generally higher in the *low-load* condition compared to the *high-load* condition as indicated by the main effect of *instructional load* ($F_{1,64}$ = 202.81; $p(F) < 0.001$; η_p^2 = 0.76) and increased linearly across *stimulus repetitions* ($F_{3,192}$ = 71.43; $p(F) < 0.001$; η_p^2 = 0.53; linear contrast $F_{1,64}$ = 111.52; $p(F) < 0.001$; η_p^2 = 0.64). The increase of accuracy was more pronounced in the *high-load* than in the *low-load* condition as indicated by the interaction of *instructional load* by *stimulus repetition* ($F_{3,192}$ = 80.14; $p(F) < 0.001$; η_p^2 = 0.56; linear contrast $F_{1,64}$ = 156.40; $p < .001$; η_p^2 = 0.71). Accuracies differed significantly between conditions even at the fourth *stimulus repetition* with the performance in *low-load* trials still being better than that in *high-load* trials (t = 6.26; $p(t) < 0.001$).

The same behavioral pattern was found when considering reaction time (RT). A significant main effect of *instructional load* indicated generally higher RTs in the *high-load* condition ($F_{1,64}$ = 175.15; $p(F) < 0.001$; η_p^2 = 0.73). A linear RT decrease from the first towards the fourth *stimulus repetition* was indicated as well ($F_{3,192}$ = 224.87; $p(F) < 0.001$; η_p^2 = 0.78; linear contrast $F_{1,64}$ = 290.67; $p(F) < 0.001$; η_p^2 = 0.82). This RT decrease was more pronounced in the *high-load* than in the *low-load* condition ($F_{3,192}$ = 137.94; $p(F) < 0.001$; η_p^2 = 0.68; linear contrast $F_{1,64}$ = 252.69; $p(F) < 0.001$; η_p^2 = 0.80) but RTs were still longer in the *high-load* than in the *low-load* condition at repetition 4 (t = 4,60; $p(t) < 0.001$).

3.1.2. Correlation between accuracy and test scores

Correlations between mean accuracy across all stimulus repetitions and the progressive matrices score were positive for the low-load instructions (r = 0.32; p = .009, two-tailed) as well as for the high-load instructions (r = 0.35; p = .005, two-tailed). Correlations with progressive matrices score were still present in the low-load condition (r = 0.26; p = .037, two-tailed) and the high-load condition (r = 0.28; p = .022,

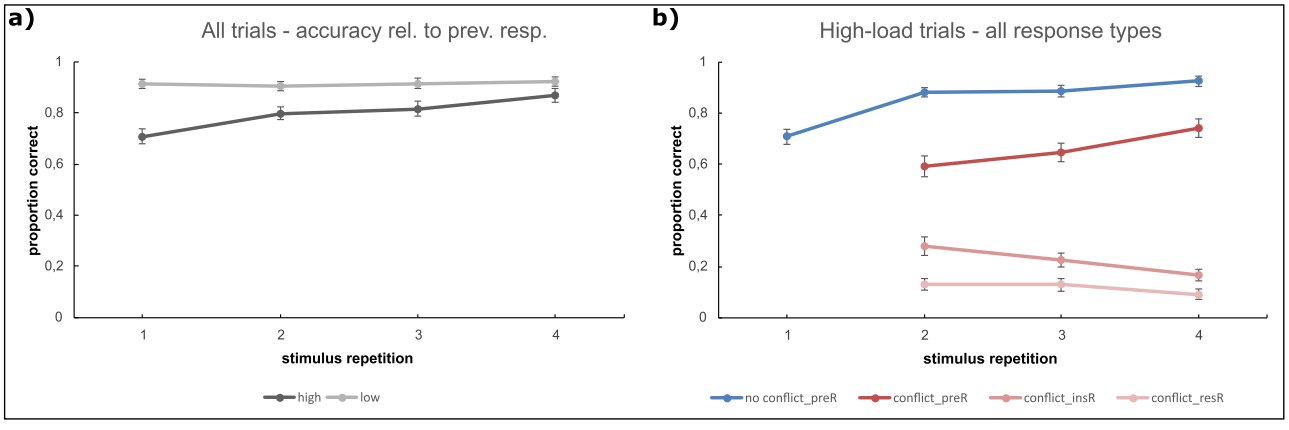


Fig. 3. a) Behavioral performance as measured by response accuracy (proportion correct). Response accuracy was defined relative to the previously executed response in repetitions 2 to 4 and relative to the originally instructed response in repetition 1. b) Response accuracy in conflict trials specifically for the high-load condition. The 'no conflict_preR' condition indicates accuracy relative to the previously executed response in non-conflict trials. The 'conflict_preR' condition indicates the same for conflict trials. The statistical analysis reported in 3.1.3 was based on these two conditions. Complementarily, the chart depicts conflict trial accuracy defined relative to the originally instructed response ('conflict_insR') as well as defined relative to the remaining third response option that was neither instructed nor previously executed ('conflict_resR'). Please note, that for non-conflict trials accuracy relative to the previously executed response equals accuracy relative to the originally instructed response. Error bars represent 95% confidence intervals.

two-tailed) when selectively analysing first implementation trial accuracy. No correlations were observed between accuracy rates and forward/backward digit span scores when considering all trials of a condition (all $p > .277$, two-tailed). When only considering the first implementation trial per stimulus of each condition, however, there was a significant positive correlation observed between forward digit span score ($r = 0.27$; $p = .032$, two-tailed) and low-load condition accuracy with a trend in the same direction in the high-load condition ($r. = 0.21$; $p = .087$, two-tailed).

3.1.3. Conflict vs. non-conflict trials

The behavioral results reported above raised the question why accuracy remained at a lower level in the high-load than in the low-load instruction condition throughout the whole implementation phase. This is somewhat remarkable as from the second stimulus repetition onwards, there is no obvious difference between the conditions (i.e., 4 S-R rules have to be implemented within the same time period). A possible explanation might be provided by the analysis of conflict trials which occur more frequently in the high-load condition (about 30% of all trials per stimulus repetition) whereas in the low-load condition accuracy levels ceiled already at stimulus repetition 1 resulting in very few conflict trials. Specifically, a lower accuracy in conflict trials than in non-conflict trials would indicate a persisting (negative) impact of the originally instructed S-R rule even though a successful transfer from instruction to implementation had originally failed in the first implementation trial. The impact of conflict was analysed for stimulus repetitions 2 to 4 (as by definition there is no conflict at stimulus repetition 1) and we focused on the high-load condition as there were too few conflict trials in the low-load condition. As before, accuracy was defined with respect to the previously *executed* response for a given stimulus. A 2-by-3 ANOVA with factors *conflict* and *stimulus repetition* yielded highly significant main effects of *conflict* ($F_{1,64} = 412.21$; $p(F) < 0.001$; $\eta_p^2 = 0.87$) and *stimulus repetition* ($F_{2,128} = 45.51$; $p(F) < 0.001$; $\eta_p^2 = 0.41$) as well as a significant interaction of both factors ($F_{2,128} = 12.84$; $p(F) < 0.001$; $\eta_p^2 = 0.17$). As shown in Fig. 3b, this reflects, as hypothesized, an overall higher accuracy in *non-conflict trials* ('no conflict_preR') than in *conflict trials* ('conflict preR'). While the accuracy increase towards the end of the implementation phase was more pronounced in *conflict trials* than in *non-conflict trials*, it is apparent that the *conflict trial* accuracy never reached the same asymptotic level as in *non-conflict* trials. This clearly shows that the persisting accuracy difference between the low-load and the high-load condition shown in Fig. 3a is indeed due to conflict trials. Moreover, Fig. 3b also shows that errors in conflict trials are mainly due to choosing the originally instructed response ('conflict_insR') rather than the remaining third response option that was neither instructed nor previously executed ('conflict_resR'). Hence, even if subjects seem to have forgotten the originally instructed response as reflected by high error rates in the first implementation trial (repetition level 1), the above-chance tendency to switch back to the originally instructed response later on, suggests that this was in fact not the case.

3.2. Standard analysis of transient and sustained bold activity

First, we conducted a standard activation analysis to assess the whole brain activation differences regarding the main effects of load (sustained and transient) and time (see Table 1)

Table 1
Results of all activation analyses. Peak voxels for each cluster are reported in MNI space.

Condition	Contrast	pFWE cluster	Voxels	pFWE peak	Tpeak	x / y / z	Regions
LOAD-sustained	LOW > HIGH	< 0.001	539	.007	5.80	−6 / 29 / −4	Bilat. MFC
		.032	48	.396	4.60	−51 / −73 / 29	Left MTG/AG
		.032	48	.888	4.08	−6 / −58 / 14	Left Prec./PCC
	HIGH > LOW	< 0.001	175	< 0.001	7.51	−30 / 23 / −1	Left ant. Ins.
		< 0.001	1395	< 0.001	7.25	−33 / −55 / 38	Left IPL/sup. Prec.
		< 0.001	641	< 0.001	6.81	45 / 32 / 32	Right MFG
		< 0.001	181	< 0.001	6.77	30 / 26 / −1	Right ant. Ins.
		< 0.001	252	.010	5.72	6 / 20 / 47	Bilat.SMA
		< 0.001	259	.012	5.67	−33 / 62 / 8	Left MFG
		< 0.001	315	.031	5.42	−48 / 20 / 32	Left IFG/MFG
		.003	78	.292	4.72	12 / 2 / 8	Right Thal.
		.003	78	.445	4.54	−12 / −13 / 11	Left Thal.
		< 0.001	124	.652	4.34	27 / 11 / 47	Right MFG/SFG
LOAD-transient	LOW > HIGH	< 0.001	424	.035	5.39	−57 / −19 / 23	Left STG/Oper.
		< 0.001	136	.399	4.60	60 / −13 / 17	Right SMG/STG
	HIGH > LOW	.001	88	.365	4.64	−42 / −67 / 47	Left AG
		.006	66	.439	4.56	−36 / 11 / 59	Left MFG
		.037	45	.526	4.48	45 / 17 / 35	Right IFG
		< 0.001	144	.533	4.47	−6 / 41 / 38	Bilat. MFC
Time	EARLY > LATE	< 0.001	12,824	< 0.001	11.69	−33 / −58 / 41	Front./par. cortex
				< 0.001	11.63	−33 / 20 / −1	Left ant. Ins.
				< 0.001	11.17	33 / 20 / −4	Right ant. Ins.
		< 0.001	1273	< 0.001	9.09	9 / −79 / −31	Right Cer.
		.002	103	< 0.001	7.13	57 / −46 / −10	Right ITG/MTG
		< 0.001	381	< 0.001	6.61	−51 / −46 / −10	Left MTG/ITG
		.011	72	.432	4.49	−18 / −25 / 59	Left para. lob.
		.006	81	.442	4.48	18 / −25 / 59	Right para. lob.
	LATE > EARLY	< 0.001	793	< 0.001	7.19	30 / −85 / 32	Right occ. cortex
		< 0.001	457	< 0.001	6.96	−3 / 35 / −10	Orb.MFC
		< 0.001	243	< 0.001	6.80	42 / 5 / −34	ITG/temp. pole
		< 0.001	255	.002	6.16	−6 / 62 / −1	Bilat. MFC
		< 0.001	457	.002	6.09	−9 / −94 / 32	Left occ. Cortex
		.001	112	.027	5.46	−45 / 5 / −37	Left ITG/temp. pole
		.050	50	.291	4.66	27 / −67 / −7	Right fusiform gyrus

Abbreviations: ant. = anterior; bilat. = bilateral; AG = angular gyrus; cer. = cerebellum; front. = frontal; IFG/MFG/SFG = inferior/middle/superior frontal gyrus; ins. = insula; IPL = inferior parietal lobule; ITG/MTG/STG = inferior/middle/superior temporal gyrus; mid. cing. = middle cingulate; MFC = medial frontal cortex; occ. = occipital; orb. = orbital; oper. = operculum; par. = parietal; para. lob. = paracentral lobule; PCC = posterior cingulate gyrus; prec. = Precuneus; SFG = superior frontal gyrus; SMA = supplementary motor area; SMG = supramarginal gyrus; sup. = superior; temp. = temporal; thal. = thalamus).

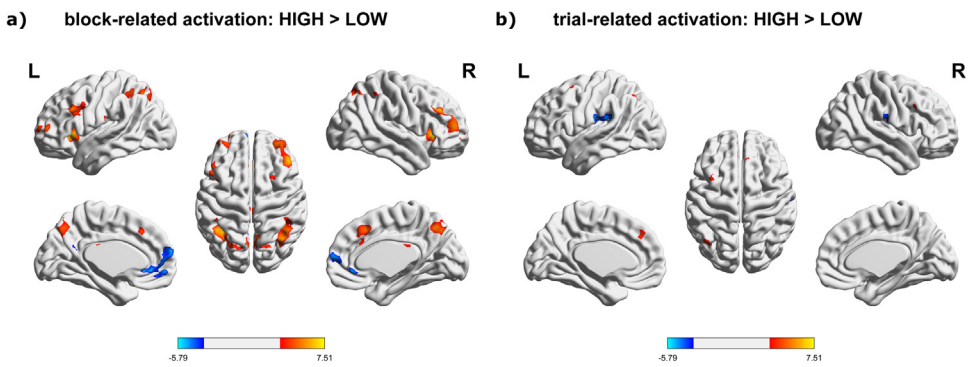


Fig. 4. Visualization of the standard activation analysis results. (a) Block-related load-dependent activation (b) trial-related load-dependent activation; positive t-values represent areas in which activation was greater in the high-load > low-load contrast and are illustrated in warm colors; negative t-values represent areas in which activation was greater in the low-load > high-load contrast and are illustrated in cold colors.

3.2.1. Effect of instructional load*: block-related sustained activation*

Only few areas, most notably medial prefrontal cortex, displayed greater sustained activation in the *low-load* as compared to the *high-load* condition. When considering the opposite contrast, *high-load* vs. *low-load*, widespread sustained activation was detected in lateral prefrontal, insular and posterior parietal as well as thalamic regions of both hemispheres (see Fig. 4a and Table 1). Strong block-related sustained effects seem to be in accordance with a persistent load-related difference in accuracy and RT (see 3.1).

3.2.2. Effect of instructional load*: trial-related transient activation*

Regarding transient event-related effects of load (see Fig. 4b and Table 1), two clusters around the bilateral parietal/temporal opercula were identified to be more strongly activated during the *low-load* than during the *high-load* condition. For the opposite contrast, *high-load* vs. *low-load*, significant activation clusters were found in lateral/medial frontal areas and around the left angular gyrus.

Discrepancies between block-related and trial-related activation might be due to power differences in favor of the block-related regressor (with a considerable share of load-dependent activation being captured by it already). Furthermore, block-related activation might also comprise (sustained) processes induced by error trials which occur more frequently in the *high-load* condition.

3.2.3. Effect of time

Besides cerebellar and posterior temporal sites, a very large cluster of voxels spanning posterior parietal and frontal regions of both hemispheres with activation foci in the anterior insulae and around angular gyri was found to be more strongly activated during the *early* part (stimulus repetitions 1 and 2) of the implementation phase as compared to the *late* part (stimulus repetitions 3 and 4). For the reversed contrast, *late* vs. *early* increased activation was predominantly detected in the bilateral occipital cortices and to a lesser degree around bilateral temporal poles and ventromedial PFC (Table 1 and supplementary figure 3).

3.3. Whole-brain connectivity: beta-series correlation for primary seeds

Our central goal in this work was to assess the load-dependent functional coupling of the 4 prefrontal seeds used in Ruge et al. (2019). As described above this was done by computing the beta-series correlation between each specific seed region and all other brain voxels. Subsequently, significantly coupled voxels were assigned to large-scale networks based on the parcellation by Yeo et al. (2011). Table 2 provides an overview for all 4 primary seed regions.

3.3.1. Whole brain: effect of instructional load

3.3.1.1. Left DLPFC seed. The left DLPFC (see Fig. 5a) was more strongly coupled to a cluster around the left inferior frontal gyrus (IFG) in the *low-load* as compared to the *high-load* condition. For the opposite contrast direction, *high-load* versus *low-load*, increased coupling was observed with ventral precuneus/posterior cingulate cortex (PCC), medial frontal cortex (MFC) and left angular gyrus (AG).

3.3.1.2. Right DLPFC seed. The right DLPFC (see Fig. 5b) exhibited increased coupling to an area around the right supramarginal gyrus (SMG) for the *low-load* vs. *high-load* contrast. For the reversed contrast (high-load versus low-load) stronger coupling to ventral precuneus/PCC, MFC and bilateral AG (reaching into posterior temporal and occipital cortices) was detected.

3.3.1.3. Left VLPFC seed. Left VLPFC (see Fig. 5c) was more strongly coupled to both left and right SMG regions for the *low-load* vs. *high-load* contrast. Other regions of increased coupling for the same contrast were localized around the bilateral orbitofrontal gyri (OFG), the bilateral supplementary motor areas (SMA), the left posterior insula (INS) and within the left cerebellum. Only one cluster covering the ventral precuneus/PCC was found to be more strongly connected in the *high-load* as compared to the *low-load* condition.

3.3.1.4. Right VLPFC seed. For the *low-load* vs. *high-load* contrast multiple locations of strengthened coupling to the right VLPFC seed (see Fig. 5d) were identified in and around: right anterior OFG, bilateral IFG, right middle frontal gyrus (MFG), right SMA, left cerebellum, right anterior insula (aIns), right and left posterior parietal cortex around the intraparietal sulcus (IPS). For the opposite contrast, *high-load* vs. *low-load*, the junction of ventral precuneus/PCC, again, was increasingly coupled as were MFC and the bilateral AG.

3.3.2. Connectivity overlap between primary seed regions

At first glance, results of the beta-series correlation analyses appear similar across all 4 prefrontal seeds. Comparing the *low-load* condition to the *high-load* condition, we observed significantly increased coupling to regions predominantly assigned to large-scale networks such as the FPN and the dorsal/ventral attention network (DAN/VAN). Areas exhibiting increased coupling to one of the 4 prefrontal seeds in the high-load condition compared to the low-load condition, in turn, were almost exclusively assigned to the DMN. However, while this pattern was consistent across all 4 prefrontal seeds there was still considerable variation in location and extent of the specific connectivity foci. In an attempt to capture and quantify this variation we performed pairwise second level t-tests in which we compared subjects' load-related connectivity contrast maps between the 4 seeds. Each seed was compared against each other seed resulting in a total of 6 paired t-tests (see supplementary Table 1 and supplementary Fig. 4).

Interestingly, especially the left VLPFC's connectivity pattern differed markedly and consistently from that of all 3 remaining seed regions' connectivity patterns. In particular, the posterior parietal, temporal and medial frontal areas showed deviating load-related functional couplings specifically with the left VLPFC seed. Although differences in load-induced whole-brain connectivity changes were also observed between some of the remaining three pairwise comparisons between seed regions (i.e., left DLPFC vs. right DLPFC, left DLPFC vs. right VLPFC, right DLPFC vs. right VLPFC) no systematic pattern emerged comparable to the one observed for left VLPFC-specific effects.

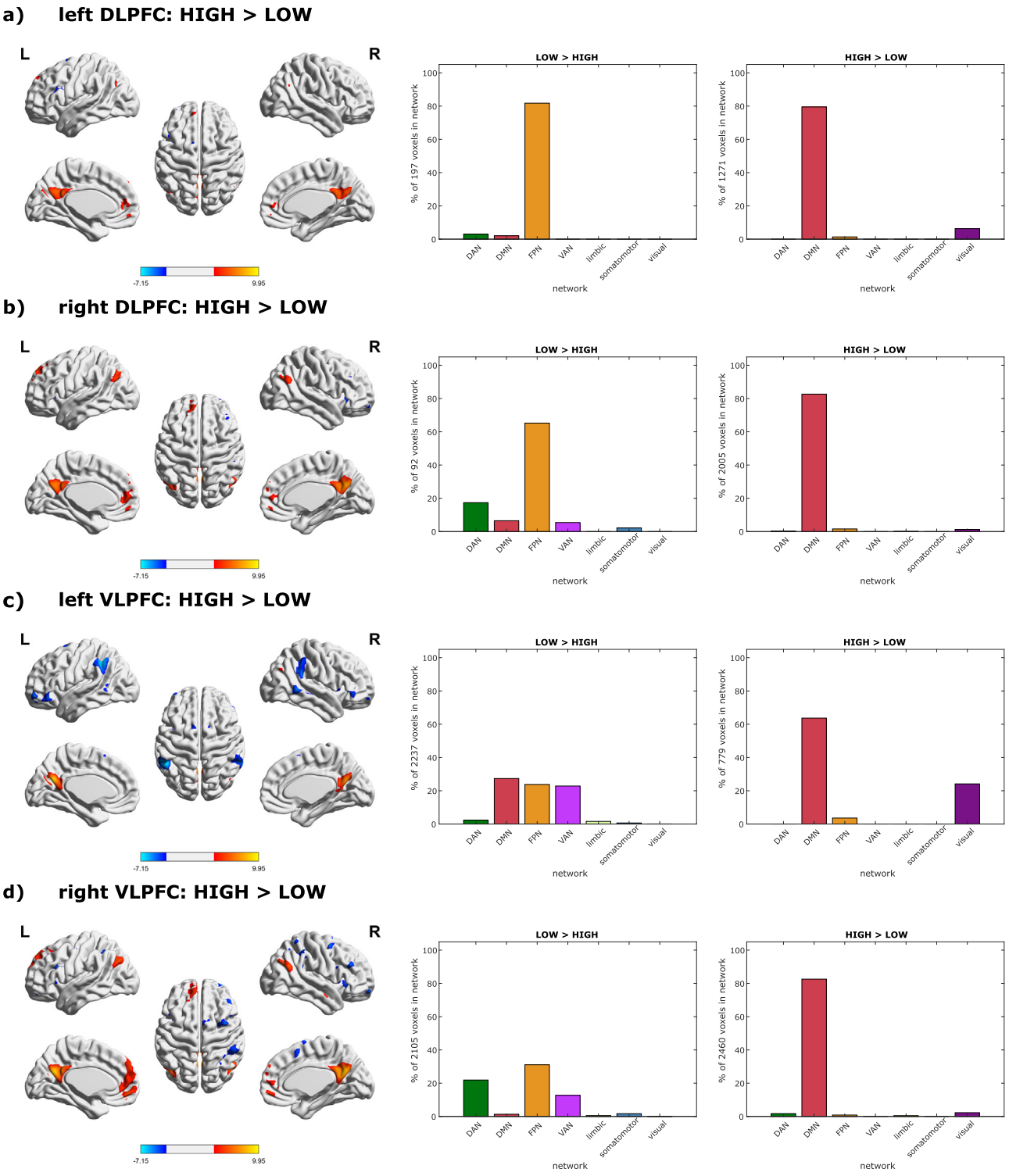


Fig. 5. Connectivity results based on a) left DLPFC; b) right DLPFC; c) left VLPFC and d) right VLPFC. T-value maps are thresholded at $p_{uncorrected} = 0.001$ for visualization purposes and were generated with BrainNet Viewer (Xia et al., 2013). Positive values depicted in warm colors represent areas in which coupling to the seed-region was greater in the high-load > low-load condition. Negative values depicted in cold colors represent areas in which coupling to the seed-regions was greater in the low-load > high-load condition. The bar graphs summarize the network assignment (Yeo et al., 2011) of all coupled voxels in the low-load > high-load contrast (middle column) and the high-load > low-load contrast (right column). DAN = dorsal attention network; DMN = default mode network; FPN = fronto-parietal network; VAN = ventral attention network.

Table 2
Connectivity results for the primary prefrontal seeds. Peak voxels for each cluster are reported in MNI space.

Connection	Seed	Contrast	pFWE *cluster*	voxels	pFWE *peak*	Tpeak	x / y / z	Regions
1	*left DLPFC*	LOW > HIGH	< 0.001	197	.312	4.49	−42 / 20 / 20	Left IFG
2		HIGH > LOW	< 0.001	617	< 0.001	6.55	−3 / −52 / 26	Bilat. Prec./PCC
3			< 0.001	565	.052	5.12	0 / 53 / −4	Bilat. MFC
4			.021	89	.300	4.50	−39 / −67 / 32	Left AG
5	*right DLPFC*	LOW > HIGH	.023	92	.864	3.86	57 / −43 / 35	Right SMG
6		HIGH > LOW	< 0.001	667	< 0.001	7.52	3 / −64 / 29	Bilat. Prec./PCC
7			.001	163	.003	5.94	51 / −64 / 29	right AG
8			< 0.001	927	.012	5.55	0 / 53 / −1	Bilat. MFC
9			< 0.001	248	.035	5.23	−39 / −70 / 38	Left AG
10	*left VLPFC*	LOW > HIGH	< 0.001	457	< 0.001	7.11	−57 / −46 / 41	Left IPL/SMG
11			< 0.001	283	.005	5.81	−42 / 56 / −10	Left ant. OFC
12			< 0.001	551	.007	5.72	60 / −43 / 29	Right SMG/MTG
13			< 0.001	278	.062	5.04	57 / 32 / −4	Right OFC
14			< 0.001	232	.072	4.99	−24 / −79 / −43	Left Cer.
15			< 0.001	281	.126	4.81	−12 / 5 / 68	Bilat. SMA
16			.002	155	.296	4.49	−48 / 5 / −10	Left Ins../temp. pole
17		HIGH > LOW	< 0.001	779	< 0.001	9.93	−6 / −64 / 23	Bilat. Prec./Cuneus
18	*right VLPFC*	LOW > HIGH	.026	88	.005	5.83	30 / 59 / −13	Right ant. OFC
19			< 0.001	289	.010	5.61	−27 / −49 / 44	Left IPS/SPL
20			< 0.001	418	.013	5.52	39 / 2 / 59	Right SMA/MFG
21			.013	105	.013	5.52	45 / 32 / 29	Right MFG/IFG
22			< 0.001	373	.035	5.23	−6 / −79 / −34	Left Cer.
23			< 0.001	487	.043	5.16	36 / −49 / 65	Right IPS/SPL
24			.007	119	.049	5.12	−42 / 20 / 26	Left IFG
25			.030	85	.254	4.55	30 / 26 / −4	Right INS
26			.003	141	.308	4.47	45 / 11 / 26	Right IFG
27		HIGH > LOW	< 0.001	835	< 0.001	9.15	0 / −61 / 26	Bilat. Prec./PCC
28			.001	190	< 0.001	6.85	51 / −64 / 26	Right AG
29			< 0.001	1188	.004	5.92	0 / 56 / −4	Bilat. MFC
30			< 0.001	247	.005	5.84	−42 / −73 / 35	Left AG/Occ. Cortex

Abbreviations: DLPFC/VLPFC = dorsolateral/ventrolateral prefrontal cortex; ant. = anterior; bilat. = bilateral; AG = angular gyrus; cer. = cerebellum; front. = frontal; IFG/MFG/SFG = inferior/middle/superior frontal gyrus; ins. = insula; IPL = inferior parietal lobule; IPS = intraparietal sulcus; ITG/MTG/STG = inferior/middle/superior temporal gyrus; MFC = medial frontal cortex; occ. = occipital; OFC = orbitofrontal cortex; par. = parietal; PCC = posterior cingulate gyrus; prec. = precuneus; SMA = supplementary motor area; SMG = supramarginal gyrus; temp. = temporal).

3.3.3. ROI-specific ANOVAs: time and interaction load X time

After we had identified the LPFC connections sensitive to different levels of instructional load in the previous analysis step, we were also interested to which degree these connections changed over the course of the implementation phase. For this purpose, we extracted subjects' mean BSC values per condition and connection identified via the two directions of the *instructional load* contrast at whole-brain level (i.e., all combinations of seed and target-ROI from 3.3.1) and entered them as dependent variables into multiple 2-by-2 repeated measures ANOVAs with factors *instructional load* and *time* (stimulus repetitions 1 and 2 as well as 3 and 4 were aggregated resulting in an *early* and a *late* implementation stage, respectively). A total of 30 ANOVAs were conducted. Correction for multiple testing was implemented by controlling the false-discovery rate individually at the level of each seed (e.g., correcting for 4 left DLPFC-related tests). Only significant effects surviving this correction procedure are reported here.

Besides the highly significant main effect of *instructional load* that obviously had to be present for each connection mirroring the original whole-brain results (all $F_{1,64} > 18.45$; all p(F) ⟨ 0.001; all η_p^2 ⟩ 0.22), a main effect of *time* was detected for several connections (all $F_{1,64} > 13.17$; all p(F) ⟨ 0.002; all η_p^2 ⟩ 0.17) in all cases reflecting a significant decrease in coupling strength towards the late implementation stage. Interestingly, all these connections (6 to 9 and 27 to 30) involved right hemispheric seed regions and were identified via the contrast direction *high-load* vs. *low-load* in the original whole-brain analysis. Additionally, 4 connections displayed a significant *instructional load x time* interaction effect (all $F_{1,64} > 6.79$; all p(F) ⟨ 0.011; all η_p^2 ⟩ 0.096). All of those connections (1, 18, 20 and 22) were identified via the *low-load > high-load* contrast direction in the original whole-brain analysis. The implications of these interaction effects were heterogeneous as, for example, at connection 1 the coupling in the two load conditions converged towards the end of the implementation phase whereas at connection 18 coupling increased more strongly over time in the high-load condition (see Table 3 for an overview and Fig. 6 for an illustration). When applying a less liberal significance criterion, i.e., using family-wise error correction (Holm, 1979) and taking into account all 30 ANOVAs at once, all reported main effects of time (connections 6 to 9 and 27 to 30) remained significant as did two of the reported interactions (1 and 18).

As all load-dependent connectivity results reported here are inherently relative, it would be informative to compare both low-load and high-load connectivity to a third 'baseline' condition. Thereby, we could determine, for instance, whether a load-dependent connectivity increase is due to increased high-load connectivity relative to baseline or decreased low-load connectivity relative to baseline. Unfortunately, the present experiment did not comprise such an unbiased baseline condition (e.g., fixation). In *supplementary analyses 2* we report the results obtained based on an auxiliary and arguably suboptimal baseline condition taken from a different experiment ('Experiment 2' from Ruge et al. (2019)).

3.4. Whole-brain connectivity: follow-up analyses

The load-dependent connectivity profiles of the primary LPFC-seeds suggested a shift from stronger coupling to areas assigned to attention/control-related networks during the low-load condition towards a stronger coupling to DMN-assigned regions during the high-load condition. It remained unclear, however, whether such a connectivity profile was specific for the pre-selected LPFC seeds (as they are considered control regions themselves) or whether this only exemplified a more general phenomenon beyond these seed regions. In par-

Table 3
ANOVA results. Connections identified in 3.3.1 were examined. P-values are printed in italic letters when trend-wise significance ($p < .1$) was indicated and in bold letters when significance ($p < .05$) was indicated. P-values that survived FDR-correction per seed (for procedure, see 3.3.3) are denoted by (*). P-values that survived family-wise error correction for all 30 seeds are denoted by (†), additionally.

SEED region	Target regions	Connection	time			load*time		
			F	*p*	η_p^2	*F*	*p*	η_p^2
DLPFC_L	Left IFG	1	1.59	.212	0.024	11.79	**.001***†	0.156
	Bilat. Prec./PCC	2	0.96	.330	0.015	0.42	.519	0.007
	Bilat. MFC	3	0.11	.747	0.002	0.66	.419	0.010
	Left AG	4	2.81	*.099*	0.042	0.53	.470	0.008
DLPFC_R	Right SMG	5	0.04	.836	0.001	2.01	.162	0.030
	Bilat. Prec./PCC	6	19.59	**<0.001***†	0.234	0.62	.433	0.010
	right AG	7	13.18	**.001***†	0.171	1.33	.254	0.020
	Bilat. MFC	8	24.54	**<0.001***†	0.277	0.07	.787	0.001
	Left AG	9	31.71	**<0.001***†	0.331	1.13	.291	0.017
VLPFC_L	Left IPL/SMG	10	0.59	.447	0.009	2.68	.106	0.040
	Left ant. OFC	11	0.01	.944	<0.001	3.88	*.053*	0.057
	Right SMG/MTG	12	1.25	.268	0.019	1.97	.165	0.030
	Right OFC	13	0.14	.714	0.002	0.29	.592	0.005
	Left Cer.	14	1.81	.183	0.028	1.12	.293	0.017
	Bilat. SMA	15	2.33	.132	0.035	7.73	**.007**	0.108
	Left Ins./temp. pole	16	2.51	.118	0.038	1.19	.280	0.018
	Bilat. Prec./Cuneus	17	2.21	.142	0.033	3.09	*.083*	0.046
VLPFC_R	Right ant. OFC	18	0.73	.396	0.011	15.52	**<0.001***†	0.195
	Left IPS/SPL	19	0.20	.660	0.003	5.97	**.017**	0.085
	Right SMA/MFG	20	1.26	.266	0.019	7.09	**.010***	0.100
	Right MFG/IFG	21	0.26	.609	0.004	4.29	**.042**	0.063
	Left Cer.	22	0.02	.903	<0.001	6.79	**.011***	0.096
	Right IPS/SPL	23	0.14	.713	0.002	4.57	**.036**	0.067
	Left IFG	24	2.99	*.089*	0.045	5.70	**.020**	0.082
	Right Ins.	25	0.01	.925	<0.001	1.82	.182	0.028
	Right IFG	26	0.58	.451	0.009	0.77	.385	0.012
	Bilat. Prec./PCC	27	29.62	**<0.001***†	0.316	3.60	*.062*	0.053
	Right AG	28	20.32	**<0.001***†	0.241	4.48	**.038**	0.065
	Bilat. MFC	29	28.74	**<0.001***†	0.310	0.23	.635	0.004
	Left AG/Occ. cortex	30	37.77	**<0.001***†	0.371	3.67	*.060*	0.054

Abbreviations: DLPFC/VLPFC = dorsolateral/ventrolateral prefrontal cortex; ant. = anterior; bilat. = bilateral; AG = angular gyrus; cer. = cerebellum; front. = frontal; IFG/MFG/SFG = inferior/middle/superior frontal gyrus; ins. = insula; IPL = inferior parietal lobule; IPS = intraparietal sulcus; ITG/MTG/STG = inferior/middle/superior temporal gyrus; MFC = medial frontal cortex; occ. = occipital; OFC = orbitofrontal cortex; par. = parietal; PCC = posterior cingulate gyrus; prec. = precuneus; SMA = supplementary motor area; SMG = supramarginal gyrus; temp. = temporal).

ticular, we were interested in whether DMN-based seeds would also exhibit elevated functional coupling to other DMN-regions during the *high-load* > *low-load* contrast. To further examine this issue we decided to conduct additional whole-brain connectivity analyses with DMN-based seeds. This time, the clusters identified in 3.3.1 via the *high-load* > *low-load* contrast – almost exclusively assigned to the DMN – were used as seed regions. Thereby, we could efficiently and simultaneously assess both, the specificity of *high-load* > *low-load* LPFC connectivity and more general load-dependent DMN-connectivity changes beyond the original LPFC seed regions. As there was a considerable overlap between these clusters, for simplification, we generated 'combined clusters' resulting in 4 new seed regions (see methods Section 2.3.4 for details and Fig. 7a for an illustration).

The connectivity patterns evoked by the high-load seeds (see Table 4 and Figs. 7b and 7c for the coupling pattern of ventral precuneus/PCC and left AG, respectively) were very similar across all four seeds: Increased coupling for the *low-load* vs. *high-load* contrast was observed exclusively within DMN-regions with an emphasis on precuneus/cuneus. For the *high-load* vs. *low-load* contrast, in turn, very large clusters were detected mostly covering areas assigned to control networks in addition to the visual network (VN). Thus, the DMN-based seeds and the LPFC-seeds appear to elicit inverse coupling profiles during the two instructional load conditions.

3.5. Correlation between connectivity changes and behavioral/trait-like variables

We used the mean difference in BSC magnitude between the low-load and the high-load condition across all voxels within a ROI for computing the correlation. Since these values were extracted from the single subjects' *low-load* vs. *high-load* contrast maps, positive difference values indicate greater connectivity during the low-load condition in a specific area whereas negative values indicate the opposite connectivity pattern. For an illustration, please see Fig. 8.

When using the Pearson correlation coefficient as measure of association, only connections based on the left VLPFC seed survived FDR-correction. Differential coupling of connection 14 (left VLPFC-left cerebellum) was positively associated with forward ($r = 0.41$; $p < .001$, two-tailed) as well as backward ($r = 0.41$; $p < .001$, two-tailed) digit span. Differential coupling of connections 12 (left VLPFC-right SMG) and 13 (left VLPFC-right OFG/IFG) was associated positively with first-trial accuracy in the low-load condition (both $r > 0.38$; both $p < .002$, two-tailed).[1]

4. Discussion

This study set out to examine load-dependent connectivity changes emerging during the implementation of instructed novel S-R associations. In particular, we focused on the functional coupling of four lateral prefrontal seeds (bilateral DLPFC and VLPFC) that had previously been

[1] Association between conn. 14 (left VLPFC-left Cerebellum) and backward memory span (r = .38; p = .002, two-tailed) as well as between conn. 13 (left VLPFC-right OFG/IFG) and first-trial accuracy in the low-load condition (r = .44; p < .001, two-tailed) was still significant when using Spearman correlation as an alternate measure of association. No additional connectivity-behavior correlation survived FDR-correction.

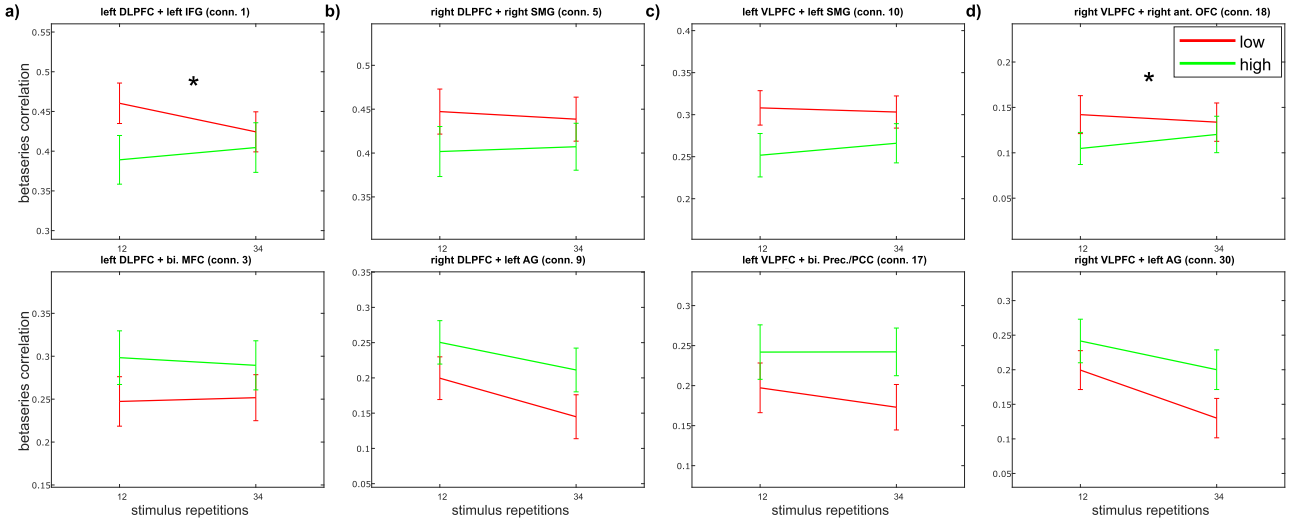


Fig. 6. Load-dependent connectivity as a function of time (early: stimulus repetitions 1 and 2; late stimulus repetitions 3 and 4). Results are illustrated for exemplary connections involving each of the four primary seed regions a) left DLPFC, b) right DLPFC, c) left VLPFC and d) right VLPFC. Connections identified via the low-load > high-load contrast are shown in the top row whereas connections identified via the high-load > low-load contrast are shown in the bottom row. Significant interactions of instructional load by time are denoted by (*). Error bars represent 95% confidence intervals.

Table 4
Follow-up connectivity analysis results based on regions with greater high-load > low-low coupling to LPFC seeds in 3.3.1. Peak voxels for each cluster are reported in MNI space.

Seed	Contrast	pFWE cluster	voxels	pFWE peak	Tpeak	x / y / z	Regions
Prec./PCC	LOW > HIGH	< 0.001	484	< 0.001	7.63	12 / −58 / 20	Bilat. Prec.
		< 0.001	247	.095	4.92	−3 / 38 / −1	Bilat. MFC
	HIGH > LOW	< 0.001	24,507	< 0.001	8.77	36 / 44 / 29	Lat. front./par./temp. Cortex, occ. cortex
				< 0.001	8.62	54 / 17 / −7	Right temp. pole/IFG
				< 0.001	8.61	−54 / −43 / 47	Left SMG
MFC	LOW > HIGH	< 0.001	294	< 0.001	6.62	9 / −58 / 23	Bilat. Prec./Calcarine
	HIGH > LOW	< 0.001	17,248	< 0.001	11.63	60 / −43 / 32	Lat. front./par./temp. Cortex, occ. Cortex
				< 0.001	10.08	−60 / −49 / 26	Left SMG
				< 0.001	9.59	57 / −40 / 44	Right SMG
Right AG	LOW > HIGH	< 0.001	314	.003	5.96	−9 / −55 / 14	Bilat. Prec./Calcarine
	HIGH > LOW	< 0.001	16,814	< 0.001	8.15	54 / 11 / 8	Lat. front./par./temp. Cortex, occ. Cortex
				< 0.001	7.72	57 / −40 / 38	Right SMG
				< 0.001	7.59	−39 / 5 / −1	Left ant. Ins.
Left AG	LOW > HIGH	< 0.001	262	.003	5.98	−6 / −58 / 17	Bilat. Prec.
		.047	75	.007	5.70	−39 / −76 / 41	Left MOG/AG
	HIGH > LOW	< 0.001	22,324	< 0.001	9.36	57 / −40 / 35	Lat. front./par./temp. Cortex, occ. cortex
				< 0.001	9.15	−60 / −49 / 29	Left SMG
				< 0.001	8.35	54 / −52 / 2	Right MTG
		.033	83	.348	4.42	15 / −7 / 14	Right Thal.

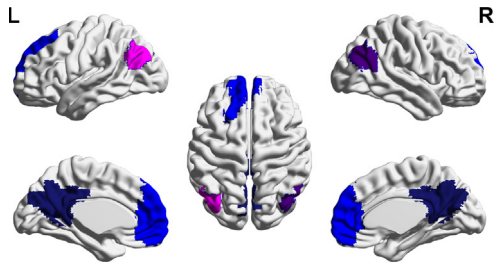
Abbreviations: ant. = anterior; bilat. = bilateral; AG = angular gyrus; front. = frontal; IFG/MFG/SFG = inferior/middle/superior frontal gyrus; ins. = insula; ITG/MTG/STG = inferior/middle/superior temporal gyrus; lat. = lateral; MFC = medial frontal cortex; MOG = middle occipital gyrus; occ. = occipital; par. = parietal; PCC = posterior cingulate gyrus; prec. = precuneus; SMG = supramarginal gyrus; temp. = temporal; thal. = thalamus).

demonstrated to bear differential relevance for successful IBL. Specifically, earlier MVPA results from Ruge et al. (2019) suggested a relevant contribution of neural rule representations in the bilateral VLPFC (but not DLPFC) to successful rapid IBL. Notably, these VLPFC representations did not seem to be compromised with increasing instructional load despite marked effects on behavior. By contrast, the present functional connectivity analysis revealed that instructional load did in fact strongly affect the functional couplings of all of these four lateral prefrontal regions. In particular, it seems that lateral PFC regions are integrated into entirely different brain systems depending on whether or not instructional load exceeded standard short term memory capacity limits.

On the one hand, across all four prefrontal seeds, we found strengthened couplings with regions within the lateral prefrontal and parietal cortices, the insulae, and the cerebellum during the low-load as compared to the high-load condition. Using a well-established large-scale network parcellation (Yeo et al., 2011) these individual regions could be assigned to large-scale networks such as the FPN, DAN and VAN. On the other hand, regions exhibiting tighter functional coupling with all four prefrontal seeds during the high-load as compared to the low-load conditions were almost exclusively assigned to the DMN including regions such as the bilateral AG, mPFC, and most prominently the junction of ventral precuneus and PCC. However, although this general pattern was observed for all four seeds, the load-dependent coupling profile of the left VLPFC stood out considerably from that of the other three prefrontal seed regions. Together with the finding that significant correlations with behavioral measures were selectively found for functional couplings involving the left VLPFC seed, this suggests a specialized role of the left VLPFC during IBL.

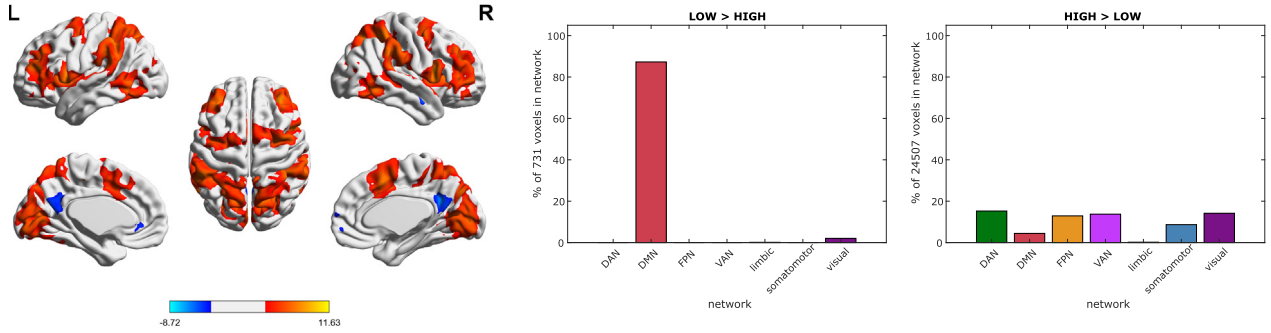
Our complementary analysis of local brain activity (instead of connectivity) changes associated with instructional load showed results that

a) follow-up seeds: DMN



cons. cluster	voxels	x / y / z	network
bilat. Prec./PCC	1154	1 / -57 / 25	DMN
bilat. MFC	1355	-5 / 49 / 19	DMN
right AG	215	48 / -65 / 29	DMN
left AG	296	-43 /-68 / 33	DMN

b) bi. Prec./PCC: HIGH > LOW



c) left AG: HIGH > LOW

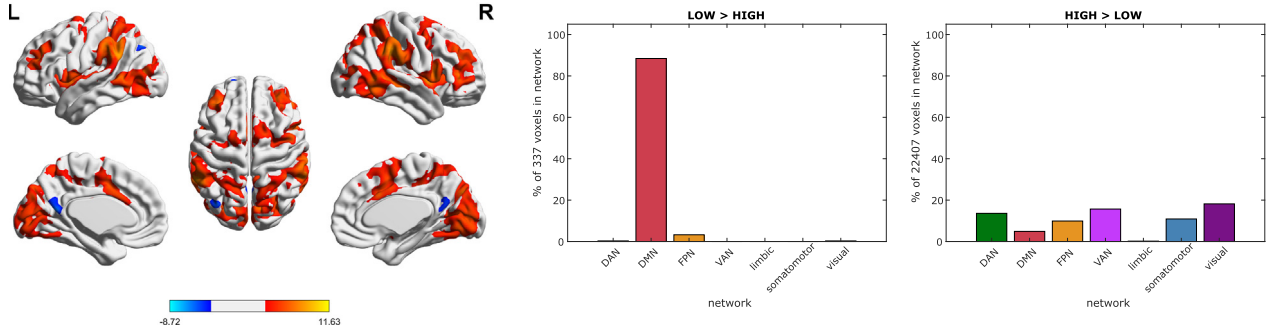


Fig. 7. a) DMN-based seeds used for follow-up analyses. Seed regions are rendered on cortex in the left column. Information about the seed regions is provided in the table in the right column. Colors in which seeds are depicted in the cortex figure correspond to the respective fond color of seeds' names in the table. Xyz-coordinates represent center of mass of the seed regions. Connectivity results based on b) bilateral ventral precuneus/posterior cingulate cortex and c) left angular gyrus: thresholded t-value maps (at $p_{uncorrected} = 0.001$) are provided in the left column – positive values represent areas in which coupling to the seed-region was greater in the high-load > low-load condition and are depicted in warm colors, negative values represent areas in which coupling to the seed-regions was greater in the low-load > high-load condition and are illustrated in cold colors; network assignment of all coupled voxels (Yeo et al., 2011) in the low-load > high-load (middle column) and the high-load > low-load (right column) contrast (DAN = dorsal attention network; DMN = default mode network; FPN = fronto-parietal network; VAN = ventral attention network). Cortex images were generated with BrainNet Viewer (Xia et al., 2013).

are generally in line with the common observation of fronto-parietal activation increases at higher levels of cognitive load (Duncan, 2010; Emch et al., 2019; Gordon et al., 2012; Woolgar et al., 2015). Specifically, these regions were more strongly activated in the higher instruction load condition and they were also more strongly activated at the beginning of the implementation phase compared to later implementation trials supposedly reflecting higher levels of cognitive control requirements early in novel task practice, which is in line with previously reported results (Cole et al., 2013b; Duncan, 2010; Mohr et al., 2016).

Not least before this background, it might seem somewhat counterintuitive that the cognitively more demanding condition (higher instruction load) was associated with a tighter coupling between our LPFC seeds and various DMN regions, which have 'historically' been conceptualized as 'task-negative' regions exhibiting lower local activity in the more demanding condition (Anticevic et al., 2012; Fox et al., 2005). However, while this 'task-negative' DMN-conceptualization is mainly based on local activation results there is a growing body of evidence from connectivity studies (Elton and Gao, 2015; Finc et al., 2017; Spreng et al., 2010; Wang et al., 2021) that the DMN actually has an active role in goal-directed behavior. Accordingly, studies have reported stronger connectivity between the DMN and other ('task-positive') large-scale networks with increasing cognitive demand and stronger within-DMN connectivity with decreasing demand (Finc et al., 2017; Vatansever et al., 2017a). Such type of connectivity changes have been suggested to reflect different levels of automaticity (Vatansever et al., 2017b). A similar automaticity-based account has also been suggested in

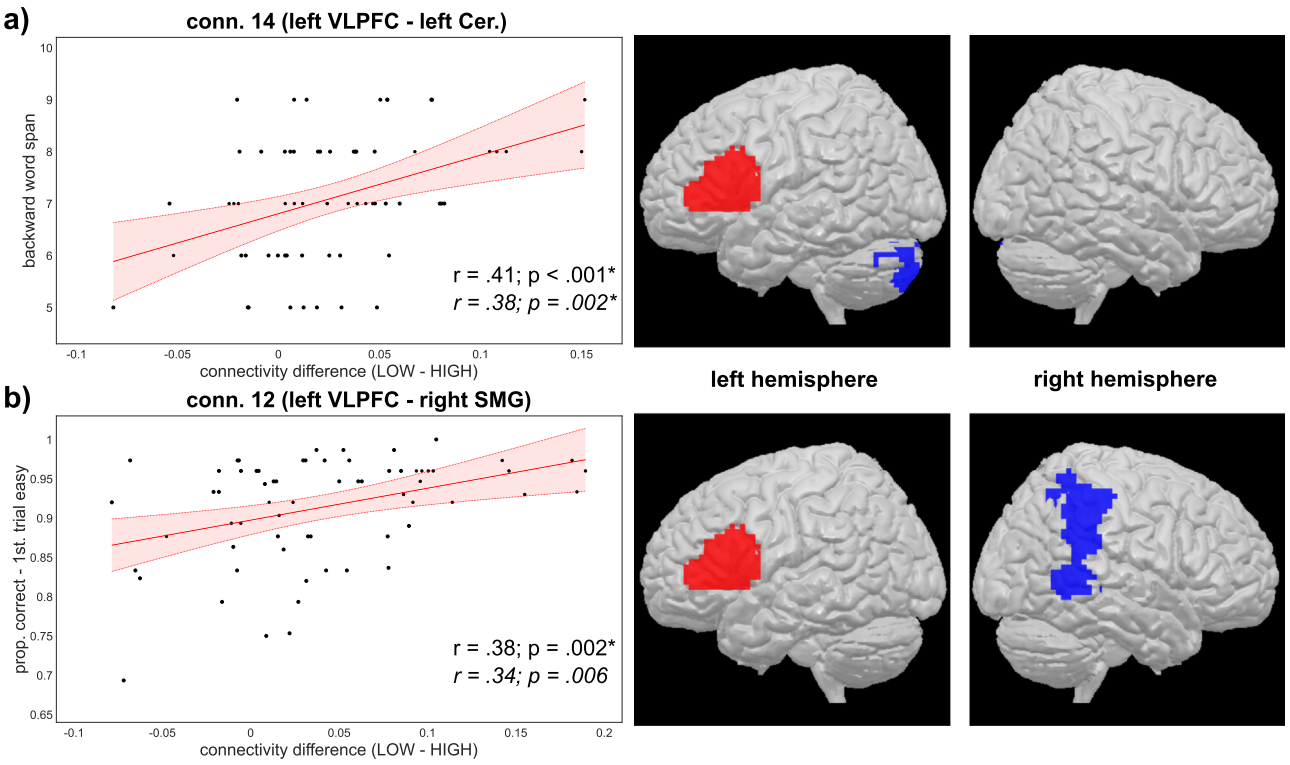


Fig. 8. a) correlation between differential coupling (low-load – high-load) of connection 14 and backward digit span; b) correlation between differential coupling (low-load – high-load) of connection 12 and 1st trial accuracy in the low-load condition; scatter plots illustrating correlations are shown in the left column, seed (in red) and 'target' (in blue) of the connections are depicted in the middle and right column. Coefficients and p-values are based on the Pearson (standard font) and the Spearman (italic) correlation, respectively. Significant values after correction are denoted by (*).

the context of practice-related connectivity changes in IBL (Mohr et al., 2016, 2018) and will be discussed in relation to our load-dependent connectivity results in greater detail further below. Moreover, we will argue that increased LPFC-DMN coupling at higher levels of instructional load might be linked to (episodic) long-term memory contributions to IBL (Meiran et al., 2017; Monsell and Graham, 2021), in line with previous studies suggesting a role of DMN regions for episodic memory retrieval (Ritchey and Cooper, 2020; Rugg and Vilberg, 2013; Vatansever et al., 2021).

4.1. Automaticity: initial instruction-induced differences

The load manipulation – applied to the instruction phase and impacting the implementation phase only indirectly – initially resulted in a considerably lower mean accuracy in the high-load as compared to the low-load condition from stimulus repetition 1 onwards.

An initial processing difference between the load conditions seems to depend on multiple factors each of which might contribute to more fluent or 'automatic' instruction implementation from the outset. First, the low-load condition benefits quite obviously from the smaller number of instructed S-R rules that arguably lies within the capacity limits that have been suggested by theoretical frameworks on 'instruction-based proceduralisation' (Brass et al., 2017). As such, low-load instructions should be easily accessible and well-prepared for their seamless implementation. Second, subjects had twice as much potential encoding time per individual S-R rule in the low-load as compared to the high-load condition. As longer preparation time during rule encoding has been shown to boost implementation of newly instructed rules (Cole et al., 2018) it is likely that our participants were able to form a highly accessibly rule representation during the instruction phase of the low-load condition possibly through motor imagery and more extensive 'covert practice' (Liefooghe et al., 2021; Ruge and Wolfenstetter, 2010; Theeuwes et al., 2018). Third, the expectation to implement an instruction might have played a role as well. While all of the instructed S-R rules (i.e. 100%) in the low-load condition were subsequently required to be implemented, this was only the case for 40% of the individual S-R rules instructed in the high-load condition. The higher expectation to actually implement a rule might have contributed to a more robust rule representation in the low-load condition possibly comparable with the distinctive effects of intentionally preparing implementation versus mere memorization of S-R rules (Formica et al., 2020; Liefooghe et al., 2012). Finally, load-dependent differences in automaticity could be related to higher levels of ambiguity during response selection in the high-load condition. Specifically, in the instruction phase, 3 to 5 stimuli are assigned to each response in the high-load condition compared to only 1 to 2 stimuli per response in the low load condition (see *supplementary analyses 1* for evidence supporting the load-dependent relevance of response ambiguity).

Benefiting from all of the above mentioned factors – probably inducing strong 'intention-based reflexivity' (Meiran et al., 2012; Meiran and Cohen-Kdoshay, 2012) – S-R rules in the low-load condition seem to be highly ready for implementation from the outset. In this case, reflexivity acquired in the instruction phase aligns with the currently relevant rule during the implementation phase (which is almost always the originally instructed rule). This contrasts with the high-load condition where the originally instructed S-R rule does not align with the currently relevant rule in a considerable amount of trials which, in turn impedes potential practice-related automatization processes in the high-load condition (see next section for elaboration).

From a functional network perspective, in the low-load condition, the instantaneously fluent application of the instructed S-R rules is accompanied by higher *within*-network connectivity regarding both the

FPN (our LPFC seeds were for the most part located within the FPN) and the DMN (as revealed by the DMN-seeded follow-up analyses). This was particularly true for the DLPFC seeds. The VLPFC-based connections were somewhat more distributed across the brain, yet still mainly associated with attention/control networks. By contrast, in the high-load condition, more effortful processing and continued rule updating demand was accompanied by increased *between*-network connectivity. Here, LPFC-based connectivity was characterized by greater coupling to DMN regions whereas the DMN-seeded follow-up analyses revealed widely distributed increased coupling with regions located within the DAN, FPN, VAN and the visual network. This can be taken as evidence for reduced brain-wide modularity driven by greater DMN integration during rule implementation in the high-load condition. Such a reduced modularity with emphasis on DMN integration has been observed before at higher load conditions (Finc et al., 2017; Vatansever et al., 2015) and is thought to provide a basis for global information integration (Vatansever et al., 2015) which, in turn, might also underlie the effects induced by the present high-load instruction condition. In line with this, another study reported a similar pattern of DMN integration and segregation during the initial acquisition and subsequent application of new rules, respectively (Vatansever et al., 2017b). This distinction resembles the aforementioned automaticity-related processing differences between high and low instructional load in the present study: Relatively automatic rule implementation common to both the present low-load condition and the learnt rule application phase in Vatansever et al. (2017b) contrasts with ongoing rule updating (see below) in the present high-load condition and the acquisition phase of Vatansever et al. (2017b). Furthermore, this rule updating might potentially bear resemblance with continuous learning or inference processes supported by the DMN (cf. Dohmatob et al., 2020).

4.2. Automaticity: the role of practice

The concept of automaticity usually involves repeated execution of the same rule even across short time spans accompanied by increasingly fluent application of the newly instructed rules towards the late implementation trials (Chein and Schneider, 2005; Mohr et al., 2016).

While mean accuracy remained relatively stable in the low-load condition (around 90%), accuracy rates in the high-load condition increased considerably over the course of the remaining implementation phase from about 70% at repetition 1 to 87% at repetition 4 (still significantly lower than in the low-load condition). This is remarkable as the two load conditions were physically identical regarding the implementation phase, that is, both involving the same number of 4 to-be-implemented S-R links. Keeping in mind that from repetition 2 onwards accuracy was defined relative to the previously executed response for a given stimulus, these results imply different levels of rule processing in the two load conditions. One the one hand, a constantly high accuracy rate in the low-load condition implies minimal change of the behaviorally relevant rule representation (i.e., the same correct rule is applied again and again). On the other hand, in the high-load condition, increasing accuracy across a learning block necessarily has to go along with at least partial updating of rule representations (i.e., adapting the previously executed 'wrong' response to the same stimulus as the relevant rule). While the former would be considered to be an example of relatively automatic processing, the latter likely requires substantial cognitive control effort. Furthermore, enduring response conflict especially in the high-load condition could additionally prevent rapidly evolving practice-driven automaticity across stimulus repetitions during the implementation phase (Mohr et al., 2016). Such response conflicts (i.e. originally instructed S-R rule and newly adapted S-R rule do not match) occur much more frequently in the high-load condition. Our finding that accuracy in high-load conflict trials was significantly lower than in non-conflict trials implies that competing S-R rule representations in the high-load condition contribute substantially to the persisting accuracy differences between the instructional load conditions. This competition needs to be overcome (e.g., by dedicating cognitive effort to the suppression of the initially instructed but now irrelevant rule) before a rule can be processed more automatically in the high-load.

Mohr et al. (2016) reported decreasing fronto-parietal activation as well as progressing segregation of the DMN from a number of attention/control networks during the transition from early to late implementation trials. This neural pattern associated with 'short-term automatization' resembles the one we observed between the different load conditions in the present study. Moreover short-term automatization across implementation trials seems to be influenced by task demands as well (Mohr et al., 2018): DMN segregation from task-related networks like the DAN was less pronounced during an arguably more demanding 'reversal learning' condition than during an initial learning condition. If both time-related (i.e. practice-related) as well as load-related automaticity differences induced similar neural (de-)coupling processes, we would have expected the connections identified via the instructional load contrast to show a time effect as well. This was, however, only partly the case: First, only the DMN-related connections based on the right hemispheric LPFC seeds showed a general connectivity decrease across repetitions. Furthermore, we would have expected the two factors to interact (i.e. more pronounced DMN decoupling towards the end of the implementation phase in the low-load compared to the high-load condition). Evidence for such an interaction effect was weak at best as only some of the DMN-related connections showed as much as a trend towards a significant interaction effect. Regarding connections originally identified via the *low-load* > *high-load* contrast, again a heterogeneous picture emerged: While all interaction effects generally indicated a less pronounced coupling difference between load conditions towards the end of the implementation phase, some of the connections (e.g., connection 1: left DLPFC – left IFG) showed a pronounced coupling decrease in the low-load condition whereas others (e.g. connection 18: right VLPFC – right anterior OFC) showed a pronounced coupling increase in the high-load condition. This rather heterogeneous results pattern might reflect the specific contributions of a variety of the previously discussed sub-processes: First, connections that do show an interaction effect, especially driven by a coupling increase in the high load condition (e.g., Fig. 6d, upper panel), might reflect reduced suffering from interference towards the end of the implementation phase - consistent with the decreasing conflict effect over time. Second, connections that do not show such an interaction effect (i.e., the magnitude of the load-difference is not affected by practice), in turn, might rather reflect a constantly ongoing S-R rule updating process that is mirrored by the constant slope of the overall linear accuracy increase in the high-load condition (e.g., Fig. 6b and 6c, upper panel). Third, interaction effects specifically driven by a time-dependent coupling decrease in the low-load condition (e.g., Fig. 6a, upper panel) might indicate a more rapid trend towards stronger modularity (see above) due to more rapid automatization as S-R rules are already fully established at the outset.

A final interesting observation was that significant practice-related effects (main effect of time as well as interaction effects) were detected predominantly at connections involving the right VLPFC seed. This was especially true when compared to the left VLPFC seed where load-induced connectivity differences seemed to remain rather static and implies that different rule features are conveyed by the VLPFC-region of both hemispheres. On the one hand, the rule representation in the left VLPFC might comprise features that are linked more closely to the original instruction which also seems consistent with the finding that exclusively the left hemispheric VLPFC connections were correlated with 1st trial performance (see below). On the other hand, the S-R rule representation in the right VLPFC might be more flexible: after initially being formed by the instruction it is reshaped during the implementation phase according to self-generated rules.

To conclude, the observed load and time-dependent changes in functional couplings seem to be consistent with changing levels of automaticity during instruction-based learning. Yet, we also found some

evidence that apparent automaticity-related changes in functional couplings come in different flavours which we argued to be reflecting a variety of underlying processing differences. While different degrees of 'intention-based reflexivity' (Cole et al., 2018; Meiran et al., 2012) might underlie initial differences between the load conditions, persistently elevated levels of conflict and response ambiguity might impair 'short-term automatization' in the high-load condition (Mohr et al., 2016, 2018). Thus, both aspects of 'automatic' behavior in IBL might cooperatively engage to a greater (low-load) or a lesser degree (high-load), but a more comprehensive characterization of their potentially multi-facetted relationship needs to be explored in future studies. Finally, at the level of individual neural connections, the presence (right VLPFC) and virtual absence (left VLPFC) of practice-related changes in functional couplings implies hemispherical differences in the type and impact of the respective S-R rule representations.

4.3. Episodic memory contributions

Another, possibly related, way to interpret our connectivity results rests on the notion of load-specific mnemonic mechanisms. On the one hand, the pattern of greater LPFC coupling to control and attention related regions, mostly within FPN and DAN, during the low-load condition might imply that successful task implementation is guided mainly by working memory (Curtis and D'Esposito, 2003; D'Esposito et al., 2000; Emch et al., 2019). On the other hand, the shift to greater coupling between LPFC and the DMN – especially its posterior components – in the high-load condition might indicate an increasing involvement of episodic (long-term) memory, taking into account the extensive literature on the neural basis of episodic memory processes (Cabeza et al., 2008; Kim, 2010; Ritchey and Cooper, 2020; Rugg and Vilberg, 2013; Sestieri et al., 2011; Westphal et al., 2017).

Working memory capacity is supposed to be severely limited (Baddeley and Hitch, 1974; Logie, 2011) with a capacity limit of around 5 or 6 words regarding the verbal component (Baddeley, 2012; Monsell and Graham, 2021). Clearly, the decline in performance in the high-load condition to around 70% accuracy indicates that this capacity gap could not be fully compensated by episodic memory mechanisms. There is even a possibility that episodic LTM does not play a role at all and that the above-chance accuracy in the high-load condition is accomplished by partial WM maintenance of the whole set of instructed S-R links: Assuming that some of the instructed 10 S-R links (e.g. 5) could be maintained within WM, above-chance performance accuracy could be due to the subset of memorized S-R links that were randomly selected to be actually implemented (e.g. 2 out of 4). In this example, roughly 70% accuracy would be expected (2 correctly memorized links out of a total of 4 to-be implemented S-R links plus, by chance, 1 out of the remaining 2 not-remembered links). This scenario, however, seems hard to reconcile with our finding of increased coupling between LPFC and DMN regions. Instead, one could argue that subjects' WM system was overwhelmed by the sheer number and rapid presentation rate of the 10 S-R links during the instruction phase. This might have caused WM to break down even below its normal capacity limits. In turn, ~70% performance accuracy might have relied more than usual on memory traces encoded within episodic LTM. Support for this latter claim comes from a recent behavioral study (Bartsch and Oberauer, 2022) in which it was shown that at larger set sizes – at a presentation rate identical to the one used for our high-load condition – memory performance was mostly determined by episodic LTM whereas only very small set sizes were actually depending on working memory. In general, rapidly formed episodic traces represented in the 'activated part of LTM' seem to affect performance in tasks that are actually designed to test WM (cf. Cowan, 2019). In the current study, also the conflict trial analysis hints towards an episodic memory contribution suggesting a detrimental impact of the still lingering memory trace of the actually instructed response. This memory trace is formed during the instruction context and interferes with the memory trace formed during the implementation context. For instance, if the response actually executed at stimulus repetition 1 for a certain stimulus deviates from the originally instructed response, the newly formed memory trace will be in competition with the originally encoded S-R link when the same stimulus occurs the next time. It has been argued that such kind of interference typically occurs within episodic LTM but not within WM where newly encoded content can be more effectively shielded against interference (cf. Bartsch and Oberauer, 2022; Cowan et al., 2005).

Coming back to our elaboration on automaticity further above, we would therefore argue that reduced automaticity due to increased interference in the high-load condition is specifically due to lingering S-R episodes within LTM. Consistent with this, it has been shown that unintended effects of instruction (here: interference due to no-longer used instructions) are unaffected by a manipulation tackling working memory capacity whereas the actual rapid instructed learning within WM limits is (Pereg and Meiran, 2019).

While the discussed memory-related effects in the present study are clearly either directly or indirectly induced by instructions, it remains to be shown whether they are specific of IBL or apply more generally to other types of learning (e.g., trial-and-error learning).

4.4. Special role of left VLPFC couplings?

The original study (Ruge et al., 2019) suggested that specifically the VLPFC but not the DLPFC comprised representations of the instructed S-R rules. Hence, we were interested in whether a similar regional specialization would also be present with regard to the load-dependent connectivity profiles. Interestingly, we found that the left VLPFC showed a connectivity profile that was distinctly different not only compared to both the left and the right DLPFC but also compared to the right-hemispheric VLPFC pendant whereas the connectivity profiles of the other three prefrontal seeds were relatively homogeneous. What distinguished the left VLPFC most from the other prefrontal seed regions was its tendency to exhibit significantly greater coupling to the bilateral SMG (extending into the MTG) during the low-load condition compared to the high-load condition. At the same time, coupling between left VLPFC and the DMN seemed to be less affected by the load manipulation than this was the case for the other three LPFC seed regions. Following our line of interpretation, this would indicate a DMN-based integration of the rule representation conveyed by the left VLPFC that is similar in the low-load and the high-load condition. However, the coupling between left VLPFC and the precuneus/PCC was a notable exception which was at least as load-sensitive (i.e., increasingly connected in the high-load condition) as the other prefrontal seeds. This is in line with the notion that the precuneus/PCC might constitute a distinct functional instance within the DMN during task performance (Utevsky et al., 2014). Since the right VLPFC showed load-dependent changes in DMN-coupling that were not limited to the precuneus/PCC our results suggest that there is a hemispheric difference in terms of the rule features the LPFC represents and how these are being integrated during instruction-based learning.

Given the verbal nature of our stimulus material, the pronounced coupling to the SMG, which has frequently been associated with phonological processing (Deschamps et al., 2014; Sliwinska et al., 2012), suggests a greater reliance on verbal rehearsal strategies for maintaining rule representation conveyed by the left VLPFC especially in the low-load condition. This connection's significant relationship to successful IBL performance exclusively in the low-load condition – participants with greater load-dependent coupling differences between left VLPFC and right SMG (connection 10) were more accurate at the first implementation trial – can be taken as evidence for that claim. Another connection, the one between the left VLPFC and the left cerebellum (connection 14), again emphasizes the many different aspects of instruction-based learning. We observed a significant relationship of load-dependent VLPFC-cerebellar coupling with performance in word span tests. While this can be taken as evidence that at least some of our

participants employed a verbal rehearsal strategy there was no direct indication of this connection being associated to better performance (neither in the first trial nor at later stages). This finding is reminiscent of earlier ambiguities concerning the relationship of classical verbal working memory tests and IBL (Ruge et al., 2018).

5. Conclusion

Using a seed-based connectivity approach we found that lateral prefrontal cortex couplings during rule implementation were strongly affected by instructional load. The low-load condition was characterized by increased cooperation between lateral prefrontal seeds and areas known to be relevant for goal-directed action whereas coupling shifted towards increased DMN coupling in the high-load condition. From a large-scale network perspective these observations likely reflect different degrees of automated processing. Concerning the VLPFC – especially when also taking into account practice-related effects and correlations with behavioral measures – our results suggest that left hemispheric VLPFC-based functional couplings are related to the enduring influence of and guidance by the initially instructed task rules whereas the right hemispheric VLPFC seems to have a more flexible role in rule updating during implementation. Our study informs the literature on (rapid) instruction-based learning by focusing on functional connectivity that might help to integrate existing ambiguities regarding rule representations in the lateral prefrontal cortex. Specifically, our results imply an interplay of episodic memory and (procedural) working memory traces with a shifting balance when working memory capacity is exceeded.

Data and code availability statement

Behavioral data and untresholded group-level contrast maps underlying the results reported in Sections 3.2 to 3.4 are publicly available at https://osf.io/uxvyd/.

Declaration of Competing Interest

We declare no conflicts of interest.

Credit authorship contribution statement

Alexander W. Baumann: Writing – original draft, Formal analysis, Visualization, Conceptualization, Methodology, Software. **Theo A.J. Schäfer:** Writing – review & editing, Conceptualization, Methodology. **Hannes Ruge:** Writing – review & editing, Conceptualization, Methodology, Software, Resources, Supervision, Project administration, Funding acquisition.

Data availability

Data will be made available on request.

Acknowledgements and funding

This work was supported by the German Research Foundation (DFG, SFB 940 project A2).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120262.

References

Abdulrahman, H., Henson, R.N., 2016. Effect of trial-to-trial variability on optimal event-related fMRI design: implications for Beta-series correlation and multi-voxel pattern analysis. Neuroimage 125, 756–766. doi:10.1016/j.neuroimage.2015.11.009.

Abrahamse, E., Braem, S., Houwer, J.de, Liefooghe, B, 2022. Tenacious instructions: how to dismantle newly instructed task rules? J. Exp. Psychol. 151 (11), 2812–2832. doi:10.1037/xge0001233.

Andrews-Hanna, J.R., Smallwood, J., Spreng, R.N., 2014. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann. N. Y. Acad. Sci. 1316, 29–52. doi:10.1111/nyas.12360.

Anticevic, A., Cole, M.W., Murray, J.D., Corlett, P.R., Wang, X.-J., Krystal, J.H, 2012. The role of default network deactivation in cognition and disease. Trends Cogn. Sci. (Regul. Ed.) 16 (12), 584–592. doi:10.1016/j.tics.2012.10.008.

Baddeley, A., 2012. Working memory: theories, models, and controversies. Annu. Rev. Psychol. 63, 1–29. doi:10.1146/annurev-psych-120710-100422.

Baddeley, A., Hitch, G., 1974. Working memory. In: Psychology of Learning and Motivation, 8. Elsevier, pp. 47–89.

Bartsch, L.M., Oberauer, K., 2022. The contribution of episodic long-term memory to working memory for bindings. Cognition 231, 105330. doi:10.1016/j.cognition.2022.105330.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Stat. Methodol. 57 (1), 289–300.

Bhandari, A., Duncan, J., 2014. Goal neglect and knowledge chunking in the construction of novel behaviour. Cognition 130 (1), 11–30. doi:10.1016/j.cognition.2013.08.013.

Brass, M., Liefooghe, B., Braem, S., Houwer, J.de, 2017. Following new task instructions: evidence for a dissociation between knowing and doing. Neurosci. Biobehav. Rev. 81 (Pt A), 16–28. doi:10.1016/j.neubiorev.2017.02.012.

Buckner, R.L., DiNicola, L.M., 2019. The brain's default network: updated anatomy, physiology and evolving insights. Nat. Rev. Neurosci. 20 (10), 593–608. doi:10.1038/s41583-019-0212-7.

Cabeza, R., Ciaramelli, E., Olson, I.R., Moscovitch, M., 2008. The parietal cortex and episodic memory: an attentional account. Nat. Rev. Neurosci. 9 (8), 613–625. doi:10.1038/nrn2459.

Chein, J.M., Schneider, W., 2005. Neuroimaging studies of practice-related change: fmri and meta-analytic evidence of a domain-general control network for learning. Brain Res. Cogn. Brain Res. 25 (3), 607–623. doi:10.1016/j.cogbrainres.2005.08.013.

Cocuzza, C.V., Ito, T., Schultz, D., Bassett, D.S., Cole, M.W., 2020. Flexible coordinator and switcher hubs for adaptive task control. J. Neurosci. 40 (36), 6949–6968. doi:10.1523/JNEUROSCI.2559-19.2020.

Cole, M.W., Braver, T.S., Meiran, N., 2017. The task novelty paradox: flexible control of inflexible neural pathways during rapid instructed task learning. Neurosci. Biobehav. Rev. 81 (Pt A), 4–15. doi:10.1016/j.neubiorev.2017.02.009.

Cole, M.W., Ito, T., Braver, T.S., 2016. The behavioral relevance of task information in human prefrontal cortex. Cereb. Cortex 26 (6), 2497–2505. doi:10.1093/cercor/bhv072.

Cole, M.W., Laurent, P., Stocco, A., 2013a. Rapid instructed task learning: a new window into the human brain's unique capacity for flexible cognitive control. Cogn. Affect. Behav. Neurosci. 13 (1), 1–22. doi:10.3758/s13415-012-0125-7.

Cole, M.W., Patrick, L.M., Meiran, N., Braver, T.S., 2018. A role for proactive control in rapid instructed task learning. Acta Psychol. (Amst) 184, 20–30. doi:10.1016/j.actpsy.2017.06.004.

Cole, M.W., Reynolds, J.R., Power, J.D., Repovs, G., Anticevic, A., Braver, T.S., 2013b. Multi-task connectivity reveals flexible hubs for adaptive task control. Nat. Neurosci. 16 (9), 1348–1355. doi:10.1038/nn.3470.

Cowan, N., 2019. Short-term memory based on activated long-term memory: a review in response to Norris (2017). Psychol. Bull. 145 (8), 822–847. doi:10.1037/bul0000199.

Cowan, N., Johnson, T.D., Saults, J.S., 2005. Capacity limits in list item recognition: evidence from proactive interference. Memory 13 (3–4), 293–299. doi:10.1080/09658210344000206.

Curtis, C.E., D'Esposito, M, 2003. Persistent activity in the prefrontal cortex during working memory. Trends Cogn. Sci. (Regul. Ed.) 7 (9), 415–423. doi:10.1016/S1364-6613(03)00197-9.

Deschamps, I., Baum, S.R., Gracco, V.L., 2014. On the role of the supramarginal gyrus in phonological processing and verbal working memory: evidence from rTMS studies. Neuropsychologia 53, 39–46. doi:10.1016/j.neuropsychologia.2013.10.015.

D'Esposito, M., Postle, B.R., 2015. The cognitive neuroscience of working memory. Annu. Rev. Psychol. 66, 115–142. doi:10.1146/annurev-psych-010814-015031.

D'Esposito, M., Postle, B.R., Rypma, B, 2000. Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. Executive Control and the Frontal Lobe: Current Issues. Springer, Berlin, Heidelberg, pp. 3–11. doi:10.1007/s002210000395.

Di, X., Biswal, B.B., 2019. Toward task connectomics: examining whole-brain task modulated connectivity in different task domains. Cereb. Cortex 29 (4), 1572–1583. doi:10.1093/cercor/bhy055.

Dohmatob, E., Dumas, G., Bzdok, D., 2020. Dark control: the default mode network as a reinforcement learning agent. Hum. Brain Mapp. 41 (12), 3318–3341. doi:10.1002/hbm.25019.

Dosenbach, N.U.F., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., Petersen, S.E., 2006. A core system for the implementation of task sets. Neuron 50 (5), 799–812. doi:10.1016/j.neuron.2006.04.031.

Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn. Sci. (Regul. Ed.) 14 (4), 172–179. doi:10.1016/j.tics.2010.01.004.

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., Bishop, S., Nimmo-Smith, I., 2008. Goal neglect and Spearman's g: competing parts of a complex task. J. Exp. Psychol. Gen. 137 (1), 131–148. doi:10.1037/0096-3445.137.1.131.

Elton, A., Gao, W., 2015. Task-positive functional connectivity of the default mode network transcends task domain. J. Cogn. Neurosci. 27 (12), 2369–2381. doi:10.1162/jocn_a_00859.

Emch, M., von Bastian, C.C., Koch, K., 2019. Neural correlates of verbal working memory: an fMRI meta-analysis. Front. Hum. Neurosci. 13, 180. doi:10.3389/fnhum.2019.00180.

Finc, K., Bonna, K., Lewandowska, M., Wolak, T., Nikadon, J., Dreszer, J., Duch, W., Kühn, S., 2017. Transition of the functional brain network related to increasing cognitive demands. Hum. Brain Mapp. 38 (7), 3659–3674. doi:10.1002/hbm.23621.

Fisher, R.A., 1921. On the 'probable error' of a coefficient of correlation deduced from a small sample. Metron 1, 1–32.

Formica, S., González-García, C., Senoussi, M., and Brass, M. (2020). Neural oscillations track the maintenance and proceduralization of novel instructions. https://www.biorxiv.org/content/biorxiv/early/2020/11/25/2020.01.20.912162.full.pdf. 10.1101/2020.01.20.912162.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. U. S. A. 102 (27), 9673–9678. doi:10.1073/pnas.0504136102.

Gerchen, M.F., Kirsch, P., 2017. Combining task-related activation and connectivity analysis of fMRI data reveals complex modulation of brain networks. Hum. Brain Mapp. 38 (11), 5726–5739. doi:10.1002/hbm.23762.

González-García, C., Formica, S., Wisniewski, D., Brass, M., 2020. Frontoparietal action-oriented codes support novel instruction implementation. Neuroimage 226, 117608. doi:10.1016/j.neuroimage.2020.117608.

Gordon, E.M., Stollstorff, M., Vaidya, C.J., 2012. Using spatial multiple regression to identify intrinsic connectivity networks involved in working memory performance. Hum. Brain Mapp. 33 (7), 1536–1552. doi:10.1002/hbm.21306.

Greve, A., Cooper, E., Kaula, A., Anderson, M.C., Henson, R., 2017. Does prediction error drive one-shot declarative learning? J. Mem. Lang. 94, 149–165. doi:10.1016/j.jml.2016.11.001.

Hampshire, A., Daws, R.E., Neves, I.D., Soreq, E., Sandrone, S., Violante, I.R., 2019. Probing cortical and sub-cortical contributions to instruction-based learning: regional specialisation and global network dynamics. Neuroimage 192, 88–100. doi:10.1016/j.neuroimage.2019.03.002.

Hartstra, E., Kühn, S., Verguts, T., Brass, M., 2011. The implementation of verbal instructions: an fMRI study. Hum. Brain Mapp. 32 (11), 1811–1824. doi:10.1002/hbm.21152.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6 (2), 65–70.

Kim, H., 2010. Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval. Neuroimage 50 (4), 1648–1657. doi:10.1016/j.neuroimage.2010.01.051.

Lee, S.W., O'Doherty, J.P., Shimojo, S., 2015. Neural computations mediating one-shot learning in the human brain. PLoS Biol. 13 (4), e1002137. doi:10.1371/journal.pbio.1002137.

Liefooghe, B., Braem, S., Meiran, N., 2018. The implications and applications of learning via instructions. Acta Psychol. (Amst) 184, 1–3. doi:10.1016/j.actpsy.2017.09.015.

Liefooghe, B., Houwer, J.de, 2018. Automatic effects of instructions do not require the intention to execute these instructions. J. Cogn. Psychol. 30 (1), 108–121. doi:10.1080/20445911.2017.1365871.

Liefooghe, B., Jim, A., de Houwer, J., 2021. Automatic effects of covert practice. Q. J. Exp. Psychol. (Hove) 74 (10), 1697–1708. doi:10.1177/17470218211007138.

Liefooghe, B., Wenke, D., Houwer, J.de, 2012. Instruction-based task-rule congruency effects. J. Exp. Psychol. Learn. Mem. Cogn. 38 (5), 1325–1335. doi:10.1037/a0028148.

Logan, G.D., 1988. Toward an instance theory of automatization. Psychol. Rev. 95 (4), 492–527. doi:10.1037/0033-295X.95.4.492.

Logie, R.H., 2011. The functional organization and capacity limits of working memory. Curr. Dir. Psychol. Sci. 20 (4), 240–245. doi:10.1177/0963721411415340.

Meiran, N., Cohen-Kdoshay, O., 2012. Working memory load but not multitasking eliminates the prepared reflex: further evidence from the adapted flanker paradigm. Acta Psychol. (Amst) 139 (2), 309–313. doi:10.1016/j.actpsy.2011.12.008.

Meiran, N., Cole, M.W., Braver, T.S., 2012. When planning results in loss of control: intention-based reflexivity and working-memory. Front. Hum. Neurosci. 6, 104. doi:10.3389/fnhum.2012.00104.

Meiran, N., Liefooghe, B., Houwer, J.de, 2017. Powerful instructions: automaticity without practice. Curr. Dir. Psychol. Sci. 26 (6), 509–514. doi:10.1177/0963721417711638.

Meiran, N., Pereg, M., Kessler, Y., Cole, M.W., Braver, T.S., 2015. The power of instructions: proactive configuration of stimulus-response translation. J. Exp. Psychol. Learn. Mem. Cogn. 41 (3), 768–786. doi:10.1037/xlm0000063.

Mohr, H., Wolfensteller, U., Betzel, R.F., Mišić, B., Sporns, O., Richiardi, J., Ruge, H., 2016. Integration and segregation of large-scale brain networks during short-term task automatization. Nat. Commun. 7, 13217. doi:10.1038/ncomms13217.

Mohr, H., Wolfensteller, U., Ruge, H., 2018. Large-scale coupling dynamics of instructed reversal learning. Neuroimage 167, 237–246. doi:10.1016/j.neuroimage.2017.11.049.

Monsell, S., Graham, B., 2021. Role of verbal working memory in rapid procedural acquisition of a choice response task. Cognition 214, 104731. doi:10.1016/j.cognition.2021.104731.

Muhle-Karbe, P.S., Duncan, J., Baene, W.de, Mitchell, D.J., Brass, M, 2017. Neural coding for instruction-based task sets in human frontoparietal and visual cortex. Cereb. Cortex 27 (3), 1891–1905. doi:10.1093/cercor/bhw032.

Mumford, J.A., Davis, T., Poldrack, R.A., 2014. The impact of study design on pattern estimation for single-trial multivariate pattern analysis. Neuroimage 103, 130–138. doi:10.1016/j.neuroimage.2014.09.026.

Mumford, J.A., Turner, B.O., Ashby, F.G., Poldrack, R.A., 2011. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 59 (3), 2636–2643. doi:10.1016/j.neuroimage.2011.08.076.

Nour, M.M., Dahoun, T., McCutcheon, R.A., Adams, R.A., Wall, M.B., Howes, O.D., 2019. Task-induced functional brain connectivity mediates the relationship between striatal D2/3 receptors and working memory. Elife 8. doi:10.7554/eLife.45045.

Palenciano, A.F., González-García, C., Arco, J.E., Pessoa, L., Ruz, M., 2019. Representational organization of novel task sets during proactive encoding. J. Neurosci. 39 (42), 8386–8397. doi:10.1523/JNEUROSCI.0725-19.2019.

Pereg, M., Meiran, N., 2019. Rapid instructed task learning (but not automatic effects of instructions) is influenced by working memory load. PLoS One 14 (6), e0217681. doi:10.1371/journal.pone.0217681.

Pereg, M., Shahar, N., Meiran, N., 2019. Can we learn to learn? The influence of procedural working-memory training on rapid instructed-task-learning. Psychol. Res. 83 (1), 132–146. doi:10.1007/s00426-018-1122-4.

Raven, J., 2003. In: McCallum, R.S. (Ed.). In: Handbook of Nonverbal Assessment, (Vol. 95,. Springer US, pp. 223–237. doi:10.1007/978-1-4615-0153-4_11.

Rigotti, M., Barak, O., Warden, M.R., Wang, X.-J., Daw, N.D., Miller, E.K., Fusi, S, 2013. The importance of mixed selectivity in complex cognitive tasks. Nature 497 (7451), 585–590. doi:10.1038/nature12160.

Rissman, J., Gazzaley, A., D'Esposito, M, 2004. Measuring functional connectivity during distinct stages of a cognitive task. Neuroimage 23 (2), 752–763. doi:10.1016/j.neuroimage.2004.06.035.

Ritchey, M., Cooper, R.A., 2020. Deconstructing the posterior medial episodic network. Trends Cogn. Sci. (Regul. Ed.) 24 (6), 451–465. doi:10.1016/j.tics.2020.03.006.

Ruge, H., Karcz, T., Mark, T., Martin, V., Zwosta, K., Wolfensteller, U., 2018. On the efficiency of instruction-based rule encoding. Acta Psychol. (Amst) 184, 4–19. doi:10.1016/j.actpsy.2017.04.005.

Ruge, H., Schäfer, T.A., Zwosta, K., Mohr, H., Wolfensteller, U., 2019. Neural representation of newly instructed rule identities during early implementation trials. Elife 8. doi:10.7554/eLife.48293.

Ruge, H., Wolfensteller, U., 2010. Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. Cereb. Cortex 20 (7), 1656–1667. doi:10.1093/cercor/bhp228.

Ruge, H., Wolfensteller, U., 2013. Functional integration processes underlying the instruction-based learning of novel goal-directed behaviors. Neuroimage 68, 162–172. doi:10.1016/j.neuroimage.2012.12.003.

Ruge, H., Wolfensteller, U., 2015. Distinct fronto-striatal couplings reveal the double-faced nature of response-outcome relations in instruction-based learning. Cogn. Affect. Behav. Neurosci. 15 (2), 349–364. doi:10.3758/s13415-014-0325-4.

Rugg, M.D., Vilberg, K.L., 2013. Brain networks underlying episodic memory retrieval. Curr. Opin. Neurobiol. 23 (2), 255–260. doi:10.1016/j.conb.2012.11.005.

Schneider, W., Shiffrin, R.M., 1977. Controlled and automatic human information processing: I. Detection, search, and attention. Psychol. Rev. 84 (1), 1–66. doi:10.1037/0033-295x.84.1.1.

Sestieri, C., Corbetta, M., Romani, G.L., Shulman, G.L., 2011. Episodic memory retrieval, parietal cortex, and the default mode network: functional and topographic analyses. J. Neurosci. 31 (12), 4407–4420. doi:10.1523/JNEUROSCI.3335-10.2011.

Sliwinska, M.W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F., Devlin, J.T., 2012. Early and sustained supramarginal gyrus contributions to phonological processing. Front. Psychol. 3, 161. doi:10.3389/fpsyg.2012.00161.

Sobrado, A., Palenciano, A.F., González-García, C., and Ruz, M. (2021). The effect of task demands on the neural patterns generated by novel instruction encoding. https://www.biorxiv.org/content/biorxiv/early/2021/06/16/2021.03.08.434338.full.pdf 10.1101/2021.03.08.434338.

Spreng, R.N., Stevens, W.D., Chamberlain, J.P., Gilmore, A.W., Schacter, D.L., 2010. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. Neuroimage 53 (1), 303–317. doi:10.1016/j.neuroimage.2010.06.016.

Theeuwes, M., Liefooghe, B., Schryver, M.de, Houwer, J.de, 2018. The role of motor imagery in learning via instructions. Acta Psychol. (Amst) 184, 110–123. doi:10.1016/j.actpsy.2017.05.002.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289. doi:10.1006/nimg.2001.0978.

Utevsky, A.V., Smith, D.V., Huettel, S.A., 2014. Precuneus is a functional core of the default-mode network. J. Neurosci. 34 (3), 932–940. doi:10.1523/JNEUROSCI.4227-13.2014.

Vatansever, D., Manktelow, A.E., Sahakian, B.J., Menon, D.K., Stamatakis, E.A., 2017a. Angular default mode network connectivity across working memory load. Hum. Brain Mapp. 38 (1), 41–52. doi:10.1002/hbm.23341.

Vatansever, D., Menon, D.K., Manktelow, A.E., Sahakian, B.J., Stamatakis, E.A., 2015. Default mode dynamics for global functional integration. J. Neurosci. 35 (46), 15254–15262. doi:10.1523/JNEUROSCI.2135-15.2015.

Vatansever, D., Menon, D.K., Stamatakis, E.A., 2017b. Default mode contributions to automated information processing. Proc. Natl. Acad. Sci. U. S. A. 114 (48), 12821–12826. doi:10.1073/pnas.1710521114.

Vatansever, D., Smallwood, J., Jefferies, E., 2021. Varying demands for cognitive control reveals shared neural processes supporting semantic and episodic memory retrieval. Nat. Commun. 12 (1), 2134. doi:10.1038/s41467-021-22443-2.

Visscher, K.M., Miezin, F.M., Kelly, J.E., Buckner, R.L., Donaldson, D.I., McAvoy, M.P., Bhalodia, V.M., Petersen, S.E., 2003. Mixed blocked/event-related designs separate transient and sustained activity in fMRI. Neuroimage 19 (4), 1694–1708. doi:10.1016/s1053-8119(03)00178-2.

Wang, X., Gao, Z., Smallwood, J., Jefferies, E., 2021. Both default and multiple-demand regions represent semantic goal information. J. Neurosci. 41 (16), 3679–3691. doi:10.1523/JNEUROSCI.1782-20.2021.

Wechsler, D., 1997. WAIS-III: Administration and Scoring Manual, 3rd ed. Psychological Corp.

Westphal, A.J., Wang, S., Rissman, J., 2017. Episodic memory retrieval benefits from a less modular brain network organization. J. Neurosci. 37 (13), 3523–3531. doi:10.1523/JNEUROSCI.2509-16.2017.

Woolgar, A., Afshar, S., Williams, M.A., Rich, A.N., 2015. Flexible coding of task rules in frontoparietal cortex: an adaptive system for flexible cognitive control. J Cogn. Neurosci. 27 (10), 1895–1911. doi:10.1162/jocn_a_00827.

Xia, M., Wang, J., He, Y., 2013. Brainnet Viewer: a network visualization tool for human brain connectomics. PLoS One 8 (7), e68910. doi:10.1371/journal.pone.0068910.

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106 (3), 1125–1165. doi:10.1152/jn.00338.2011.