

In Press in Developmental Psychology

RUNNING HEAD: Creativity Training

Training in the adolescent brain: an fMRI training study on divergent thinking

Sietske W. Kleibeuker\*, Claire E. Stevenson\*, Laura van der Aar, Sandy Overgaauw, Anna  
C. van Duijvenvoorde, Eveline A. Crone

\* Shared first authors

Department of Psychology, Leiden University, Leiden, the Netherlands

Leiden Institute for Brain and Cognition (LIBC), Leiden University, Leiden, the Netherlands

## **Abstract**

Prior research suggests that adolescence is a time of enhanced sensitivity for practice and learning. In this study we tested the neural correlates of divergent thinking training in 15-16-year-old adolescents relative to an age-matched active control group. All participants performed an alternative uses task, a valid measure to test divergent thinking, while fMRI images were acquired before and after a training program. In between the two scanning sessions the experimental group completed 2 weeks of divergent thinking training (8 sessions) and the control group completed 2 weeks of rule switching training (8 session). A group x time interaction demonstrated stable divergent thinking performance for the experimental group, whereas in the control group performance declined. Generating alternative uses (experimental task condition) relative to generating ordinary characteristics of objects (control task condition) was associated with increased activation in the supramarginal gyrus (SMG), angular gyrus (AG) and middle temporal gyrus (MTG). Test-retest analyses showed that within-individuals-activation in these regions was stable over time in both groups. Changes in alternative uses fluency over time, however, were positively associated with changes in superior lateral PFC activation over time. Together, the results indicate that core brain regions for creativity (SMG, AG, MTG) are consistently recruited in adolescence, and that changes in performance are associated with changes in activation in lateral PFC.

Keywords: adolescence, training, creativity, neuroimaging, prefrontal cortex

## **Introduction**

Some parents may wonder why adolescents have difficulty planning their homework while at the same time they are experts in switching between social media devices, designing clothes or redecorating their rooms. Adolescence is the transition period between childhood and adulthood during which individuals gain their independence from their parents and rapidly adjust to new social contexts (Crone and Dahl, 2012). Even though cognitive control functions still increase between adolescence and adulthood (Luna et al., 2001), both animal (Johnson and Wilbrecht, 2011) and human research (Kleibeuker et al., 2013a) has shown that adolescence relative to adulthood is a period of increased potential for flexible thinking. That is, adolescence seems to be a period of increased flexibility that is well suited for novel insights and creative problem solving, which is beneficial in a period which asks for rapid adjustments to changing social demands and gaining independence (Crone and Dahl, 2012).

Adolescence is also a period of life that is identified by a significant amount of time involved in training and education, both in- and outside school settings. As such, a better understanding of learning mechanisms and training effects is especially useful in this age period. Prior research indicates that the adolescent brain is indeed sensitive to the effects of training of different cognitive functions including working memory and mathematical skills (Jolles et al., 2010; Quin et al, 2004), showing increased activation patterns in prefrontal and parietal regions after training. These brain regions have previously been found to develop relatively late in adolescence (Gogtay, 2004; Sowell, 1999; see Crone and Ridderinkhof, 2010; Blakemore and Choudhury, 2006). A question that remains, is how the adolescent brain adapts to training of cognitive functions that require flexible and divergent thinking rather than controlled and convergent thinking. In the current study, we test the benefits of training

creativity in adolescents, by examining neural responses to problems that require divergent thinking, before and after two weeks of divergent thinking training.

Divergent thinking is an important component of creativity and it involves the ability to think of novel solutions for encountered problems (Torrance, 1965). A well-known task to measure divergent thinking is the Alternative Uses Task (AUT) where individuals are asked to think of as many possible ways to use an object, such as an umbrella (e.g., ‘storage place for stuffed animals’). Solutions should be novel and appropriately useful (Guilford, 1967; Kim, 2008). The answers can be categorized in terms of fluency (number of possible answers), flexibility (number of times individuals switch between different categories; i.e., storage, protection, decoration, etc.), and originality (uniqueness of answers). Prior studies revealed that adults and adolescents perform equally well on the AUT in terms of flexibility (Kleibeuker et al., 2013a), but adults outperform adolescents in baseline fluency (Kleibeuker et al., 2013b) and originality (Kleibeuker et al., 2013b; Stevenson et al., 2014).

Research has shown that divergent thinking performance can be enhanced by two weeks of training, whereas longer duration has only minimally positive effects on training outcomes (Scott et al., 2004; Fink et al., 2006). These findings have been demonstrated in adults (Kienitz et al., 2014), adolescents (Stevenson et al., 2014) and children (Scott et al., 2004), although transfer effects to other domains are debated (Kaufman and Baer, 2009). A recent developmental study demonstrated training benefits after eight sessions during two weeks of divergent thinking training for AUT fluency and originality. Intriguingly, this study revealed larger changes for adolescents aged 13-16 years than for adults for the domain of originality, which transferred to a new divergent thinking task (Stevenson et al., 2014). These findings led to the hypothesis that adolescence is an important time window for creativity development. A key question concerns which neural regions contribute to these training

related changes in adolescence; we utilize the same training paradigm employed by Stevenson et al. (2014) to investigate this.

Recently, several studies have examined the neural correlates of divergent thinking by examining neural activity while individuals perform the AUT (Abraham et al., 2012; Cousijn et al., 2014a,b; Fink et al., 2006, 2009, 2010; Kleibeuker et al., 2013b). These findings have consistently shown activity in a network of regions including the supramarginal gyrus (SMG), middle temporal gyrus (MTG), angular gyrus (AG), and prefrontal cortex (PFC) (Abraham et al., 2012; Fink et al., 2009, 2010; Kleibeuker et al., 2013b). Prior studies related activity in angular gyrus to insight (Aziz-Zadeh et al., 2009; Bechtereva et al., 2004); MTG to imagination of object use (Beauchamp and Martin, 2007; Johnson-Frey et al., 2005; Lewis, 2006) and PFC to executive semantic control processes including processes that enable context appropriate semantic retrieval (Whitney et al., 2011). The region that most strongly correlated with fluency performance of the AUT is the lateral PFC, such that higher fluency scores were associated with increased neural activation in the lateral PFC (Kleibeuker et al., 2013b). Developmental comparisons between adolescents aged 15-17 years and adults showed that adolescents recruited the same network as adults (Kleibeuker et al., 2013b). However, adults had a relatively higher activation in the lateral PFC than adolescents, which was accounted for by performance differences. Intriguingly, a divergent thinking training study using EEG showed higher synchronization in frontal alpha activity after two weeks of AUT training (Fink et al., 2006), but it is not yet known how this relates to neural activation changes in the different brain regions involved in divergent thinking. In a recent fMRI study, in which adult participants were subjected to an extensive verbal creativity training for three weeks, activations increased mainly in temporo-parietal regions including bilateral SMG and left (posterior) MTG (Fink et al., 2015). These results were interpreted as suggesting that

training increased semantic control that is necessary to effectively combine available semantic information to produce novelty. A more commonly used approach to enhance creative performance in neuroimaging research involves cognitive stimulation by providing (moderately) creative ideas. Several studies have found this approach to be effective and training has been associated with functional changes of the (left) temporo-parietal cortex, including the left MTG and PFC regions (Fink et al., 2012; Wei et al., 2014). An important question addressed in the current study concerns how activity in these regions changes through creative thinking training in adolescence. This will allow us to have a better understanding of how training-related changes in adolescents take place.

In this study we compared behavior and neural activity in 15-16-year-old adolescents while performing an AUT task in the scanner before and after two weeks of divergent thinking training, and we compared this to AUT activity in an active control group who performed a rule switching training program (see also Stevenson et al., 2014). The control training was similar in terms of effort and time investment. The use of such an active control group diminishes possible confounding effects (see Klingberg, 2010 for a discussion on this approach). This design allowed us to test the following questions. First, given the small number of longitudinal studies on creativity (Claxton et al., 2005; Scott et al., 2004), we tested whether neural activity was stable within individuals over time, which would inform us about whether creativity is a stable person specific ability or varies within persons over time. This was done by computing the IntraClass Correlation Coefficients for regions which are commonly active during the AUT, which include the SMG, MTG, and AG. This was found to be a valid method for determining test-retest stability in prior research (Van den Bulk et al., 2013). Second, we tested whether neural activity was enhanced by training in these same regions. Finally, we tested the hypothesis that the lateral PFC is sensitive to individual

differences in performance improvements, given that prior studies showed that this region is most sensitive to individual differences in AUT fluency (Kleibeuker et al., 2013b).

## **Methods**

### **Participants**

A total of 32 adolescents (18 male) aged 15 to 16 years participated in this study. Participants were recruited through local advertisements. The sample size was based on prior studies using the same task showing robust neural activity in adolescents using a cross-sectional design (Kleibeuker et al., 2013) and on working memory training studies with adults (Jolles et al., 2010; Olesen & Klingberg, 2004). Power analyses using effect sizes based on Kleibeuker et al. (2013) revealed that a total sample size of 32 participants would have enough power (.80 with  $\alpha = .05$ ) to identify similar effects.

All participants were healthy, right-handed, and MRI compatible (i.e., no braces or metal implants). None of the participants reported a history of neurological or psychiatric disorders. Participants were recruited from postal code areas with average to high social economic backgrounds. However, no direct information on social economic status (SES) was obtained.

The participants were randomly divided into two groups matched for gender; a Divergent Thinking training group (DT-group) and a Rule Switching training group (RS-group). There was a small but significant difference in age between the two groups; the DT-group was a few months younger than the RS-group ( $M_{AU} = 15.84$ ,  $SD = .11$ ;  $M_{RS} = 16.2$ ,  $SD = .14$ ;  $t_{30} = 2.40$ ,  $p = 0.02$ ). To check for group differences in intelligence, IQ-scores were estimated based on two subtests from the Wechsler Intelligence Scale for Children (WISC;

Wechsler 2004): Similarities and Digit Span. Groups did not significantly differ in IQ-scores ( $t_{30} = 1.16, p = .26$ ).

All participants as well as their primary caregiver signed informed consent before participation. Participants were financially rewarded for their participation. The Medical Ethics Committee from the Leiden University Medical Center (LUMC) approved the study.

### **Cognitive assessments during pretest and posttest**

Participants completed a battery of tests for which some of the results are presented elsewhere (Cousijn, et al., 2014). For this study, we examined performance on the two training tasks during the scan session and in between scan sessions. These tasks are described in detail below.

#### AU/OC-scanner task

At pretest and posttest the participants performed an adapted version of the Alternative Uses Test (AUT; Guilford, 1967) inside the MRI scanner while neural activity was measured (see also Kleibeuker et al. 2013b for a detailed task description). This task measures divergent thinking in the verbal domain. The task consisted of two conditions: 1) the free-association-related Alternative Uses (AU) condition during which participants had to think of as many appropriate alternative and original uses of common objects as possible (i.e., use a shoe as a baseball bat); and 2) the more general verbal-ability-related Object Characteristics (OC) condition during which participants had to think of as many ordinary characteristics of common objects as possible (i.e., a shoe fits on a foot). Each trial started with a 3 second instruction screen. Then, a written item was presented in the middle of the screen for 15 s with the text ‘Ordinary Characteristics’ or ‘Alternative Uses’ on the top of the screen during OC and AU trials, respectively, to remind participants of the trial condition (see Figure 1).



Immediately after the target screen, an evaluation screen appeared for 3 s. Participants indicated how many solutions they had found by pressing one of four buttons on a left/right button-box that was attached to their left/right leg respectively; the left middle finger for 0 or 1 solution, the left index finger for 2 solutions, the right index finger for 3 solutions and the right middle finger for 4 or more solutions. Each trial was preceded by a fixation cross that was presented for a variable duration (1.1–7.7 s) to optimize the event.

--Figure 1 about here--

The AU/OC-scanner task consisted of 60 trials (30 alternative uses and 30 ordinary characteristics) divided over three blocks of 8.2 minutes, during which 30 unique words were presented (once in the AU and once in the OC condition). Two sets of 30 words were created, matched on word length, number of syllabi, and word frequency. Participants saw different word sets during the pretest and posttest. The order of the two word sets was counterbalanced across participants. The total task time was 30 minutes.

The task was programmed in E-Prime (version 2.0). The performance measures for both the AU and OC condition were the number of times within a session that participants indicated (with button press) that they generated 0 or 1 solution, 2 solutions, 3 solutions, and 4 or more solutions. A composite fluency score was required to use as a regressor in data analysis; this was calculated as follows:  $1 * \text{button\_0/1 responses} + 2 * \text{button\_2} + 3 * \text{button\_3 responses} + 4 * \text{button\_4 responses}$  (see also Kleibeuker et al., 2013b). This was calculated for each condition (AU-score and OC-score) for the pretest and the posttest

session. This composite score was correlated with the fluency score of a different AUT administered outside of the scanner ( $r=.54, p<.01$ ) (Cousijn et al., 2014).

### Local Global Task (LGT)

Rule switching was assessed behaviorally at pretest and posttest after scanning with the LGT (adapted from Huizinga et al., 2006) that was used as the active control task in the control training condition. During the task, the target stimuli were large squares and rectangles (global figures) consisting of small squares or rectangles (local figures). Participants were instructed to focus on either the global or local aspect of the target shape (i.e., square or rectangle) given the presence of a global or local cue. The global cue consisted of a large square and rectangle presented on the left and right side of the target respectively. The local cue consisted of a small square and rectangle presented on the sides of the target. Participants pressed a right or left response button corresponding to the correct (global or local) shape of the target that coincided with the cue on the left or right side of the target. Each trial started with a 500ms presentation of the global or local cue, after which the target stimulus appeared. Participants had 3500ms to respond to the target stimulus (see also Huizinga et al., 2006).

The task consisted of a global and a local block (each 50 trials, presentation order counterbalanced over participants), followed by a switch block (160 trials). From the switch block trials, median response times (RTs) and proportion correct of local-global switch and control trials were computed.

## **Training**

During the two weeks between the pretest and posttest the DT-group followed an 8-session AUT-training and the RS-group followed an 8-session Local-Global Task switching (LGT); both groups trained online from home.

The AUT-training task resembled the AU trials of the AU/OC-scanner task; participants had to generate as many appropriate alternative and original uses of common objects as possible. Nine different objects were presented during each training session. The objects were different from the AU trials from the AU/OC-scanner task objects. Word length, number of syllabi, and word frequency were matched across sessions. Participants were given 2 minutes to enter their solutions for each object.

The RS-training was similar to the LGT at pretest and posttest. Each training session consisted of 8 blocks of 40 trials with self-paced breaks in-between blocks. The blocks contained alternating global and local mini blocks of 4 trials.

An online training schedule was created with the participants and their primary caregivers. The experimenters monitored the training progress using an online system. In case a training session was forgotten, a text-message was sent to the participant's cell phone to suggest another date to catch up. If a second training session was missed, the participant was called to discuss a new schedule. Each training session lasted approximately 20 minutes.

Four participants missed one AUT-training and two participants missed one LGT-training. All participants completed at least 7 training sessions.

## **MRI data collection**

A 3T MRI scanner (Philips Intera, Best, The Netherlands) with a standard whole-head coil was used for image acquisition. Resting state fMRI data were acquired at the start of the

pretest and posttest session. A total of 140 volumes were acquired resulting in a scan time of 5 minutes. The results from the resting state scan have been reported elsewhere (Cousijn, Zanolie et al., 2014).

Next, participants completed three runs of the AU-OC task, each lasting 8.3 minutes, during which 226 volumes per run were acquired. Bold signal was measured with a T2\* gradient-echo EPI sequence (TR 2.2s, TE 30ms, 38 slices, slice thickness 2.75mm, FOV 220x220mm, in-plane resolution 2.75x2.75mm, flip angle 80°, sequential slice acquisition). The first 2 volumes of each run were discarded in order to allow for equilibration of T1 saturation effects. Finally, a high resolution T1 structural scan was acquired for anatomical reference (T1 turbo field echo, TR 9.8ms, TE 4.6ms, 140 slices, slice thickness 1.2mm, FOV 224x178mm, in-plane resolution 0.88x0.88mm, flip angle 8°). Head motion was restricted by using foam inserts between the head and the head coil. Visual stimuli were projected onto a screen in the magnet bore that could be viewed through a mirror attached to the head coil.

SPM8 software ([www.fil.ion.ucl.ac.uk](http://www.fil.ion.ucl.ac.uk)) was used for image preprocessing and analyses. Images were corrected for slice-time differences, followed by rigid body motion correction. Functional volumes were spatially normalized to individual T1 scans and subsequently to T1 templates based on MNI305 stereotaxic space (Cocosco et al., 1997) using a 12-parameter affine transformation together with a nonlinear transformation involving cosine base functions. Data were resampled to 3 mm cubic voxels. Functional volumes were smoothed using an 8mm full-width half-maximum 3D Gaussian kernel.

## **fMRI data analyses**

SPM8 software was used for image analyses. For each participant, the functional time series were modeled by a series of events convolved with a canonical hemodynamic response function (HRF). Trials were modeled separately based on condition (AU or OC), with the presentation as onset time and a duration of 15 seconds, and entered in a general linear model along with a basic set of cosine functions to high-pass filter the data, and a covariate for run effects. In addition, the instruction screen preceding the AU and OC trials and the evaluation screen after trials were modeled separately (each with duration of 0 ms).

The least square parameter estimates of height of best fitting canonical HRF for each condition were used in pairwise contrasts (AU - OC). Resulting first level contrast images, computed on a subject-by-subject basis, were calculated for pretest (T1) and posttest (T2) data separately. Additional images were created for activation changes from T1 to T2 for the contrast AU - OC using the ImCalc tool in SPM8. These calculations resulted in another set of first level contrast images  $((AU_{T2}-OC_{T2})-(AU_{T1}-OC_{T1}))$  which will be referred to as  $AU > OC_{T2-T1}$ . The first level contrast images were submitted to group analyses. At the group level, contrasts between conditions were computed by performing one-tailed t-tests on these contrasts, treating participants as a random effect, and two-sample t-tests to compare training groups. Whole brain fMRI analyses were FWE cluster corrected for multiple comparisons at an initial threshold of  $p < .001$  uncorrected (following Woo et al., 2014). We further conducted whole-brain regression analyses on the contrast  $AU > OC$  and  $AU > OC_{T2-T1}$  to test for general brain-behavior relations and relations between brain activation and behavioral changes, respectively. AU-scores at T1 were entered as covariate of interest to find regions showing general brain-behavior relations. Difference-scores ( $AU\text{-}score_{T2-T1}$ ) were entered as covariate of interest to find regions showing relations between brain- and behavioral changes. We applied the threshold of  $p < .001$  uncorrected with at least 10 contiguous voxels to overcome

the relatively low power inherent to analyses of individual differences/ type II error. Results are reported in the MNI305 stereotaxic space. Brain regions were determined based on the SPM anatomy toolbox v1.8 (Eickhoff et al., 2005, 2006, 2007).

### **Region-of-interest (ROI) Analyses**

Region-of-interest (ROI) analyses were performed with MarsBaR toolbox in SPM8 (Brett et al., 2002). The output ‘contrast estimates’ was used. Contrast estimates were derived for each condition relative to baseline (i.e., OC-baseline, AU-baseline).

ROIs were analyzed to test for possible training related effects within divergent thinking related brain regions. This ROI approach was chosen in addition to the whole-brain analyses, because analyses at the ROI level have more power to detect smaller differences in the task-related brain regions. These analyses included Time (pretest T1, posttest T2) and Condition (AU, OC) as within-subject variables and Training Group (DT, RS) as between-subjects variable. We applied repeated measures ANOVAs on ROIs derived from the contrast AU>OC at T1. If the region spanned a large area, the region was masked using anatomical ROIs derived from the MarsBaR anatomical toolbox.

## **Results**

### **Initial comparisons**

Initial comparisons comprised of tests for differences between training conditions on each of the pretest tasks using Analysis of Variance (ANOVAs) with Training as between-subjects factor. Descriptive statistics for all measures per training condition are shown in

Table 1. None of the measures revealed differences between the two groups (all main effects and interaction effects:  $p$ 's > .05) showing that random assignment was successful and that there were no differences between the two groups prior to the training.

--Table 1 about here--

### **Pretest to posttest changes on training tasks**

#### *AU/OC scanner task*

To examine the behavioral effects of training on alternative uses (AU) and ordinary characteristics (OC) retrieval, we examined whether the pattern of responses (button choice) in Number Solutions (0/1, 2, 3, 4+), differed between Training Groups (DT, RS) in the two Task Conditions (AU, OC) over Time (T1, T2) with a repeated measures ANOVA (Huynh-Feldt estimates reported when sphericity was violated). Results are depicted in Figure 2. Participants generated significantly more solutions for OC than AU trials (Number Solutions x Task Condition:  $F(2.4, 90) = 48.77, p < .001, \eta_p^2 = .62$ ; Number Solutions:  $F(2.1, 90) = 4.64, p = .01, \eta_p^2 = .13$ ). In addition, there was a tendency to provide fewer solutions at T2 than T1 for both AU and OC trials (Number Solutions x Time:  $F(3, 90) = 2.50, p = .064, \eta_p^2 = .08$ ; Number Solutions x Task Condition x Time:  $F(3, 90) = 2.64, p = .054, \eta_p^2 = .08$ ). Although not significant, Figure 2 also shows a marginal difference in performance change between Training groups from T1 to T2 which appears to indicate relatively more stable performance for the DT Training group compared to the RS Training group from T1 to T2 (Training x Time:  $F(1, 90) = 2.06, p = .10, \eta_p^2 = .09$ ). No other reliable Time or Training Group differences were found (main and interaction effects, all  $p$ 's > .05).

---Figure 2 about here---

In addition to the analyses on button choices, we investigated the Training and Time effects on the AU-composite scores, which represent divergent thinking performances and are, as such, applied in the fMRI analyses. To this end, we conducted a repeated measures ANOVAs using AU-composite scores as dependent variable with Time (T1, T2) as within-subject variable and Training Group (DT, RS) as between-subject variable. . A significant main effect of Time ( $F(1,30) = 6.31, p = .02, \eta_p^2 = .17$ ), and an interaction effect of Time x Training Group ( $F(1,30) = 4.14, p = .05, \eta_p^2 = .12$ ) were found (see Figure 3). Post hoc analyses of the interaction effect show that performance in the RS group decreased from T1 to T2 ( $F(1,30) = 13.49, p < .01, \eta_p^2 = .47$ ), whereas results for the DT group remained stable ( $F(1,30) = .09, p = .77, \eta_p^2 = .01$ ). Thus, rather than increases for the DT group, we observed decreases for the RS group. This has also been reported in other studies on creativity in adults (Fink et al., 2015; De Dreu, Baas and Nijstad, 2008; see discussion).

---Figure 3 about here---

### *LGT*

To test for training related effects on LGT accuracy (% correct) and reaction times, repeated measures ANOVAs were performed with Condition (Switch, Control) and Time (T1, T2) as within-subject variables and Training Group (DT, RS) as between-subject variable (see Table 1 for descriptive statistics) for reaction times and accuracy.



For reaction times, there were main effects for Time ( $F(1,30) = 30.3, p < .001, \eta_p^2 = .50$ ), showing a general decrease of RTs from T1 to T2, and for Condition ( $F(1,30) = 38.02, p < .001, \eta_p^2 = .56$ ), showing longer RTs for switch relative to control trials. In addition, there was a significant interaction effect of Time x Condition ( $F(1,30) = 8.23, p < .01, \eta_p^2 = .22$ ), showing larger decreases for switch than for control trials at T2 relative to T1. No main or interaction effects were present for Training Group (all  $p$ 's  $> .05$ ). For accuracy, no significant effects were observed for Condition or Time and no interactions with Training Group were observed (all  $p$ 's  $> .05$ ).

Analyses focusing specifically on switch trial performance were conducted with  $p$ -values corrected for multiple comparisons. ANCOVAs with posttest performance on switch trials as dependent variable revealed significant effects for pretest performance as covariate (RT:  $F(1,29)=15.74, p<.001$ ; accuracy:  $F(1,29)=12.53, p_{\text{corr}}=.002$ ) and a significant effect for Training Group as between-subject factor for RT ( $F(1,29)=19.94, p<.001$ ), but not accuracy ( $F(1,29)=0.34, p=.57$ ). This indicates that participants were faster and generally more accurate on the posttest switch trials; furthermore, the RS group was significantly faster than the DT group on posttest switch trials after correcting for pretest performance.

#### *DT and RS training and relations*

Figure 4 shows how the two training groups progressed in their respective tasks across the eight sessions. Although there was some initial fluctuation in performance there were no significant differences in performance between the first and last sessions. The analyses of this data are reported in Cousijn et al. (2014).

To test for possible relations between AU and LGT performance, bivariate correlation analyses were performed on AU fluency scores and LGT switch cost accuracy and RT performances at pretest. None of the correlations were significant (all  $p$ 's > .05).

--Figure 4 about here--

## **fMRI results**

### *T1 (pre-test) analyses*

To extract the neural correlates of creative idea generation we conducted whole-brain voxel-wise t-tests on activation levels for the contrast alternative uses (AU) > ordinary characteristics (OC) across all participants (N = 32) at T1. Results revealed a number of regions including left supramarginal gyrus (SMG), left and right middle temporal gyrus (MTG), and left angular gyrus (AG) (see Figure 5; Table 2), which is in line with prior studies (Kleibeuker et al, 2013b; Fink et al., 2009, 2010).

--Figure 5 about here—

---Table 2 about here—

To test for group differences in brain activation patterns during the AU/OC scanner task at pretest, we performed whole-brain two sample t-tests on the contrasts AU>OC. No

significant results were observed for the contrasts AU>OC, reassuring that groups did not differ at T1.

### *Test-retest stability*

To determine whether the activations related to AU>OC remained relatively stable over time after AU and RS training, we performed test-retest reliability analyses on the ROIs derived from the T1 AU>OC contrast. To this end, we calculated IntraClass Correlation Coefficients (ICCs) for the contrast AU>OC for all participants together as well as for the two training groups (DT, RS) separately. Results are presented in Table 3A (all participants) and 3B (training groups separately). For regions that spanned clusters encompassing different anatomical regions, the cluster was masked with anatomical regions from the Marsbar AAL atlas (reported separately in Table 3).

Anatomical masking was performed for three large clusters and this resulted in the following subclusters. Cluster 1 (-24, -39, -15) was separated in left fusiform gyrus, left inferior temporal gyrus, and left precuneus. Cluster 2 (-51, -66, 3) was separated in left middle temporal gyrus, left angular gyrus, and left occipital gyrus. Cluster 3 (57, -66, 15) was masked with right middle temporal gyrus. Three other clusters that were active in this contrast fell within anatomical boundaries and were therefore not masked. These included the left medial superior frontal gyrus (-3, 63, 18), left supramarginal gyrus (-63, -27, 36) and parahippocampal gyrus (21, -48, 9) (see Figure 5b).

Across all participants, ICCs ranged from 0 to .65 with highest correlations for the right MTG (ICC = .60,  $p = .006$ ), left AG (ICC = .65,  $p = .003$ ), and left SMG (ICC = .64,  $p = .003$ ).

--Table 3 about here--

### *T2>T1 analyses*

Our analyses on brain activation training effects were twofold. First, whole-brain analyses were applied on the T2>T1 contrast images for the contrasts AU>OC and OC>AU. No significant results were observed for the contrasts AU>OC or for the opposite contrast OC>AU. Possibly, with the current threshold and testing across the whole brain results in underpowered analyses for detected subtle training effects.

Second, we analyzed possible training related effects within divergent thinking related brain regions using ROI analyses. Therefore we applied repeated measures ANOVAs on ROIs from the contrast AU>OC at T1.

The Time (2) x Condition (2) x Training Group (2) repeated measures ANOVA for the separate ROIs only showed a time effect at trend level for left MTG/angular gyrus/occipital lobe cluster (cluster with coordinates: -51 -66 3;  $F(1,30) = 4.42$ ,  $p = .058$ ,  $\eta_p^2 = .13$ ) as well as a trend level Time x Training Group interaction effect ( $F(1,30) = 3.30$ ,  $p = .08$ ,  $\eta_p^2 = .10$ ). The interaction showed a larger increase of activation over time for the DT-training group relative to the RS-training group for both the AU and OC trials. Post hoc analyses on the two groups separately showed significant changes over time for the DT group ( $F(1,15) = 15.95$ ,  $p = .001$ ), but not for the RS group ( $F(1,15) = .027$ ,  $p = .87$ ). No Training effects or interactions

were found for the other regions. Given that the left MTG/AG/occipital lobe cluster spanned several regions, separate analyses were performed with masked regions with anatomical ROIs from Marsbar AAL to separate these effects. These separate analyses showed that the interaction effect was driven by the MTG (Time x Training Group interaction;  $F(1,30) = 3.89$ ,  $p = .058$ ,  $\eta_p^2 = .12$ ). Post hoc analyses on the two groups separately showed significant changes over time for the DT group ( $F(1,15) = 14.25$ ,  $p = .002$ ,  $\eta_p^2 = .49$ ), but not for the RS group ( $F(1,15) = .003$ ,  $p = .96$ ,  $\eta_p^2 = .00$ ). (see Figure 5c).

### *T2>T1 analyses, relations with performance change*

Our final aim was to test for brain activation patterns that were related to changes in AUT performance over time. To this end, we conducted whole brain regression analyses on the contrasts AU>OC<sub>T2-T1</sub> with AU difference scores (T2-T1). The analyses did not survive cluster corrections. Given that we had a priori hypothesis about brain-behavior correlates in the prefrontal cortex, the threshold was lowered to  $p < .001$ , a minimum 10 contiguous voxels of 10. Using this threshold, a significant positive relation was observed in the left cerebellum (peak coordinate: -30, -75, -21; 150 voxels) and left Middle Frontal Gyrus (MFG; peak coordinate: -39, 24, 9; 14 voxels; see Figure 6a,b).

Post hoc ROI analyses on the left MFG showed that the effect remained significant when controlling for Training Group, for activation at T1 (AU>OC), and for AU fluency scores at T1. These results indicate that the regression analyses outcomes are not due to individual differences at T1 or to group differences in training related changes. The independence of training group is also indicated by whole brain regression analyses for the two training groups separately ( $p_{\text{uncorr}} < .001$ ; > 10 consecutive voxels). For both groups the

performance regression results showed regions within the left MFG that overlapped with the region found in the analyses that included all participants (see Figure 6c; DT-training: peak voxel at -42, 51, 0; RS-training: peak voxel at -33, 54, 12).

---Figure 6 about here---

#### *Additional tests for lateral PFC performance relations*

To test for commonalities with an earlier study on AUT neural responses using the same task in both adolescents and adults, bivariate correlation analyses were performed on a predefined ROI in the lateral PFC. This ROI was derived from a whole brain regression analysis on the contrast AU>OC with AU-score as covariate (see Kleibeuker et al., 2013b), which resulted in a cluster in right middle frontal gyrus (coordinates: 42, 42, 21). Using the contrast values of this ROI, correlations were estimated in the current study between AU>OC changes from T1 to T2 and AU difference scores (AU-score at T2 minus AU-score at T1). We found a significant positive correlation between right middle frontal gyrus and AU difference scores ( $r=.46$ ,  $p=.008$ ). These results support the importance of right lateral frontal functioning in creative idea generation.

## **Discussion**

The aim of the study was to examine changes in neural activations related to divergent thinking practice in adolescents. To this end, 15-16-year-olds were scanned while performing an adapted version of the AUT (Guilford, 1956, 1967; Kleibeuker et al., 2013b) before and after two-weeks of alternative uses training. Behavioral and neural changes were compared to

an age-matched active control group who practiced with a rule-switching task for the same training period. Behavioral results showed a decline in divergent thinking performance after training for the control group (the rule switching training group). The divergent thinking group showed no change in performance before and after the training. The fMRI results yielded three important findings: 1) activations of the creative ideation related network remained relatively stable over time, 2) were not sensitive to training manipulations, except for MTG at trend level; and 3) right lateral frontal cortex activation changes were associated with changes in divergent thinking over time. The discussion is organized in line with these findings.

### **Training effects of divergent thinking training**

The behavioural results showed a decrease for AU scores for the rule switch group, whereas performances for the divergent thinking group remained stable over time. Similar findings have previously been reported in adults (Fink et al., 2015). The behavioural outcomes could be interpreted as a negative effect of rule switch training on divergent thinking performance. However, the lack of significant correlations between AU scores and LGT performances from both the current and a previous study (Cousijn et al., 2014) is in contradiction with such an interpretation: If rule switch training affects divergent thinking negatively, one might expect to find negative correlations between rule switch performances and divergent thinking performances. A second possible explanation for the behavioural interaction effect is that the RS group lacked motivation to perform during the second session. However, if this was the case, one might have expected a similar negative effect for other test

performances. Results of additional tests administered outside the scanner (results are described in Cousijn et al. 2014) do not support this explanation.

Alternatively, the results might be influenced by secondary aspects such as the test environment. Indeed, MRI scanner environment has been related to changes in cognitive functioning showing both increases and decreases in cognitive control performances (Van Maanen, Forstmann, Keuken, Wagenmakers, Heathcote, 2016). Results have been interpreted e.g., in terms of stress-induced increase in attentional control (Hommel, Fischer, Colzato, van den Wildenberg, & Cellini, 2012; see also Plessow, Fischer, Kirschbaum, & Goschke, 2011) or arousal (Koten, Langner, Wood, & Willmes, 2013), and decreased attentional focus (Assecondi et al., 2010). Interestingly, both these interpretations (increased arousal and decreased attentional focus) could account for better divergent thinking performance during a first scanner session relative to a second scanner session. Indeed, prior studies have revealed positive effects of moderate levels of arousal on creative fluency and originality (De Dreu, Baas, Nijstad, 2008; Byron et al., 2010), whereas attentional focus is negatively associated with creativity performances (e.g., Wiley and Jarosz, 2012, Ansburg and Hill, 2003). As such, it is reasonable that potential training related increases in creative thinking are hidden by negative session effects, and future research should focus on possible environmental effects in training studies.

## **4.2 Stability of creative ideation network activations**

The stability (ICC) analyses showed that activations of the core network related to creative ideation, represented by regions that are more active during alternative uses generation than ordinary characteristics retrieval at T1, remain relatively stable over time. The



network is dominated by parietal-temporal regions including bilateral MTG, left AG and left SMG. Thus, the results are consistent with prior research that indicates critical involvement of parieto-temporal regions in verbal creative thinking (Fink et al., 2010,; see also Arden et al., 2010 for an overview). Several studies have shown that the AG, SMG (posterior), and MTG are involved during semantic tasks (e.g., Binder et al., 2009; Jung-Beeman, 2005; Vigneau et al., 2006). In part, the semantic information processing might be focused specifically on tools. Indeed, the MTG and SMG have repeatedly been related to tool use and action knowledge, including semantic information of tools and imaginative tool use (Beauchamp and Martin, 2007; Johnson-Frey et al., 2005; Johnson-Frey, 2004; Lewis, 2006). Accordingly, processing of these types of information are likely profitable when thinking about alternative uses of objects. Recent experimental research also indicates a specific function for the left posterior MTG together with the left inferior frontal gyrus (IFG) in semantic processing. These regions have been related to executive mechanisms that direct semantic retrieval according to the task or context at hand (e.g., Jung-Beeman, 2005; Whitney et al., 2012). Possibly, semantic retrieval is demanded more during thinking about alternative uses than thinking about ordinary characteristics. Notably, especially the MTG, AG and SMG showed strong stability within individuals over time in both training groups, suggesting that this is a core network that is engaged for divergent thinking performance. The stability in this network may also explain why training effects were limited. Future studies should investigate if activation in this network is related to specific individual traits.

Taken together, the within-individual consistency in neural activity across the two measurements was highest in those areas that are also most consistently reported in single session measurements, including the AG, SMG and (right) MTG, further confirming the

relevance of this network for divergent thinking (Abraham et al., 2012; Bechtereva et al., 2004; Benedek et al., 2014; Fink et al., 2009, 2010, 2014; Kleibeuker et al., 2013b).

### **Effects of divergent thinking training on left MTG activity**

The comparison of neural activation before and after training indicates that two weeks of divergent thinking training leads to a relative increase in activity in the left MTG during idea generation. These findings are consistent with prior research on verbal creative thinking training. One possible interpretation is that the left MTG is important for demanding executive semantic processes (Whitney et al., 2012). The change in left MTG activity was found for both the alternative uses and the ordinary characteristics trials, suggesting a general increase in a process that is demanded for both types of trials, which was enhanced by two weeks of divergent thinking training. The change in left MTG activity was not related to individual differences in performance change. It is possible that these relations are due to small sample sizes and that relations will be found in studies with larger samples. Yet, the absence of a performance relation is consistent with the hypothesis that the change represents a general ideation effect which is not specific to thinking of alternative uses. Given that the interaction between Training Group x Time was at trend level ( $p=.058$ ), they should be replicated in future studies.

### **PFC activity and performance change**

In separate comparisons for both training groups, performance correlations were found for regions within the lateral PFC. These findings are consistent with an earlier study that showed that especially lateral PFC activity was associated with individual differences in

alternative uses fluency (Kleibeuker et al., 2013b). In the current study, change in activity in a region in the left middle lateral PFC, adjacent to (minimally overlapping) the region in Kleibeuker et al. 2013b, was related to a change in alternative uses fluency. Thus, those individuals increasing in fluency over time showed more activity in this region of lateral PFC over time, whereas those individuals decreasing in fluency over time showed less activity in lateral PFC over time. These change by change correlations have also been found in a prior study examining longitudinal changes in performance monitoring in children, adolescents and adults (Koolschijn et al., 2011).

Interestingly, in a prior study this region of lateral PFC was less active in adolescents than adults during divergent thinking as measured with the AUT, which was accounted for by differences in fluency performance (Kleibeuker et al., 2013b). Although the exact function of the lateral PFC in the applied paradigm and in divergent thinking in general is still under debate, it has been related to working memory, monitoring, and inhibition of common answers (Carlson et al., 2000; Chavez et al., 2004; Chavez-Eakle et al., 2007; Martindale, 1999; Zabelina & Robinson, 2010). It should be noted that divergent thinking benefits by divergent thinking training were characterized by an absence of a decrease in fluency, whereas the rule switch training group showed a decrease in fluency over time. More studies are necessary to better understand the dynamics of divergent thinking performance across multiple measurements.

## **Limitations**

This study had several limitations which should be addressed in future research. First, the behavioral outcome fluency, i.e. number of ideas generated, was indicated in the scanner by a button press; this did not allow us to validate the responses. In-scanner fluency indicators

are related to fluency scores on AU tasks outside of the scanner (Kleibeuker et al., 2013b) and fluency scores are by definition related to originality ratings (Silvia et al., 2008). However, originality is another important AUT measure and we recommend future studies to apply voice recording in the scanner (e.g., Benedek et al., 2014) to enable originality scoring. Second, the active control condition performed a rule switching task which helped us control for the cognitive flexibility component of the AU/OC task; however, a verbal fluency task is closer in nature to the AUT and may form an alternative control condition. Third, the study only included adolescents, and future studies should test if training effects are different for children and adults (e.g., Stevenson et al., 2014). Fourth, the study focused on practice, and future studies should use other training designs as well, such as using feedback including exposure to others' ideas (Dugosh and Paulus, 2005; Fink et al., 2010; Paulus and Nijstad, 2003). Finally, this study examined effects over a relatively short time period of two weeks, measuring neural correlates before and after training. Future studies should examine long term effects of training as well, for example across several months (e.g., Bott et al., 2014; Scott et al., 2004), and incorporate more measurements across time to increase the understanding of the training process and account for possible session effects.

## **Conclusion**

Taken together, this study was the first to show that over time, there is stability in the neural responses during divergent thinking, as indicated by moderate to high ICCs. This result shows that creativity can be reliably assessed and has some level of stability that is difficult to show based on behavioral findings alone. Individual differences in performance changes were most strongly related to changes in lateral PFC activity. These findings are interesting in relation to studies that have shown that especially executive control processes, which rely

heavily on PFC, develop further during adolescence. Future studies should test the hypothesis that adolescence is a sensitive period for training related changes in neural activity by comparing adolescents and adults, given that prior behavioral studies have suggested that training benefits in the domain of divergent thinking are larger for adolescents than for adults (Stevenson et al., 2014). These training effects may rely more on strategy use related to recruitment of PFC than in recruitment of the temporal-parietal network, which seems more reflective of differences between and not within individuals.

## References

- Abraham, A, Pieritz, K, Thybusch, K, Rutter, B, Kröger, S, Schweckendiek, J, ... Hermann, C (2012): Creativity and the brain: uncovering the neural signature of conceptual expansion. *Neuropsychologia*, 50(8):1906-1917.
- Arden, R, Chavez, LGT, Grazioplene, R, Jung, RE (2010): Neuroimaging creativity: a psychometric view. *Behav Brain Research*, 214(2):143-156.
- Assecondi, S, Vanderperren, K, Novitskiy, N, Ramautar, J R, Fias, W, Staelens, S, Lemahieu, I (2010). Effect of the static magnetic field of the MR-scanner on ERPs: Evaluation of visual, cognitive and motor potentials. *Clinical Neurophysiology*, 121: 672-85.
- Aziz-Zadeh, L, Kaplan, JT, Iacoboni, M (2009): “Aha!”: The neural correlates of verbal insight solutions. *Human Brain Mapping*, 30(3):908-916.
- Beauchamp, MS, Martin, A (2007): Grounding Object Concepts in Perception and Action: Evidence from FMRI Studies of Tools. *Cortex*, 43(3):461-468.
- Benedek, M, Jauk, E, Fink, A, Koschutnig, K, Reishofer, G, Ebner, F, Neubauer, AC (2014): To create or to recall? Neural mechanisms underlying the generation of creative new ideas. *NeuroImage*, 88:125-133.
- Binder, JR, Desai, RH, Graves, WW, Conant, LL (2009): Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12):2767-2796.

- Bott, N, Quintin, E-M, Saggat, M, Kienitz, E, Royalty, A, Hong, DW-C, ... Allan L. Reiss (2014): Creativity training enhances goal-directed attention and information processing. *Thinking Skills and Creativity*, 13:120-128.
- Brett, M, Anton, JL, Valabregue, R, Poline, JB (2002): Region of interest analysis using the MarsBar toolbox for SPM99. In 8th International Congress on Functional Mapping of the Human Brain (pp. 2-6): Retrieved from <http://sourceforge.net/projects/marsbar/files/>
- Byron, K, Khazanchi, S Nazarian, D (2010). The relationship between stressors and creativity: A meta-analysis examining competing theoretical models. *Journal of Applied Psychology*, 95(1): 201-212.
- Carlsson, I, Wendt, PE, Risberg, J (2000): On the neurobiology of creativity. Differences in frontal activity between high and low creative subjects. *Neuropsychologia*, 38(6):873-885.
- Chavez RA, Graff-Guerrero A, Garcia-Reyna JC, Vaugier V, Cruz-Fuentes C (2004). Neurobiology of creativity: preliminary results from a brain activation study. *Salud Mental* 27: 38–46.
- Chávez-Eakle, RA, Graff-Guerrero, A, García-Reyna, J-C, Vaugier, V, Cruz-Fuentes, C (2007): Cerebral blood flow associated with creative performance: a comparative study. *NeuroImage*, 38(3):519-528.
- Claxton, AF, Pannells, TC, Rhoads, PA (2005): Developmental Trends in the Creativity of School-Age Children. *Creativity Research J*, 17(4):327-335.

- Cocosco, CA, Kollokian, V, Kwan, RK-SG, Pike, B, Evans, AC (1997): BrainWeb: Online Interface to a 3D MRI Simulated Brain Database. *NeuroImage*, 5:S425.
- Cousijn, J, Koolschijn, PCMP, Zanolie, K, Kleibeuker, SW, Crone, EA (2014): The Relation between Gray Matter Morphology and Divergent Thinking in Adolescents and Young Adults. *PloS One*, 9(12):e114619.
- Cousijn, J, Zanolie, K, Munsters, RJM, Kleibeuker, SW, Crone, EA (2014): The relation between resting state connectivity and creativity in adolescents before and after training. *PloS One*, 9(9):e105780.
- Crone, EA, Dahl, RE (2012): Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews. Neuroscience*, 13(9):636-50.
- Crone, EA, Donohue, SE, Honomichl, R, Wendelken, C, Bunge, SA (2006): Brain regions mediating flexible rule use during development. *J Neuroscience*, 26(43):11239-47.
- Dietrich, A (2004): The cognitive neuroscience of creativity. *Psychonomic Bulletin & Review*, 11(6):1011-1026.
- Dugosh, KL, Paulus, PB (2005): Cognitive and social comparison processes in brainstorming. *J Experimental Social Psychology*, 41(3):313-320.
- Eickhoff, SB, Heim, S, Zilles, K, Amunts, K (2006): Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *NeuroImage*, 32(2):570-582.



- Eickhoff, SB, Paus, T, Caspers, S, Grosbras, M-H, Evans, AC, Zilles, K, Amunts, K (2007): Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *NeuroImage*, 36(3):511-521.
- Eickhoff, SB, Stephan, KE, Mohlberg, H, Grefkes, C, Fink, GR, Amunts, K, Zilles, K (2005): A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4):1325-1335.
- Fink, A, Grabner, RH, Benedek, M, Neubauer, AC (2006): Divergent thinking training is related to frontal electroencephalogram alpha synchronization. *European J Neuroscience*, 23(8):2241-6.
- Fink, A, Grabner, RH, Benedek, M, Reishofer, G, Hauswirth, V, Fally, M, ... Neubauer, AC (2009): The creative brain: investigation of brain activity during creative problem solving by means of EEG and fMRI. *Human Brain Mapping*, 30(3):734-48.
- Fink, A, Grabner, RH, Gebauer, D, Reishofer, G, Koschutnig, K, Ebner, F (2010): Enhancing creativity by means of cognitive stimulation: evidence from an fMRI study. *NeuroImage*, 52(4):1687-95.
- Fink, A, Weber, B, Koschutnig, K, Benedek, M, Reishofer, G, Ebner, F, ... Weiss, EM (2014): Creativity and schizotypy from the neuroscience perspective. *Cognitive, Affective & Behavioral Neuroscience*, 14(1):378-387.
- Fink, A, Benedek, M, Koschutnig, K, Pirker, E, Berger, E, Meister, S, Neubauer, A.C., Papousek, I. & Weiss, E.M. Training of verbal creativity modulates brain activity in regions associated with language and memory-related demands. *Human Brain Mapping*, 36(10), 4104-4015.

- Genovese, CR, Lazar, NA, Nichols, T (2002): Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870-878.
- Groborz, M, Necka, E (2003): Creativity and Cognitive Control: Explorations of Generation and Evaluation Skills. *Creativity Research J*, 15(2-3):183-197.
- Guilford, JP (1956): Creativity Research: Past, Present and Future. *Frontiers in Creativity Research*:33-66.
- Guilford, JP (1967): Creativity: Yesterday, Today and Tomorrow. *J Creative Behavior*, 1(1):3-14.
- Gutchess, A ., & Park, D C (2006). fMRI Environment Can Impair Memory Performance in Young and Elderly Adults. *Brain Research*, 1099(1), 133–140.
- Hommel, B, Fischer, R, Colzato, LS., van den Wildenberg, WPM., & Cellini, C (2012).The effect of fMRI (noise) on cognitive control. *Journal of Experimental Psychology: Human Perception and Performance*, 38: 290-301.
- Howard-Jones, PA, Blakemore, S-J, Samuel, EA, Summers, IR, Claxton, G (2005): Semantic divergence and creative story generation: an fMRI investigation. *Cognitive Brain Research*, 25(1):240-250.
- Huizinga, M, Dolan, CV, van der Molen, MW (2006): Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia*, 44(11):2017-36.
- Johnson, C, Wilbrecht, L (2011): Juvenile mice show greater flexibility in multiple choice reversal learning than adults. *Developmental Cognitive Neuroscience*, 1(4):540-51.

- Johnson-Frey, SH (2004): The neural bases of complex tool use in humans. *Trends in Cognitive Sciences*, 8(2):71-78.
- Johnson-Frey, SH, Newman-Norlund, R, Grafton, ST (2005): A distributed left hemisphere network active during planning of everyday tool use skills. *Cerebral Cortex*:15(6):681-695.
- Jolles, DD, Kleibeuker, SW, Rombouts, SARB, Crone, EA (2011): Developmental differences in prefrontal activation during working memory maintenance and manipulation for different memory loads. *Developmental Science*, 14(4):713-24.
- Jolles, DD., Grol, M, Van Buchem, MA, Crone EA, Rombouts SARB (2010). Practice effects in the brain: Changes in cerebral activation after working memory practice depend on task demands. *Neuroimage*, 52, 658-668.
- Jung-Beeman, M (2005): Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, 9(11):512-518.
- Kaufman, JC, Baer, J (2009): Is One Dimension Enough? A Response to Simonton's Varieties of (Scientific) Creativity. *Perspectives on Psychological Science*, 4(5):453-454.
- Keating, DP (2004): Cognitive and brain development. In R. M. Lerner & L. Steinberg (Eds.): *Handbook of Adolescent Psychology* (p. 864): John Wiley & Sons.
- Kienitz, E, Quintin, E-M, Saggar, M, Bott, NT, Royalty, A, Hong, DW-C, ... Reiss, AL (2014): Targeted intervention to increase creative capacity and performance: A randomized controlled pilot study. *Thinking Skills and Creativity*, 13:57-66.

- Kim, KH (2008): Meta-Analyses of the Relationship of Creative Achievement to Both IQ and Divergent Thinking Test Scores. *J Creative Behavior*, 42(2):106-130.
- Kleibeuker, SW, De Dreu, CKW, Crone, EA (2013a): The development of creative cognition across adolescence: distinct trajectories for insight and divergent thinking. *Developmental Science*, 16(1):2-12.
- Kleibeuker, SW, Koolschijn, PCMP, Jolles, DD, De Dreu, CKW, Crone, EA (2013b): The neural coding of creative idea generation across adolescence and early adulthood. *Frontiers in Human Neuroscience*; 7; 905.
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, 14: 317-324.
- Koolschijn, PCMP, Schel, MA, de Rooij, M, Rombouts, SARB, Crone, EA (2011): A three-year longitudinal functional magnetic resonance imaging study of performance monitoring and test-retest reliability from childhood to early adulthood. *J Neuroscience*, 31(11):4204-4212.
- Koten, J W, Langner, R, Wood, G, & Willmes, K (2013): Are reaction times obtained during fMRI scanning reliable and valid measures of behavior? *Experimental Brain Research*, 227 :93-100.
- Lewis, JW (2006): Cortical networks related to human use of tools. *The Neuroscientist : A Review J Bringing Neurobiology, Neurology and Psychiatry*, 12(3):211-231.

- Luna, B, Thulborn, KR, Munoz, DP, Merriam, EP, Garver, KE, Minshew, NJ, ... Sweeney, JA (2001): Maturation of widely distributed brain function subserves cognitive development. *NeuroImage*, 13(5):786-93.
- Martindale, C (1999): Biological basis of creativity. In R.J.Sternberg (Ed.), *Handbook of creativity*. Cambridge: Cambridge University Press.
- Nachev, P, Wydell, H, O'Neill, K, Husain, M, Kennard, C (2007): The role of the pre-supplementary motor area in the control of action. *NeuroImage*, 36(2):155-163.
- Olesen, P, Westenberg, H, Klingberg, T (2004) Increased prefrontal and parietal brain activity after training of working memory. *Nature Neuroscience*, 7(1), 75-79.
- Paulus, PB, Nijstad, BA (2003): *Group creativity: Innovation through collaboration* (B. A. Nijstad & P. B. Paulus, Eds.): New York: Oxford University Press. Retrieved from <http://psycnet.apa.org/psycinfo/2003-88061-000>
- Qin Y, Carter CS, Silk EM, Stenger VA, Fissell K, et al. (2004):. The change of the brain activation patterns as children learn algebra equation solving. *Proceedings of the National Academy of Sciences of the United States of America* 101: 5686–5691.
- Scott, G, Leritz, LE, Mumford, MD (2004): The Effectiveness of Creativity Training: A Quantitative Review. *Creativity Research J*, 16(4):361-388.
- Schtereva, NP, Korotkov, AD, Pakhomov, SV, Roudas, MS, Starchenko, MG, Medvedev, S. V (2004): PET study of brain maintenance of verbal creative activity. *Intrnl J Psychophysio*, 53(1):11-20.

- Stevenson, CE, Kleibeuker, SW, de Dreu, CKW, Crone, EA (2014): Training creative cognition: adolescence as a flexible period for improving creativity. *Frontiers in Human Neuroscience*, 8:827.
- Torrance, EP (1965): *Rewarding Creative Behavior; Experiments in Classroom Creativity* (p. 353): Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Van den Bulk, BG, Koolschijn, PCMP, Meens, PHF, van Lang, NDJ, van der Wee, NJA., Rombouts, SARB, Vermeiren, RRJM & Crone, EA (2013). How stable is activation in the amygdala and prefrontal cortex in adolescence? A study of emotional face processing across three measurements. *Developmental Cognitive Neuroscience*, 4: 65-76.
- van Maanen, L, Forstmann, BU, Keuken, MC, Wagenmakers, E-J, & Heathcote, A (2016): The impact of MRI scanner environment on perceptual decision making. *Behavior Research Methods*, 48(1):184-200.
- Vigneau, M, Beaucousin, V, Hervé, P. Y, Duffau, H, Crivello, F, Houdé, O, ... Tzourio-Mazoyer, N (2006): Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*, 30(4):1414-1432.
- Whitney, C, Kirk, M, O'Sullivan, J, Lambon RMA, Jefferies, E (2012): Executive semantic processing is underpinned by a large-scale neural network: revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. *J Cognitive Neuroscience*, 24(1):133-147.
- Wiley J & Jarosz AF (2012): Working Memory Capacity, Attentional Focus, and Problem Solving. *Current Directions in Psychological Science* 21 (4): 258-262.
- Woo, C. W., Krishnan, A. & Wager, T. D. (2014) 'Cluster-extent based thresholding in

fMRI analyses: Pitfalls and recommendations', *Neuroimage*, **91**, pp. 412-419.

Zabelina,DI & Robinson,MD(2010):Creativity as flexible cognitive control. *Psychology of Aesthetics Creativity & Arts*, 4: 136–143.

## Figure captions

Figure 1. Examples of Alternative Uses (AU) and Ordinary Characteristics (OC) trials in the AU/OC scanner task.

Figure 2. Performance on Alternative Uses / Ordinary Characteristics (AU/OC) scanner task for the Divergent Thinking (DT) Training Group and Rule Switch (RS) Training Group at pretest (T1) and posttest (T2). Participants indicated the number of solutions generated with a button press (Number Solutions: 0/1, 2, 3, 4+). More OC than AU solutions were given and fewer solutions on the whole were indicated on posttest (T2) than pretest (T1). Performance over Time for the DT Training Group appears more stable than that of the RS Training Group - the Group x Time interaction was not significant ( $p = .10$ )

Figure 3. Performance on alternative uses trials for the Divergent Thinking (DT) Training Group and Rule Switch (RS) Training Group at pretest (Time 1) and posttest (Time 2). Analyses of composite scores indicated a decrease in provision of alternative uses (AU) solutions for the RS Training Group and stable performance for the DT Training Group.

Figure 4. Divergent thinking (DT) and Rule switching (RS) training performance (see also Cousijn et al. (2014)). (A) Mean (standard error) fluency scores for each AUT-training session in the DT group ( $n=16$ ). (B) Median (standard error) reaction time of local-global switch trials of each LGT-training session in the RS group ( $n=16$ ). \*  $p<0.05$ , \*\*  $p<0.001$ .



**Figure 5.** (A) Neural regions showing activity for Alternative Uses (AU) > Ordinary Characteristics (OC) at Pretest (T1), averaged across all participants ( $N=32$ ). This resulted in activation in left supramarginal gyrus (SMG), bilateral middle temporal gyrus (MTG) and left angular gyrus, thresholded at FWE cluster correction,  $p < .001$ . See Table 2 for coordinates, and Supplementary Figure 1 for analyses per training group. (B) Regions of interest based on larger cluster presented in Figure 5A and Table 2/3. Regions that spanned multiple areas were masked using the Marsbar anatomical atlas. Regions are presented in the colors: **Cluster 1:** Left fusiform gyrus in dark blue (masked:  $x=-27, y=-38, z=-16$ ), left inferior temporal gyrus in blue (masked:  $x=-38, y=-37, z=-15$ ), left precuneus in light blue (masked:  $x=-11, y=-54, z=14$ ). **Cluster 2:** left middle temporal gyrus in green (masked:  $x=-48, y=-64, z=11$ ), left angular gyrus in yellow (masked:  $x=-46, y=-70, z=28$ ), left occipital gyrus in orange (masked:  $x=-42, y=-72, z=20$ ). **Cluster 3:** Right middle temporal gyrus in dark red (masked:  $x=50, y=-60, z=13$ ). **Region 4:** left supramarginal gyrus in red ( $x=-63, y=-27, z=36$ ), **Region 5:** Left superior medial gyrus in sea blue ( $x=-3, y=63, z=18$ ). **Region 6:** Parahippocampal gyrus is not displayed. (C) Contrast values for ROI analyses of left MTG (anatomically masked, displayed in green in B). Separate analyses for each Training Group revealed a significant increase in activation for  $T2 > T1$  for the Divergent Thinking (DT) Training Group for both conditions (AU and OC), but no Time effect for the Rule Switch (RS) Training Group. \*  $p < .05$ .

**Figure 6.** (A) Regression analysis for Alternative Uses (AU) > Ordinary Characteristics (OC) Post test ( $T2$ ) > Pre test ( $T1$ ) with Alternative Uses (AU) test fluency  $T2 > T1$  as predictor, averaged across all participants ( $N=32$ ). Threshold:  $p < .001$ , uncorrected for multiple comparisons, at least 10 contiguous voxels. (B) Data of ROI analyses for the cluster in the left

Middle Frontal Gyrus (MFG) are presented with circles for DT Training Group and triangles for RS Training Group for illustration purposes. (C) Results for the regression analyses across all participants ( $N = 32$ ; red), for the DT training group ( $N = 16$ ; yellow), and RS training group ( $N = 16$ ; blue) separately.

Figure 1

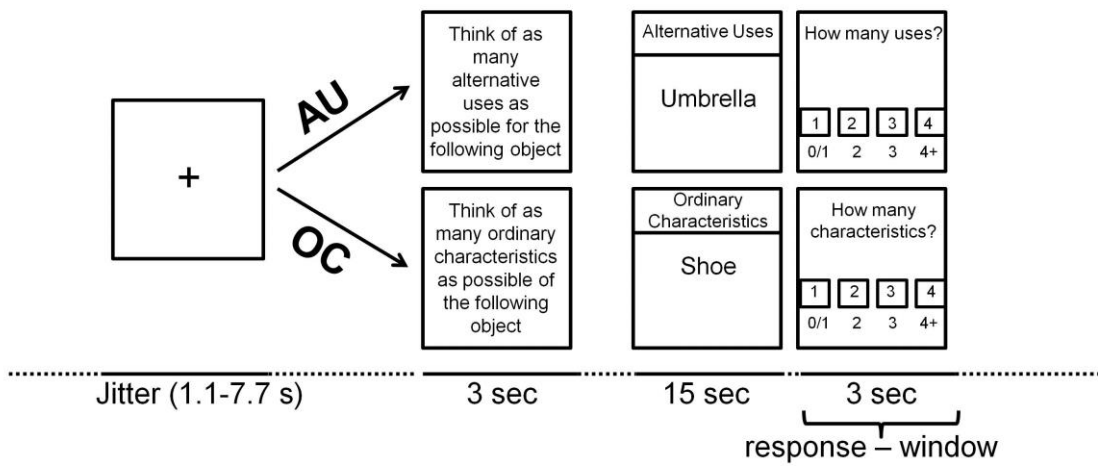


Figure 2

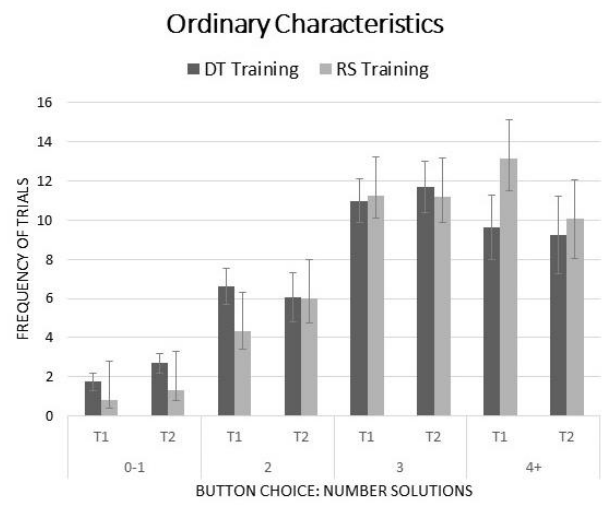
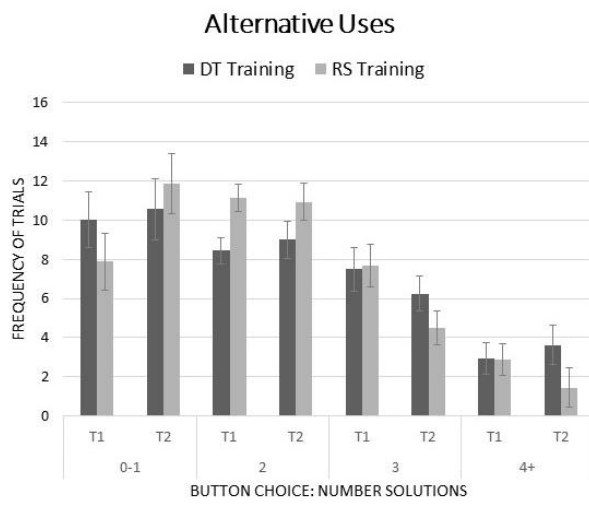


Figure 3

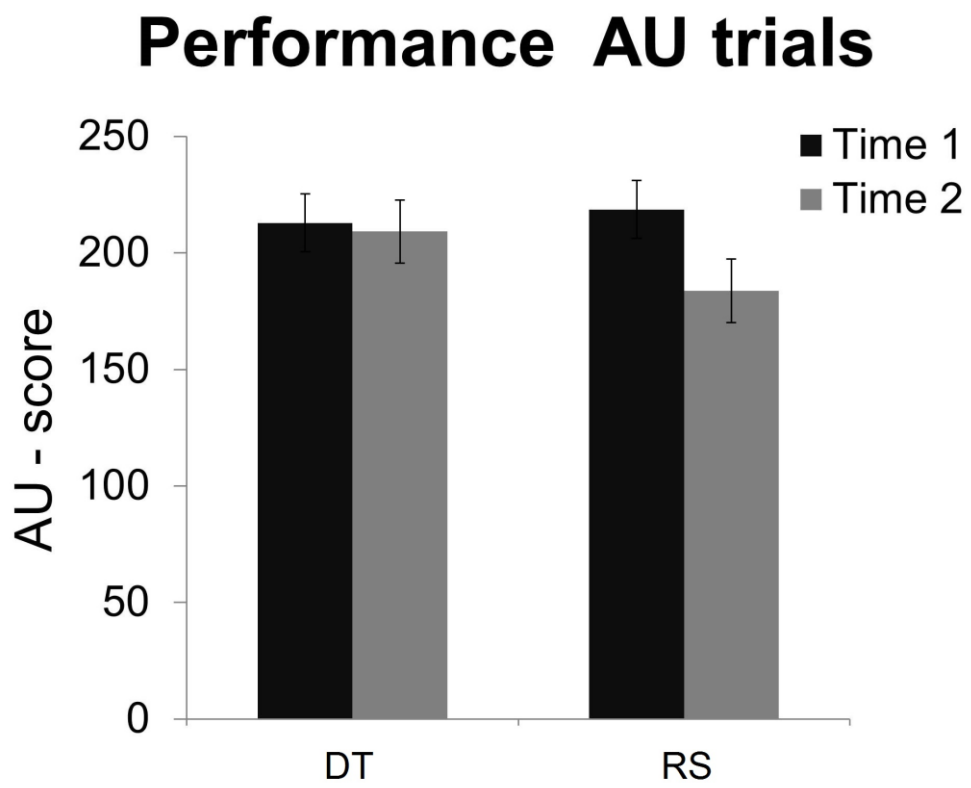


Figure 4

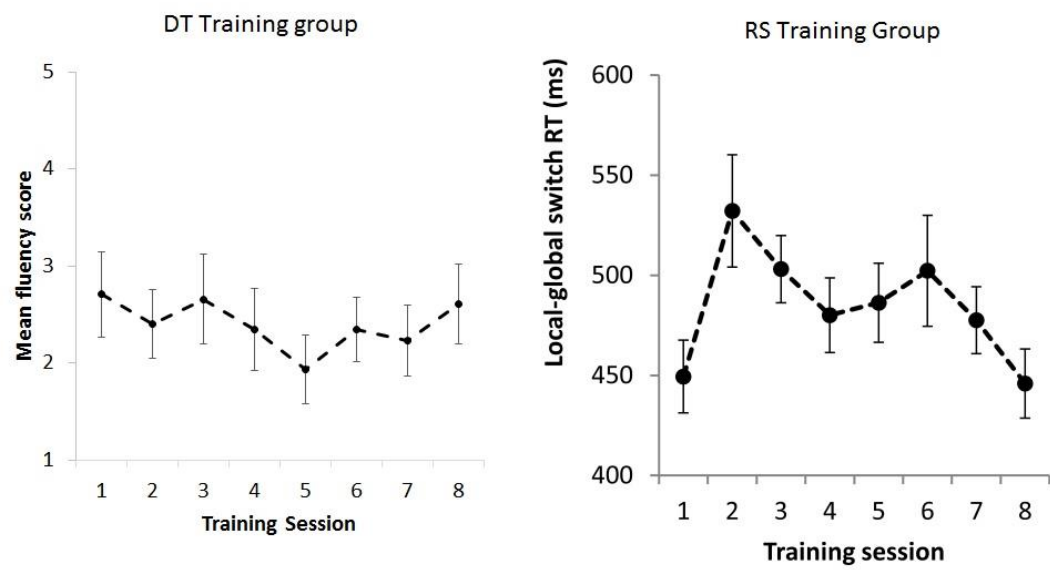


Figure 5

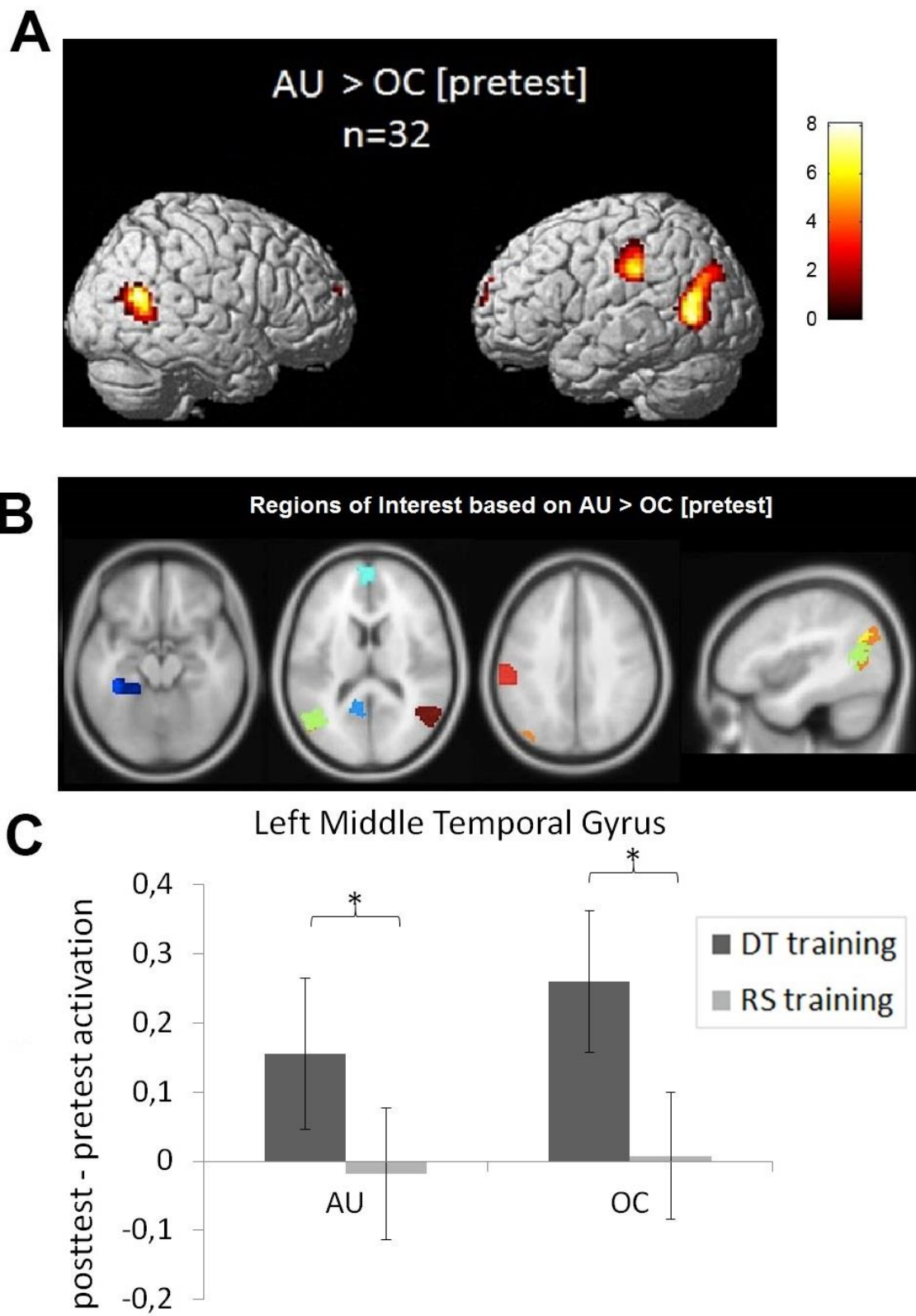


Figure 6

## Regression $AU > OC_{T2-T1}$ with $AU\text{-}score_{T2-T1}$

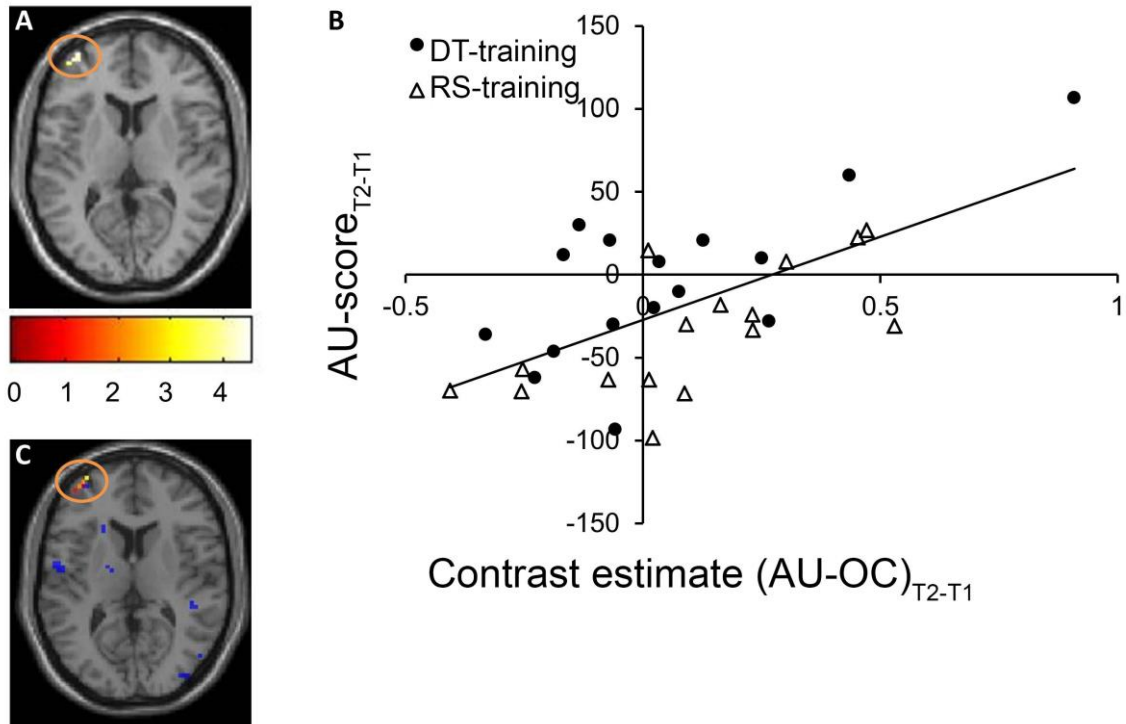




Table 1.

Descriptive statistics for the behavioural measures (AU/OC scanner task and local global task) by training group and testing session. DT=Divergent Thinking, RS = Rule Switching, AU= Alternative Uses, OC = Ordinary Characteristics, RT= Reaction time, ms = Milliseconds, M=Mean, SD= Standard Deviation

Task	DT-training group		RS-training group	
	pre-test	post-test	pre-test	post-test
	M (SD)	M (SD)	M (SD)	M (SD)
AU/OC-scanner task				
AU-score	2.13 (0.59)	2.09 (0.68)	2.19 (0.38)	1.84 (0.36)
OC-score	3.00 (0.46)	2.93 (0.52)	3.24 (0.32)	3.04 (0.50)
Local Global Task				
% correct Control	94.04 (3.06)	96.19 (2.16)	90.55 (8.64)	90.85 (7.08)
% correct Switch	94.41 (6.70)	94.67 (8.07)	89.95 (8.50)	90.85 (7.09)
RT <i>ms</i> Control	443.75 (96.67)	385.50 (57.09)	416.41 (79.15)	329.03 (49.80)
RT <i>ms</i> Switch	509.22 (129.11)	428.16 (69.26)	525.25 (211.31)	354.66 (47.23)

Table 2.

Coordinates for the contrast Alternative Uses (AU) > Ordinary Characteristics (OC) at Time point 1 (T1) across participants (N=32), thresholded at  $p < .001$  FWE cluster corrected.

Brain regions	L/R	K	Z-value	MNI coordinates		
			peak voxel	x	Y	z
<b>Cluster 1:</b> fusiform gyrus, precuneus, calcarine gyrus	L	317	4.86	-24	-39	-15
			4.65	-9	-57	15
			4.45	-36	-36	-15
<b>Cluster 2:</b> middle temporal gyrus, angular gyrus, middle occipital lobe	L	344	5.72	-51	-66	3
			4.78	-42	-78	33
			4.21	-42	-69	27
<b>Cluster 3:</b> middle temporal gyrus, precuneus, calcarine gyrus	R	202	5.89	57	-66	15
supramarginal gyrus	L	141	5.22	-63	-27	36
medial superior frontal gyrus	L	118	4.84	-3	63	18
parahippocampal gyrus	R	96	4.36	21	-48	9

Abbreviations: MNI = montreal neurological institute; L = left hemisphere; R = right hemisphere.

Table 3.

A. IntraClass Correlations (ICC) for Regions of Interest (ROIs) to indicate stability over time derived from the contrast Alternative Uses (AU) > Ordinary Characteristics (OC) at Time point 1 (T1; pre-training) across all participants.

	All participants	
	ICC	significance
<b>Cluster 1:</b> l fusiform gyrus, l precuneus, l calcarine gyrus (-24 -39 -15)	<b>0.46</b>	<b>0.046</b>
l fusiform gyrus*	0	0.63
l inferior temporal gyrus*	0	0.74
l precuneus *	<b>0.55</b>	<b>0.02</b>
<b>Cluster 2:</b> l middle temporal gyrus, l angular gyrus, l middle occipital lobe (-51 -66 3)	<b>0.52</b>	<b>0.02</b>
l middle temporal gyrus*	<b>0.57</b>	<b>0.01</b>
l angular gyrus*	<b>0.65</b>	<b>0.003</b>
l occipital lobe*	<b>0.46</b>	<b>0.045</b>
<b>Cluster 3:</b> r middle temporal gyrus, r precuneus, r calcarine gyrus (57 -66 15)	<b>0.60</b>	<b>0.006</b>
r middle temporal gyrus*	<b>0.60</b>	<b>0.007</b>
<b>Region 4:</b> l supramarginal gyrus (-63 -27 36)	<b>0.64</b>	<b>0.003</b>
<b>Region 5:</b> l superior medial gyrus (-3 63 18)	<b>0.54</b>	<b>0.02</b>
<b>Region 6:</b> r parahippocampal gyrus (21 -48 9)	<b>0.60</b>	<b>0.006</b>

A. IntraClass Correlations (ICC) for Regions of Interest (ROIs) to indicate stability over time derived from the contrast Alternative Uses (AU) > Ordinary Characteristics (OC) at Time point 1 (T1; pre-training) for the training groups separately.

	AU-training group	RS-training group
--	-------------------	-------------------

	ICC	significance	ICC	significance
<b>Cluster 1:</b> l fusiform gyrus, l precuneus, l calcarine gyrus (-24 -39 -15)	0.31	0.24	<b>0.61</b>	<b>0.04</b>
l fusiform gyrus*	0	0.74	0	0.60
l inferior temporal gyrus*	0	0.72	0	0.56
l precuneus *	0.39	0.18	<b>0.80</b>	<b>0.002</b>
 <b>Cluster 2:</b> l middle temporal gyrus, l angular gyrus, l middle occipital lobe (-51 -66 3)	0.46	0.12	<b>0.63</b>	<b>0.03</b>
l middle temporal gyrus*	0.37	0.19	<b>0.75</b>	<b>0.006</b>
l angular gyrus*	<b>0.73</b>	<b>0.008</b>	0.51	0.09
l occipital gyrus*	0.52	0.08	0.32	0.23
 <b>Cluster 3:</b> r middle temporal gyrus (57 -66 15)	<b>0.59</b>	<b>0.048</b>	0.58	0.05
r middle temporal gyrus*	0.58	0.05	0.58	0.05
<b>Region 4:</b> l supramarginal gyrus (-63 -27 36)	<b>0.67</b>	<b>0.02</b>	0.57	0.06
<b>Region 5:</b> l superior medial gyrus (-3 63 18)	0.49	0.10	<b>0.61</b>	<b>0.04</b>
 <b>Region 6:</b> r parahippocampal gyrus	0.40	0.17	<b>0.67</b>	<b>0.02</b>

Note: ICC = intraclass correlation; l = left; r = right; MNI coordinates of peak activations are presented between parentheses, Grey = Anatomical Masking;

\*subROI derived from ROI: masked with anatomical ROI from SPM Anatomical Toolbox v1.8.