Age-related differences in task-induced brain activation is not task specific: Multivariate pattern generalization between metacognition, cognition and perception

Esther H.H. Keulers a,*, María Björk Birkisdóttir a, Luciana Falbo a, Anique de Bruin b, Peter L.J. Stiers a

a Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, the Netherlands
b Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, Maastricht University, the Netherlands

ARTICLE INFO

Keywords:
Adolescence
Development
Functional MRI
(meta)cognition
Multivariate pattern classification analysis

ABSTRACT

Adolescence is associated with widespread maturation of brain structures and functional connectivity profiles that shift from local to more distributed and better integrated networks, which are active during a variety of cognitive tasks. Nevertheless, the approach to examine task-induced developmental brain changes is function-specific, leaving the question open whether functional maturation is specific to the particular cognitive demands of the task used, or generalizes across different tasks. In the present study we examine the hypothesis that functional brain maturation is driven by global changes in how the brain handles cognitive demands. Multivariate pattern classification analysis (MVPA) was used to examine whether age discriminative task-induced activation patterns generalize across a wide range of information processing levels. 25 young (13-years old) and 22 old (17-years old) adolescents performed three conceptually different tasks of metacognition, cognition and visual processing. MVPA applied within each task indicated that task-induced brain activation is consistent and reliably different between ages 13 and 17. These age-discriminative activation patterns proved to be common across the different tasks used, despite the differences in cognitive demands and brain structures engaged by each of the three tasks. MVP classifiers trained to detect age-discriminative patterns in brain activation during one task were significantly able to decode age from brain activation maps during execution of other tasks with accuracies between 63 and 75%. The results emphasize that age-specific characteristics of task-induced brain activation have to be understood at the level of brain-wide networks that show maturational changes in their organization and processing efficacy during adolescence.

1. Introduction

Adolescence is an important transitional period between childhood and adulthood, which is characterized by changes in cognitive, socio-emotional and behavioral processes. These changes have been associated with prolonged maturation in brain structure and functioning in this age period, due to a complex interplay between (epi)genetic, environmental, and biological (e.g., hormonal changes) factors (Spear and Silverti, 2016). Adolescence is both a time of opportunities, because of improved complex cognitive functioning, and a time of vulnerabilities indicated by increased risk-taking behavior, incidence of psychiatric disorders, and mortality (e.g., Dahl, 2004; Ernst, 2014; Shulman et al., 2016). Neurobiological maturational processes and specifically diverging developmental trajectories between different brain systems or networks (i.e., motivational and affective versus cognitive control neural systems) are thought to underlie these characteristics of adolescence (Ernst, 2014; Shulman et al., 2016). Neuroimaging studies have indicated that the structure of the cerebral cortex undergoes widespread changes during adolescence (e.g., Giedd et al., 1999; Gogtay et al., 2004; Mills et al., 2016). Grey matter volumes show, after initial increases in early childhood, nonlinear decreases in late childhood and throughout adolescence (Gogtay et al., 2004; Mills et al., 2016). These cortical volume reductions are mainly determined by cortical thinning (Tamnes et al., 2017). In contrast, white matter volumes and the integrity of white matter fibers

* Corresponding author. Maastricht University, Faculty of Psychology and Neuroscience, Department of Neuropsychology and Psychopharmacology, P. O. Box 616, 6200 MD, Maastricht, the Netherlands.
E-mail address: esther.keulers@maastrichtuniversity.nl (E.H.H. Keulers).

https://doi.org/10.1016/j.neuroimage.2018.12.014
Received 9 October 2018; Received in revised form 26 November 2018; Accepted 7 December 2018
Available online 8 December 2018
1053-8119/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
increase with age until mid-to-late adolescence (Giedd et al., 1999; Giorgio et al., 2016; Mills et al., 2016). Underlying changes in synaptic reorganization, pruning, and increased myelination and axonal diameter are thought to explain this increased white matter and decreased grey matter volumes throughout the brain in adolescence (Huttenlocher and Dabholkar, 1997; Paus, 2010; Petanjek et al., 2011).

Widespread maturational changes are also seen in the functional organization of the brain. Functional connectivity is often used as a measure to indicate how strong distant brain areas co-engage in order to accomplish cognitive processes. The functional re-organization during adolescence is described as a shift from local to distributed functional connectivity profiles, mainly driven by increased long-range connections (e.g., Fair et al., 2009; Stevens, 2016; Vogel et al., 2010). Additionally, within networks areas tend to increase their connectivity. Also during this age period the correlations between networks increase, while their hierarchical relationship changes. This maximizes the efficiency of intra-and inter-network communication (Grayson and Fair, 2017; Stevens, 2016). Although the majority of aforementioned research is based on resting state functional connectivity, these networks also shape functional connectivity during a wide variety of cognitive tasks (Cole et al., 2014).

Nevertheless, the main approach to examine developmental changes in task-induced brain activation is function-specific. Developmental fMRI studies tend to describe age-related changes in task-induced activation patterns during the performance of specific cognitive tasks, such as response inhibition, working memory, or reward processing (e.g., Luna et al., 2010; Rubia et al., 2013; van Duijvenvoorde et al., 2015). Such an approach does not respond to the question whether the described task-induced maturational changes are specific to the particular cognitive demands of the task used, or whether similar functional changes would be found if participants performed other cognitive tasks with different demands. The latter might be more likely in the light of the widespread organization of brain areas into different functional networks that are active during multiple tasks (Dosenbach et al., 2007; Duncan, 2010; Stiers et al., 2010), as well as the increased efficiency in and between these networks during adolescence. A recently proposed model by Luna et al. (2015) used this increased integration between different networks to explain cognitive control development. Specifically, brain areas that are functionally connected to multiple networks and that serve different cognitive processes, such as the dorsal anterior cingulate cortex, show prolonged developmental changes (Luna et al., 2015). In line with this explanation, we previously found empirical evidence that functional brain maturation in adolescence is driven by common processes across cognitive tasks as opposed to task specific processes (Keulers et al., 2012). Multivariate pattern classification analysis was able to discriminate between 13, 17 and 21 year-old adolescents based on task-induced activation maps, thus indicating developmental differences in the responsiveness of a wide range of task positive and default mode regions. Moreover, this distributed age-distinctive pattern generalized from a simple go/no-go task to a very different gambling task, and vice versa. A relative limitation of this study was that despite being cognitively and motivationally very different tasks, both tasks were implementations of the traditional stimuli-response mapping paradigm and therefore likely to activate overlapping cognitive processes and brain networks (Keulers et al., 2012).

In the present study we further substantiate the view that functional brain maturation is driven by global changes in the way the brain processes information that go beyond specific cognitive demands and therefore are common across different task designs. The aim was to extend the findings of Keulers et al. (2012) by examining whether the generalization of age discriminating task-induced activation patterns expanded beyond the typical stimulus-response paradigm. Therefore, we selected three tasks associated with different functional brain networks: a meta-cognitive, a cognitive and a visual processing task. The cognitive task was a word pair learning task in which participants encoded series of two independent words, in order to remember them later (encoding phase). Encoding tasks are cognitively effortful, and engage the fronto-parietal and cingulo-opercular task positive networks (Dosenbach et al., 2007; Duncan and Owen, 2000; Stiers et al., 2016). The metacognitive task was integrated in the word pair learning task. Each encoding trial was preceded by a metacognitive trial in which participants had to evaluate and decide whether or not they wanted to be tested later for a recall on the upcoming word pair, based on information regarding points associated with correct recall and difficulty of the word-pair (decision making phase). Metacognition refers to the ability to reflect on your own cognitive processes and is associated with activation of the default mode network (Chua et al., 2009; Well et al., 2013). Lastly, participants performed a passive visual viewing task, in which blocks of pictures of objects, movement scenes and faces were presented without requiring any response. This task mainly activates visual processing areas (Stiers et al., 2006). If developmental differences in task-induced brain activity reflect age-related changes in the efficacy of neural communication between functional networks due to increased long-range connectivity as hypothesized, then the multivariate pattern classification analysis should be able to predict the participants’ age class from their task-induced activation maps regardless of the specific task that was performed to give rise to the activation maps. To address this hypothesis we trained the learning algorithm on examples from one task and tested its classification ability of examples from the other tasks. Successful generalization of age classification across these three very different and independent tasks would confirm that age-related changes in task-induced brain activation are not confined to the specific demands of the task, but are instead a manifestation of a global change in the way the brain processes information. Such knowledge would increase our understanding of typical functional brain maturation during adolescence, which is necessary for studying developmental and adolescence-onset psychiatric disorders. Further, findings related to the generalization between cognitive functions might also have implications for related topics within psychology such as transfer effects of cognitive training, and/or treatments.

2. Materials and method

2.1. Participants and procedure

A total of 47 participants divided into two narrow age groups participated in the study: 25 young adolescents (mean age 13.1 ± 0.42 years; range 12.2–13.8 years; 15 females) and 22 old adolescents (mean age 17.0 ± 0.27 years; range 16.5–17.5 years; 15 females). Due to scanning time limitations and additional exclusion based on excessive head motion, a few participants of this sample were not included in analyses which included the visual task (i.e., for 42 (21 young and 21 old) out of the total 47 adolescents visual task data were included). This unequal samples for the word learning versus the visual task was not considered a problem, because the multivariate pattern classification analysis aimed to demonstrate generalization over participants (i.e., training on some subjects and testing on others). Participants were recruited through an advertisement in a local newspaper. All participants were screened for MRI contra-indications, had normal or corrected-to-normal vision, never repeated or skipped a school grade, had no psychiatric or neurobiological abnormalities and did not use medication that could influence cognitive functioning. Written informed consent was obtained from all participants and their parents. Participants received both travel expenses reimbursement and vouchers with monetary value, after completion of the experiment. The ethical committee of the Faculty of Psychology and Neuroscience of the Maastricht University approved the study. All participants attended one training session and one scanning session. During the training session participants were familiarized with the scanning environment and the tasks. All participants received head motion training in a mock scanner, aimed at teaching the participants how to minimize head motion during scanning. During this training participants viewed a cartoon in the mock scanner twice for 10 min each, with the movie being paused whenever the participants moved their head.
more than 4 mm in any direction. Additionally, during the training ses-
sion participants were trained on the (meta)cognitive word learning task. To
reduce the potential variability in encoding techniques used by the
different participants, all were instructed and trained to use a visualiza-
tion technique for encoding word pairs. Lastly, participants completed
two neuropsychological tests to estimate intelligence. The two age
groups did not differ on verbal intelligence (M_young = 109.9 ± 9.7; M_ol-
d = 110.4 ± 7.2; F(1, 39) = 0.04; p = 0.850) measured with the Pea-
body Picture Vocabulary Test-III-NL (Dunn and Dunn, 2005) or
non-verbal intelligence (M_young = 117.5 ± 8.9, M_old = 119.6 ± 5.7; F(1,
45) = 0.95; p = 0.335) measured with Raven Standard Progressive
Matrices (Raven et al., 1998). To screen for behavioral problems that
could interfere with performance, participants filled in the Youth Self
Report and a parent of the participant filled in the Child Behavior
Checklist (Achenbach and Rescorla, 2001). Scores of all participants on
those total scales were within 1 SD of the mean of a normalized stan-
dardized sample. The age groups did not differ on the total problem scale
on either of the questionnaires (YSR: F(1, 45) = 2.18; p = 0.147; CBCL: F(1,
45) = 1.24; p = 0.273).

2.2. Task paradigms

2.2.1. (Meta-)cognitive word learning task

The (meta)cognitive task was a word-pair memory learning task.
Participants were shown a cue and target word and were instructed that
on the test they had to recall the target word when presented with the
cue. The word-pairs differed in difficulty level in terms of concreteness/
abstractness of the words (2 levels: easy vs. difficult) and in reward level
upon correctly recalling the word pair (2 levels: low, gaining 1–2 points
vs. high, gaining 9–10 points). Subsequently, word pairs can be divided
into four different classes: a) easy word-pair, low reward, b) easy word
pair, high reward, c) difficult word-pair, low reward and d) difficult
word-pair, high reward. Before the presentation of each word-pair, par-
ticipants were given information about its difficulty and reward level,
and had to decide based on this information to either “play” (i.e., earning
points if correctly recalling the target word afterwards) or not to
“play” (i.e., not earning points if correctly recalling the target word after-
wards). The required self-evaluation of future memory performance makes
this a metacognitive judgement (Chua et al., 2009). After this metacognitive
decision phase, the actual word-pair was presented for participants to encode. Blocks of 20 trials of metacognitive decision-making combined with the cognitive encoding were presented, after which a recall test was administered. In this recall test, the cue word
of a word-pair was shown and the participants had to retrieve and
verbally express the associated target word. During the recall, the scan-
ner was turned off. If a participant had earlier decided to “play” on that
word-pair s/he gained the associated reward points if the word was
correctly recalled, but nothing happened to the participants total point
score if s/he made a mistake or had decided not to “play”. At the end of
the experiment the points earned were exchanged for a small amount of
money. This whole procedure of metacognitive decision, cognitive
recalling, and after the word-pairs was repeated for a total of five runs. The word lists used in these five runs were randomized across
participants.

The task had a slow event-related design. The information during the
metacognitive decision phase as well as the word-pair during the
encoding phase was presented for 6s each. Both the transition from a
decision making trial to a corresponding encoding trial, as well as from
an encoding trial to a new decision making trials was interrupted by an
inter stimulus interval (ISI) that pseudo-randomly varied between 6 and
10s. Consequently, the phase onset asynchronicity varied between 12 and
16s and was controlled by scanner triggers. This long interval
ensured that the BOLD signal modulations induced by the neurocognitive
processes in the two phases were only minimally overlapping, while the
jittering in the duration of the intervals further ensured adequate
deconvolution of the remaining overlap of these modulations (Dale,
1999; Ollinger et al., 2001). To further optimize independent statistical
estimation of the BOLD changes in each task phase, during each ISI an
easy visual detection task was presented, in which participants were
asked to detect a 3-D face configuration in an array of seemingly random
moving dots. The face configuration was present in 40% of the stimuli,
whereas in 30% a cylinder was visible and in another 30% a box. The ISI
task forced participants to disengage from encoding the word-pair, by
asking them to engage in another, unrelated task. Each face detection
stimulus lasted 2s, and a range from 1 to 3 of such stimuli were shown
in the ISI of 6–10 s, respectively. During the remaining 4s of the ISI the
fixation cross was shown. Duration per task run was 9.46 min.

2.2.2. Visual processing task

A block design passive visual task was performed after the five runs of
the word-pair memory task. Participants were instructed to view a
stimulus sequence consisting of blocks of different conditions; a) black
and white pictures of faces b) coloured pictures of static objects, c) col-
oured movies of natural moving scenes and d) a fixation point that
changed its shape to increase ease of attention engagement. No response
was required. Each block lasted 24s and consisted of 12 items of a specific
condition (i.e., faces, objects or movement) that were presented for 2s
each. During the control fixation condition, the screen was blank with a
central red fixation spot that changed shape between a larger circle,
smaller circle, a left tilted and a right tilted ellipse every 2s. The four
conditions were all repeated four times in the same order. Starting and
ending the sequence with a black screen for 2s brought the total duration
of the experiment to 6.47 min. This task was originally developed to map
a large number of visual areas during passive viewing (Stiers et al.,
2006).

2.3. Behavioral analyses

Behavioral analyses were conducted using the statistical package
SPSS 24.0. Analyses were conducted to examine age differences in the
decision-making to “play” or not in the metacognitive phase, in correct
recall and in reaction times. Due to non-normal distributions of the
number of times participants decided to “play” and the number of
correctly recalled items, the analyses were performed using the non-
parametric Mann-Whitney U test, with the alpha level set at 0.05. For
the analysis of normally distributed reaction times, a one-way ANOVA
test was use with the alpha level set at 0.05.

2.4. Functional magnetic resonance imaging data

2.4.1. Image acquisition and preprocessing

Data collection was performed on a Siemens MAGNETOM Allegra 3 T
MRI head-only scanner. Head motion was minimized by the use of foam
padding. A TI-weighted anatomical scan was acquired for normalization
purposes for each participant (TR = 2250 ms, TE = 2.6 ms, flip
angle = 9°, FOV = 256 mm, slice thickness = 1 mm, matrix
size = 256 × 256, number of slices = 192). The voxel size of the
anatomical scan was 1×1×1 mm. A total number of 32 axial slices
covering the whole brain including the cerebellum were imaged by using
a T2*-weighted gradient echo planner pulse sequence (TR = 2000 ms,
TE = 30 ms, FA = 90, FOV = 224, slice thickness = 4 mm, matrix
size = 64 × 64, flip angle = 90°). Voxel size was 3.5 × 3.5 × 4 mm. Slice
scanning order was ascending interleaved.

Data was preprocessed using SPM version 5 (Welcome Department of
Cognitive Neurology, London, UK) running on Matlab, version R2013b
(The MathWorks Inc., Natick, Massachusetts). First, images were cor-
correct for slice scanning time differences, followed by rigid realignment
of images to the first scan of participants and further co-registering to the
structural image of the participants. Segmentation was used to create a
spatial normalization file that was further used to normalize the func-
tional images. Finally images were spatially smoothed using a 4 mm
Gaussian kernel.
2.4.2. Head motion handling

In fMRI studies using between subject designs, group differences in the amount of head motion of participants can seriously affect group differences in the BOLD signal. This is of high importance in developmental studies as younger participants tend to move more while in the scanner than older ones (Keulers et al., 2011, 2012). In functional connectivity studies it was shown that at least part of the reported stronger short range and weaker long range connectivity in younger compared to older adolescents, can be attributed to these head motion related biases (Van Dijk et al., 2012; Power et al., 2012). In order to observe whether there is a genuine activation differences between age groups it is therefore important to correct for head motions.

In order to control for the confounding effect of head motion we applied three corrections to our data (Keulers et al., 2011, 2012). First, participants were excluded in case of excessive head movement, defined as a run-start-to-end head translation in the z direction of >3 mm and/or a head rotation around the x axis of >2.33° in any of the five (word learning task runs and/or the visual task. These are the two realignment parameters most affected by head movement (Keulers et al., 2011; Moyer et al., 2007; Yoo et al., 2005). For the word learning task, five participants in the young adolescents group and two in the old adolescents group had to be discarded, leading to the described final sample of 47 participants. For the visual task, another four young and two old adolescents were excluded because of excessive head motion. The excluded participants were not younger than the included participants in the young age group (t (30) = 1.8, 1-tailed p = 0.957) or in the old group (t (22) = 0.2, 1-tailed p = 0.403). Excluded participants also did not score lower on measures of intelligence or behavioral problems (lowest 1-tailed p value: 0.125). Secondly, due to heterogeneity in magnetic field strength, head position changes cause fluctuations in the measured signal intensity over time (Friston et al., 1996). To remove these fluctuations from the time series data, we added the six affine parameters estimated during the realignment step of preprocessing as events of no interest to the design matrix (Friston et al., 1996). Lastly, we identified individual volumes during the acquisition of which the head position changed more than 1:10th of the slice thickness (i.e., 0.4 mm). These fast movements relative to the gradient setting during acquisition distort the signal measured. Task trials whose induced BOLD response overlapped with such volumes (taking into account a hemodynamic response delay window of 8 s) were modelled as separate events of no interest. The number of fast motion contaminated volumes identified in the younger group was 3.9% and in the older group 4.4%. This procedure confined BOLD-informed inferences to those trials during which participants’ head position was sufficiently stable. For usable trials, there was no significant age-related differences in the absolute intra-volume z-translation movement (t(45) = 0.7; p 1-tailed = 0.246), or the intra-volume x-rotation (t(45) = 0.1; p 1-tailed = 0.464). This third motion correction could not be applied to the visual task, as this task used a block design in which it is not possible to separate model individual trials contaminated with head motion as events of no interest.

2.4.3. First level imaging analyses

A first level, or within participant, analysis was conducted on the preprocessed images of individual participants using a traditional univariate general linear model (GLM) in SPM5. The results of this analysis constitute the starting point for both the group-wise or second level univariate analysis and for the multivariate pattern classification analysis. A grey matter mask was set up to model all task conditions and the metacognition related parameters. Within the word learning task four conditions could be distinguished based on the prior word-pair information, i.e., difficulty level (easy vs. hard) and reward level (low vs. high). That resulted in four metacognitive decision making regressors (i.e., trial phase when making the choice to “play” or “not” and four cognitive encoding regressors (i.e., trial phase when encoding the word-pair). All decision making and encoding regressors were modelled as separate events of interest marked by their onset time (but without a duration in time), and convolved with a canonical hemodynamic response function combined with time and dispersion derivatives. The attention capturing face perception movies that were presented in the interstimulus period were also included in the design matrix as a separate regressor, in which each two second stimulus was represented by its onset time and convolved with the theoretical hemodynamic response function (HRF). Additionally, the six motion parameters as well as trials contaminated with fast head movement were modelled as events of no interest. For subsequent second level group analysis and multivariate pattern classification analysis individual percent signal change maps (PSC maps) were generated for each of the eight regressors of interest, i.e., one for each of the four subconditions within the metacognitive decision making and the cognitive encoding phase (easy, low reward; easy, high reward; difficult, low reward; difficult, high reward). PSC maps show intensity changes in a given voxel as a percentage of the baseline signal intensity and are computed from the regressor weights estimated in the univariate GLM analysis described above (Maziak, 2009). The design matrix for the visual task consisted of three conditions of interest (faces, objects and movement scenes), marked by their onset and a duration of 24 s, convolved by the HRF only and the six head motion parameters. For subsequent group analysis and multivariate pattern classification analysis PSC maps were generated for each of the visual conditions (faces, objects, and movements).

2.4.4. Group-level imaging analyses

2.4.4.1. Univariate analysis. A whole brain, voxel-wise general linear model with task condition as the within-subjects factor and two age groups as the between-subjects factor was performed separately for the word learning task and the visual task. The 8 PSC maps of the word learning task were used in a 2 (metacognitive decision making vs cognitive encoding phase) x 2 (difficulty level) x 2 (reward level) design. The three PSC maps of the visual task were used in a one-way factor (task condition) design. In both GLM models, gender, verbal intelligence and non-verbal intelligence estimates were added as covariates of no interest. First, task specific (i.e., metacognitive decision making, cognitive encoding and visual processing) activation patterns were identified using a contrast that pooled over task conditions within each task and age groups. These task-induced activation maps were compared in order to indicate differences between and the independence of the different task paradigms used. Second, age differences were examined per task in order to make comparisons possible between the traditional univariate approach and the age-discriminative patterns resulting from the multivariate pattern classification analyses. The statistical t-maps were evaluated at the voxel-wise significance level of 0.05, corrected for multiple comparisons based on the Gaussian field theory (FWE<0.05). In addition, only clusters of 5 or more voxels were reported.

2.4.4.2. Multivariate pattern classification between age groups. A multivariate pattern classification analysis (MVPA) was performed on the PSC maps created in the first level analyses, using the two age classes as the classification criterion. For every task we included PSC maps for every task condition (Keulers et al., 2012; Stiers et al., 2016). This resulted in four PSC maps per participant for the metacognitive decision making task (4 x 47 = 188 PSC maps), four PSC maps per participant for the cognitive encoding task (4 x 47 = 188 PSC maps) and three PSC maps per participant for the visual task (3 x 42 = 126 maps).

A grey matter mask was created for the voxels to be included in the analyses. First, to make sure that the same grey matter voxels were selected in all three tasks, the grey matter density prior associated with the MNI 151 T1 template was thresholded by applying a 0.33 density high-pass filter, with the resulting image being resliced to the voxel grid of the PSC maps. Second, to avoid mistaking structural maturation changes for functional maturation, all grey matter voxels were excluded from the created grey matter mask that showed evidence of structural...
The other participants were used to train the classifier. These individuals were used as test data set, while all PSC maps from all participants that participant belonged to. The Spider implementation of the support vector machine-learning algorithm (www.kyb.mpg.de/bs/people/spider/main.html) was used as the classifier for the MVPA. A linear kernel was used and classes were assumed unbalanced.

The MVPA was conducted in several steps. First, the data were rescaled within each voxel to a range of 0–1, to focus the analyses on the relative variations in values within each voxel and to avoid that voxels with large absolute values would have more importance in the end result. Second, independent test examples were created by putting all the PSC maps from one young and from one old participant aside. The PSC maps of these individuals were used as test data set, while all PSC maps from all the other participants were used to train the classifier to discriminate the age of the participant. To make the end result independent of the particular participants chosen as test examples, a cross-validation was applied in which data from each participant of the young and old adolescent group were set aside once as test example. Each resulting training set is referred to as one fold of the data set. Third, to avoid overfitting the training data, the training set at each fold was split into a number of smaller subsets and each subset was left out once from the training - i.e., at each fold the classification was repeated for as many times as there were subsets to leave out. The default number of splits on the training set was 15, but the influence of smaller or larger numbers of splits was systematically examined. Split subsets were not used for testing. For each split the trained classifier was tested on the test examples and the average accuracy obtained over all the splits of the training set is reported. Similar, the final weight of the voxels in the analysis was the average weight of each voxel over the split analyses. Lastly, because there was no prior region of interest that could limit the feature space for classification, and because MVPA results are sensitive to the number of informative features relative to the informative features in a data set, a recursive feature elimination (RFE) strategy was applied (De Martino et al., 2008). This means that the classification analysis was repeated 50 times, and each time the 10% least contributing voxels (i.e., with their weights closest to zero) in the previous step were eliminated in the subsequent analysis. Each subsequent iteration thus started with a smaller number of features and yielded a new accuracy value when the classifier trained on these features was tested on the test examples. This recursive feature elimination process was nested within cross-validation folds on the data set, as described above. Consequently, for each fold (each set of data set apart for testing) the recursive feature elimination was applied. To obtain one final result, the results were first averaged over the folds, yielding an average accuracy and voxel weights at each iteration of the RFE process, and then the result at the iteration with the highest classification accuracy was reported as the final result.

2.4.4.3. Control analyses. In order to investigate the dependency of the results on the choice of MVPA parameters, the principle MVP analyses described above were repeated with different parameter settings. First, the effect of initial feature selection was examined. The analyses described above started with all the voxels in the analysis mask. However, an initial selection of features based on univariate data analysis is often performed in an attempt to exclude noise voxels. The initial feature reduction step comprised of selecting the 25 k, 10 k or 5 k voxels that discriminated most between the age groups. This selection was based on the unsigned t statistic computed for each feature separately (De Martino et al., 2008). Importantly, this selection step only included trials belonging to the training set of the particular fold. Second, we varied the number of splits applied to each fold of the training set, from 15 in the principle analysis to 3, 5, and 25 splits. Third, the role of voxels actively engaged by the tasks on the classification results was investigated. On the one hand, the analyses were repeated while including only voxels that were significantly activated by the task in the univariate group analysis at the significance level of FWE<0.05, and on the other hand by including only voxels that were not significantly activated by the task (i.e., p < 0.001 uncorrected for multiple comparisons).

2.4.4.4. Multivariate pattern generalization across different task paradigms. In order to answer the critical investigation whether an age-discriminative brain activation pattern was specific to a particular task or general to task-induced activations across multiple tasks paradigms, the classification algorithm was trained on PSC maps from one task and subsequently tested on PSC maps from another task performed by individuals not part of the training data set - i.e., we asked how accurately the trained classifier could decode age from brain activation maps obtained during the other tasks. The same parameters as described above for the principle MVP analyses (paragraph 2.4.4.2) were used in the current analyses. In addition, similar control analyses were conducted for these generalization analyses, to establish the dependency of the classification results on the choice of MVPA parameters.

2.4.4.5. Statistical significance of classification accuracies. Statistical significance was estimated based on 125 re-analyses of the data following the exact same steps as described above (initial reduction, RFE and best iteration selection), but with random redistributions of the training labels. The random reassignment of labels (young/old adolescent group) mimics the null-hypothesis that there is no systematic association between feature values and age classes, so that the class labels are interchangeable. Random accuracies were on average higher than 50.0% due to selecting the best of all iteration in each of the 125 randomization analyses. The significance level was 0.05, one-tailed. The significance was determined by the centile position of the observed accuracy of the original analysis amongst the 125 accuracies obtained with the repeated randomization procedure.

2.4.4.6. Classification weight maps. Classification weights numerically express the importance of each voxel to the classification. At each fold and iteration, the classification weight is the average weight across the splits on the training data. For the purpose of visualizing the distribution of age-discriminative voxels in Fig. 6 we created for each of the three task-condition data sets a classification weight map at the 30th iteration of the recursive feature elimination procedure. At this iteration only 8671 of the 184165 starting voxels are still included. The 30th iteration was chosen because averaged over all task conditions, this iteration yielded the highest within task classification accuracy (60.9 ± 6.8%). The weight maps were created by simply adding the voxel weights at each fold of the analysis. MVP Parameters were as in the basic leave-one-participant-out analyses. The resulting classification weights reflect both the importance of the voxel for classifying at a specific fold and the number of folds in which the voxel contributed to the classification at that stage in the analysis. An additional weight map was created for the first iteration of training on the metacognitive condition, i.e., with 184165 voxels (Fig. 6-C).

3. Results

3.1. Univariate GLM analysis results

3.1.1. Behavioral performance

Analyses of the behavioral data of the word learning task showed that
the two age groups did not differ significantly in their task performance. There was no difference in the number of times the young adolescents decided to “play” for rewards (75.3 ± 21.1% of trials) versus the old adolescents (79.0 ± 22.7% of trials; Mann-Whitney U = 243; p = 0.493). There was also no difference in the average time required to make the metacognitive decision making in the young (1619.0 ± 330.0msec) and the old adolescents (1648.5 ± 412.4msec; F(1, 45) = 0.074; p = 0.787). The number of world-pairs correctly recalled was 73.8 ± 17.4% in the young and 81.4 ± 15.7% in the old adolescents, which was not statistically different (Mann-Whitney U = 199; p = 0.107). The reaction times for correctly recalling a word was 3379.5 ± 560.4msec in young and 3367.7 ± 719.1msec in old adolescents and also did not differ significantly (F(1, 45) = 0.001; p = 0.969). Given that the age groups did not differ in task performance, it can be assumed that any BOLD difference observed between age groups truly reflects differences caused by maturation, rather than caused by differences in task performance.

As the visual task was a passive task, there were no behavioral measures to compare between age groups.

3.1.2. Univariate task effects

A voxel-wise GLM analysis was performed to investigate to what extent the 3 tasks differed in their activation of the brain. The results of these analyses are visually presented in Figs. 1 and 2. Fig. 1- A shows brain activation during the metacognitive decision making and cognitive encoding task phase, respectively, compared to baseline. Both tasks tend to activate to various degrees the typical nodes of the frontoparietal and cingulo-opercular task positive networks (Dosenbach et al., 2007; Duncan and Owen, 2000; Stiers et al., 2016), in the intraparietal sulcus, on the precentral gyrus and on the medial part of the superior frontal gyrus. For both tasks activity was stronger in the left hemisphere (presented in Fig. 1A), in accordance with the use of verbal stimuli in this task paradigm. Both tasks also induced strong activity in posterior visual-processing cortical areas. Despite this general similarity in activation pattern (visualized in purple in Fig. 1A), several task-specific activations were also present during either the decision making (visualized in red) or encoding (visualized in blue) phase of the task compared to baseline.

Fig. 1. Task-induced activation during the metacognitive decision making and cognitive encoding task phases in the left hemisphere, pooled across age groups. (A) Activity in each task phase relative to baseline (metacognitive decision making activity in red; cognitive encoding activity in blue). The overlap in activation patterns between the two task phases is shown in purple. (B) Areas that are significantly more active during the metacognitive decision making compared with the cognitive encoding phase (cyan) and the other way around (green). All results are visualized at the family-wise error corrected level (alpha = 0.05, critical t-value = 4.98).

Fig. 2. Task-induced activation during the visual task is presented in red. Task activation during the metacognitive (black) and cognitive (light grey) phase of the word learning task are presented as a reference frame to compare task activations. All results are visualized at the family-wise error corrected level (alpha = 0.05, critical t-value = 4.74 for visual task, critical t-value = 4.98 for metacognitive decision making and cognitive encoding task).

To test more objectively for task-unique activation we directly contrasted the activation differences during the two task phases. Fig. 1B shows voxels that were significantly more active during the decision making phase compared to the encoding phase (visualized in cyan). These voxels formed smaller clusters located bilaterally in the occipital pole and lower intraparietal cortex, as well as in the right inferior parietal lobule (supramarginal gyrus). In addition, two larger bilateral clusters of voxels with significantly higher activation during the decision making phase were observed in the anterior and posterior cingulate cortex region, although these voxels did not reach significance in the analyses of either of the two task-phases alone. These cortical regions are central nodes of the default mode network, which is systematically de-activated during task execution. Therefore, these “activation” clusters were significantly less deactivated during the decision making compared to the encoding phase. The voxels responding stronger during encoding than during decision making are visualized in green in Fig. 1B. The most conspicuous cluster was in the left inferior frontal gyrus, extending from the ventral part of the precentral sulcus until the (posterior) lateral orbital frontal cortex. Only a few small clusters were observed in the same regions of the right hemisphere. In addition, smaller clusters of
voxels with more activation during encoding than decision making were also present posterior in the lateral superior frontal gyrus (frontal eye fields) and in the medial superior frontal gyrus (preSMA), in the cuneus extending into the cortex around the ictus and lastly in the cortex around the left hemisphere MT+ region extending in the most anterior aspect of the fusiform gyrus.

Fig. 2 shows brain activation during the visual processing task compared to baseline, with the regions activated by the metacognitive decision making and cognitive encoding task phases indicated as outlines for the sake of comparison.

The observed differences between the three tasks are in line with existing literature, which state that metacognitive processes are associated with the default network, while cognitive encoding task performance causes more activation in the frontoparietal and/or the cingulo- opercular task positive networks. Visual processing however, is known to be mainly associated with activation in posterior areas such as the precuneus, occipital-temporal, and parietal-occipital cortices (e.g., Buckner et al., 2008; Chua et al., 2009; Dosenbach et al., 2007; Sridharan et al., 2008; Stiers et al., 2006). These differences provide validation that the three used tasks in the current study rely (partly) on different subnetworks in the brain. The three task paradigms are therefore not only cognitively very different but also differ at the neurobiological level.

3.1.3. Univariate age group differences

In order to examine whether univariate analyses would reveal significant age effects we contrasted the PSC maps between the two adolescent age groups. For the metacognitive decision making and the cognitive encoding task, no voxels with significant age effects were observed at the FWE corrected significance level when looking at each task separately. When the decision making and encoding data were pooled, 13 very small loci of age related differences were found, distributed throughout the brain, mostly comprising only one or two voxels (See Supplementary Information, Fig. S1). The largest cluster, comprising 27 voxels, was located on the rostral bank of the left ventral precentral sulcus (MNI [-50, 6, 18]), in the inferior frontal gyrus region that was strongly activated by the encoding task. A cluster of 10 voxels was located somewhat higher on the caudal bank of the precentral sulcus (MNI [-60, 10, 36]), just outside a region activated by both tasks (presumably premotor cortex). Lastly a cluster of 9 voxels was located in the lateral bank of the dorsal intraparietal sulcus (MNI [−40, −54, 52]), also near the boundary of a cluster of voxels activated by both tasks. No clusters of brain voxels had a stronger BOLD response in the young group compared to the older group. We conclude therefore that subtle age related differences in BOLD response strength might be present at the univariate voxel level, but almost all are too weak to reach statistical significance and their wide distribution bars a straightforward interpretation. For the visual task, no significant age differences were found at the FWE corrected level (See Supplementary Information, Fig. S1).

3.2. Multivariate age effects within and between different task paradigms

3.2.1. Classifying age from metacognitive decision making and cognitive encoding task-induced activation maps

To demonstrate the feasibility of decoding age from task-induced brain activation maps, and to have a reference point for the level of accuracy that can be obtained in classifying age, we first applied the MVP analysis within the two word learning phases separately. The age classification was significantly above chance level for the metacognitive decision making maps (70.4%, p < 0.01) and the cognitive encoding activation maps (68.0%, p < 0.05). This accuracy peak was reached at the 29th iteration for the encoding activation maps, with 9635 of the 184165 starting voxels still in the running, while the highest accuracy for the decision making activation maps was reached in the 52nd iteration with 850 voxels.

The variation of accuracy for both the metacognitive decision making and cognitive encoding task with the progression of the recursive feature elimination procedure is illustrated in Fig. 3-A. Relatively high accuracy is obtained at all iterations, suggesting that the classification is not dependent on a specific and confined subset of the voxels. The results were also not critically dependent on other analysis parameters. As can be seen in Fig. 3-B, increasing or decreasing the number of splits (i.e., related to the chance of overfitting the training data) away from the 15 splits used in the basic analysis reported above had only a marginal effect on the accuracies obtained, with somewhat lower but still significant accuracies when only three splits were used (metacognitive decision making: 66.2%; p < 0.05; cognitive encoding: 66.3%; p < 0.05). Although the relative independency of accuracies on the number of voxels in the running (see Fig. 3-A) suggests that a wide range of voxels contain relevant information, we nonetheless investigated the effect of an initial feature reduction to 25, 10 or 5 thousand voxels. As shown in Fig. 3-C, this procedure lowered somewhat the accuracy in the metacognitive decision making task (lowest accuracy was for 25 thousand starting voxels: 62.7% p<0.05), but slightly boosted the classification accuracy in the cognitive encoding task (highest accuracy with 10 thousand starting voxels: 72.0% p<0.001). Lastly, we explored the importance for successful classification of voxels that were activated by the task (Fig. 3-D). First, the analysis was repeated including only voxels that were significantly activated by the specific task as shown by univariate group analysis. As can be seen in Fig. 3-D, this reduced somewhat classifications accuracies, more so for the metacognitive decision making phase (63.7%; p<0.055) than for the cognitive encoding phase (65.7%; p<0.05). Second, the analysis was repeated excluding all task activated voxels. Fig. 3-D shows that this had almost no effect on the classification: for decision making maps the accuracy slightly lowered from 70.4% to 67.4% (p<0.001), but for the encoding activity maps it slightly increased, from 68.0% to 69.6% (p<0.001). These results confirm that it is not per se the brain regions that are directly induced by the task that reflect age-specific brain activity. Instead the classification seems substantially driven by the voxels outside of these activated regions.

3.2.2. Generalization of age classification across metacognitive decision making and cognitive encoding task paradigms

The critical question addressed in this paper is whether the described age-related brain activation differences are specific for the task requirements and cognitive processes participants engage into. This generalizability of the age discriminative pattern between different tasks was examined by training the classifier on examples from one task and testing its classification ability of examples from another task. This classification was significantly above chance level in both directions. While the accuracy was slightly lower compared to within task age decoding, for training on cognitive encoding examples and testing on metacognitive decision making examples (66.8%; p<0.01), it increased with 5% for training on decision making examples and testing on encoding examples (75.1%, p<0.01). The latter peak accuracy was obtained after 96 iterations with only four voxels still in the running. However, as Fig. 4-A shows, the accuracies were high over a wide range of iterations. In fact, the lowest accuracy obtained over all iterations was 62.3%, and accuracies exceeded 70% from iteration 51 (with 945 voxels) to iteration 58 (with 450 voxels). For training on encoding examples and testing on decision making examples, the highest accuracy was obtained at the 30th iteration. Fig. 4-A shows that around the 50th iteration a break point is reached after which classification became more and more difficult. In the first 50 iterations, however, accuracy was well over 60%, with the highest accuracies being reached at the 50th iteration with 945 voxels. The independency of the classification accuracies from the initial number of voxels and from the number of splits on the training set are illustrated in Fig. 4-B and Fig. 4-C.

The best accuracies reported above were based on the iteration that yielded the best classification of test examples. However, since the generalization from one type of example to another is investigated, an alternative approach is to use the results from the iteration in which the classifier is optimized for classifying the training examples. For instance,
when training on metacognitive decision making and testing on cognitive encoding examples, the iteration that yielded the highest accuracy for classifying the decision making examples from the participant that was left out of the training was selected and the corresponding voxel configuration and weights were used to test the encoding examples of the same left out participant. This yielded very similar results, as can be seen in Fig. 4-A. For generalization from cognition to metacognition, the best accuracy dropped from 66.8% at iteration 30 to 66.4% at iteration 29 (p < 0.01), when switching from the best iteration to classify metacognitive decision making examples to the best for classifying the cognitive encoding examples (i.e., the category on which the algorithm was trained) (Fig. 4-A, thin grey arrow). For the reverse generalization,
from metacognition to cognition, the accuracy for classifying cognitive encoding maps at the best iteration for classifying metacognitive decision making maps, as used in training, was 71.5% (p < 0.01), at iteration 52 (see thin black arrow in Fig. 4-A). While this is lower than the highest accuracy obtained, which was 75.1% at iteration 96 with only four voxels (thick arrow in Fig. 4-A), it is very close to the accuracy peak that was evident at around the 50th iteration – i.e., 72.4% at iteration 53. We can conclude from this that the optimal classifier for decoding age from one type of task-induced activation map is also (near) optimal for classifying age from another type of task-induced activation maps.

To investigate the role of voxels actively engaged by the task in the classification results, MVPA was repeated with either only task engaged voxels included or with task engaged voxels excluded as features. The effects of these follow up analyses are summarized in Fig. 4-D. In general, classification results suffered much more from restricting the features to task-activated voxels, regardless of the direction of generalization (i.e., from metacognition decision making to cognitive encoding or the reverse) and regardless of whether the restriction was based on the activation maps of the training task, the test task or both. In contrast, restricting the voxels to non-task-activated voxels had only small effects on the generalization accuracies.

3.2.3. Generalization of age classification from (meta)cognitive to visual domain

The brain activity maps acquired during the passive visual task were used to further broaden the scope of our investigation into the task-generality of age specific brain activation patterns. The classifier was trained on activation maps derived from either the metacognitive decision making or the cognitive encoding task phase, and subsequently tested on the passive viewing PSC maps of participants not included in the training set. The age class of the participants was predicted with 64.8% accuracy (p < 0.05) when the classifier was trained on decision making data and with 63.6% accuracy (p < 0.05) when trained on encoding data. At the best iteration for age-classification of the training examples, the accuracies for classifying the passive viewing examples dropped to 60.5% (p = 0.066) and 56.5% (p > 0.1), respectively. Again, control analyses were performed to examine the dependency of generalization accuracies on parameter settings. Generalization of age discriminative patterns from metacognition decision making to visual processing remained significantly accurate for all split numbers, although generalization accuracies from encoding to visual processing dropped below 60% (p > 0.1) for the lowest and the highest split numbers (see Fig. 5-C). Similarly, while generalization accuracies from metacognition decision making to visual processing were not affected by the initial number of features included in the analyses, the generalization accuracies from cognition to visual processing dropped below 60% (p > 0.1) when the initial number of features is restricted (see Fig. 5-B).

The reverse generalization, from passive viewing PSC maps to either the cognitive encoding or the metacognitive decision making task phase, yielded comparable results. The age class of the participants was predicted with 65.2% accuracy (p < 0.05) when the test examples was decision making data and with 61.0% accuracy (p < 0.066) when encoding data was used for testing. When the training and testing was done within the passive viewing data, the accuracy rose to 73.0% (p < 0.01).

To further explore which voxels were relevant for the age generalization from encoding to visual tasks a series of follow-up analyses were performed in which univariate task-induced activation maps were used to include or mask voxels for the multivariate pattern analyses. The results are summarized in Fig. 5-D. The pattern that emerged from this exploration differed in an important respect from
the generalizations between metacognitive decision making and cognitive encoding discussed earlier. In these earlier across tasks generalizations we found that non task-engaged voxels importantly contributed to age classification and were necessary to retain the significant classification accuracies obtained when including all voxels. In contrast, limiting the multivariate space to task-engaged voxels (based on either train or test tasks) reduced age classification accuracies. Compared to this, when generalizing age-discriminative patterns from metacognition decision making to visual processing maps, the voxels engaged during the visual processing task played a crucial role in the classification. Excluding these visual processing voxels from the feature space reduced the classification of age, whereas confining the feature space to these visual processing voxels tended to preserve or even boost the age classification accuracies (Fig. 5-D, dark grey bars on the right). In contrast to this, including or excluding voxels that were active during the metacognitive decision making task had no substantial effect on age classification accuracies (Fig. 5-D, dark grey bars in the middle). A somewhat different pattern was observed in generalization of age-specific activity patterns from the cognitive encoding to the visual processing task. Here both the voxels that were engaged by the visual task and those that were non-engaged by the visual task were important, since classification accuracies dropped when either of them were excluded as analysis features (Fig. 5-D, light grey bars on the right). In contrast, the voxels not engaged by the training task (encoding task) were crucial for age decoding from passive visual PSC maps, because including only these voxels preserved the age generalization to visual examples, whereas excluding them reduced the age classification accuracy (Fig. 5-D, light grey bars in the middle).

3.2.4. Spatial pattern of brain structures contributing to age group classification

The classification weight maps presented in Fig. 6-A show the spatial pattern by which age groups differed for each task condition at the 30th iteration (retaining 8671 voxels) of the recursive elimination procedure. Voxel clusters contributing to age group classification were distributed throughout the whole brain, rather than being confined to some select regions. While some clusters overlapped with areas activated by the tasks and others overlap with task de-activated regions, many are located outside of task-related brain structures. Secondly, while there is some overlap in the contributing voxels from the three tasks, the overlap is relatively small. This doesn’t mean that there is no shared pattern of brain maturation across the three data categories. The overlap is low because there is a large number of voxels carrying age specific information, as was evident from the range of iterations in the recursive elimination process that yielded above chance classification. Consequently, each task-specific analysis can pick its most optimal subset from a large pool of informative voxels. While this subtest of voxels is somewhat less optimal for classifying data from the other task conditions, it does carry enough age-specific information to allow above chance classification of the other data sets. To illustrate this, Fig. 6-B and 6-C compare the set of voxels at the 30th iteration of training on the visual data to the much larger set of voxel with discriminative potential (i.e. weights exceeding $>0.5$) during the initial training (iteration 1) on the metacognitive data. This comparison shows that most of the voxel clusters selected in the visual data analysis at iteration 30 overlap with or are close to age discriminative voxels in the metacognitive data analysis (Fig. 6-B). Moreover, the average age-related difference in percent signal change in these ‘visual’ voxel clusters is in the same direction during execution of the visual task and the metacognitive task, and in line with the valences of the weights given to these voxels in both multivariate analyses (e.g., positive weights signify higher BOLD signal in older than younger participants, etc.). It should be noted, however, that there are also visual age discriminative clusters located in isolated (e.g., Fig. 6-C, clusters 5 and 6), and even clusters where the visual and metacognitive weight valences are opposite (Fig. 6-C, clusters 7 and 8).

4. Discussion

The present study aimed to examine whether age discriminative brain activation patterns can be generalized between metacognitive, cognitive and visual processes. A previous study showed that developmental differences between age groups in brain activation during task execution are general across different cognitive functions measured by stimulus-response paradigms (Keulers et al., 2012). The new angle in the present study was to test this generalizability of functional brain development across a wider range of information processing levels, from abstract metacognition over intermediate level cognition to low level visual processing.

The multivariate pattern classification applied within each of the tasks separately confirmed that brain activation during the execution of a specific task is consistent and reliably different between the ages of 13 and 17 years. The accuracy with which the trained classifier differentiated between the age classes of 13 from 17 years based on activation maps obtained during the performance of metacognitive decision making, cognitive encoding and visual tasks is in the same range (i.e., 68-70%) as a previous study using different cognitive tasks (Keulers et al., 2012). Moreover, our results confirm previous studies that showed the feasibility of multivariate pattern classification to decode age-classes from developmental imaging data in normal and pathological samples (Cohen et al., 2010; Hart et al., 2014; Keulers et al., 2012). The current results additionally show that age-typical characteristics of brain activation are not specific to the particular task performed, but instead are common to brain activation induced by a range of different tasks. This result confirms the previous finding by Keulers et al. (2012), who showed that age-discriminative activation patterns are comparable between a simple go/nogo task and a complex and engaging gambling task. While the two tasks used by Keulers et al. (2012) adhered to the typical stimulus-response mapping paradigms of cognitive tasks, the current results extend these findings to a broader set of tasks. Not only are the specific cognitive demands in the tasks used here very different, but voxel-wise univariate analyses showed that the brain areas engaged by each of the three tasks were considerably different. Despite this variation, multivariate pattern classifiers trained to detect age-discriminative patterns in brain activation from one task were significantly able to decode age classes from brain activation maps during execution of other tasks. This generalization was demonstrated for the metacognitive decision making task to the cognitive encoding task, from the cognitive encoding task to the metacognitive decision making task, and from each of these to the passive viewing task.

The generalizability of age-discriminative patterns might seem to suggest a common process across tasks as the driving force. A possible shared component might be the ability to engage in tasks in a very general sense: participants have to focus attention on a subpart of the available information with a specific goal or behavioral outcome in mind, arrived at according to some rules. All three tasks require the engagement at least in part of the task-positive networks (e.g., Dosenbach et al., 2007; Duncan, 2010; Sridharan et al., 2008). However, the maturational differences found are not confined to the task positive networks: excluding the activated regions of this network from the analyses did not affect the classification and generalization accuracies. On the contrary, age classification within tasks as well as the generalizability across tasks seems more driven by voxels that are not activated during task execution than by task activated voxels, since the classification accuracies tended to decrease when only task-activated voxels were used in training the classifier. A second indication that the generalization of age-classification across tasks is not driven by the degree of overlap in neurobiological processes between tasks is the less efficient generalization between cognitive encoding and passive viewing compared to generalization between metacognitive decision making and passive viewing, both in terms of the accuracies obtained and of resilience against manipulations of the MVPA parameters. This difference may be surprising given that the cognitive encoding task, in which participants were trained to use a
visual memorization technique, showed more overlap in task activation with the visual task than the metacognitive decision making task (Fig. 2).

That task-induced developmental brain differences are more complex than a maturational change in the task-activated networks is further indicated by the different role that visually active voxels played in the across tasks generalization. While the task-activated voxels were not critical for the generalization between cognitive encoding and metacognitive decision making data, the active voxels during the passive viewing task were essential to generalization success. Apparently, while there is enough age-characteristic brain activation available in visually active voxels, the voxels not specifically engaged in processing the visual stimuli contain too little age-specific activation, or at least, it does not overlap with such activation during the (meta)cognitive tasks. This could have to do with the fact that passive viewing is much less engaging and generally more similar to a state of rest – i.e., the parts of the brain that are not committed to the visual processing can engage in neurobiological processes that are more similar to those in a state of rest. The activity pattern associated with this alternative state would then be too dissimilar, particularly in the frontal brain, to allow generalization from these non-visually committed features to (meta)cognitive activation patterns. A further hint in that direction is that the encoding task is generally experienced by the participants as more cognitively engaging than the metacognitive decision making task. In terms of switching between complementary networks, this increased effort would require stronger suppression of resting state processes and would explain why the generalization between encoding and passive viewing is more difficult than between the less engaging resting metacognitive decision making and the passive viewing activity maps. These considerations could be brought in alignment with the recent finding that the whole brain functional connectivity during execution of an emotional go/nogo task shifts to a somewhat younger age pattern (<1 yr) when participants in the age range 12–16 yrs (but not younger or older participants) performed the same task for reward (Rudolph et al., 2017). The fact that this added reward induced a change towards a younger “neutral” (i.e., not rewarded) connectivity pattern suggests that the reward increased the participants’ motivation and therefore induced stronger cognitive effort accompanied by stronger default mode network inhibition. Future studies will have to establish whether such an interaction between the spatial manifestation of the developmental pattern and the degree of engagement of particular large-scale brain networks exists.

The considerations above make clear that the age-specific characteristics of task induced brain activation have to be understood at the level of brain-wide networks and the maturational changes in their anatomical organization and processing efficacy that take place during adolescence (Grayson and Fair, 2017; Luna et al., 2015; Mills et al., 2016; Stevens et al., 2016; Vogel et al., 2010). The pattern of voxels identified by the MVPA that carry relevant information to distinguish between the two adolescent age groups involve many voxels, distributed throughout the brain. As discussed above, the pattern is not confined to task-responsive voxels, but also involves voxel clusters in default mode regions and in many structures that are not specifically task-responsive.

---

**Fig. 6.** Classification weight maps showing the spatial distribution of voxels contributing to age classification from metacognitive, cognitive and visual tasks. (A) Task-specific absolute weights (>±0.1) derived from training the classifier to discriminate age at the 30th iteration (with 8671 voxels still included) of the recursive feature elimination procedure. Weights for training on each of the three task datasets are overlaid in three different colors on the Collin brain template in MRICron (http://people.cas.sc.edu/rorden/mricron/index.html). (B) Voxel weights from training on the visual task dataset at the 30th iteration (>±0.05; positive weights in orange; negative weights in blue; A threshold >±0.1 would reveal almost all voxels). Most of the clusters revealed by the visual training are located at or close by voxel clusters that are also discriminative in age classification training based on the metacognitive data. While numbered arrows indicate exemplar clusters for which a ROI-based group analysis of the percent signal change (PSC) data was performed (M±SE bar diagrams on the right). These bar diagrams show that the underlying age-related PSC data are in agreement for the visual and the metacognitive data. (C) Different brain slice on the same data as in (B), showing that not all visual task derived clusters are in agreement with those obtained from training on the metacognitive data; exemplar clusters 5–6 only discriminate age when training on the visual dataset; exemplar clusters 7–8 have opposite discriminative weights for training on the visual dataset and on the metacognitive data set. Exemplar clusters are indicated by white, numbered arrows. These weight differences are in line with the age-related direction of PSC differences in the underlying PSC data for the two tasks, as indicated by the accompanying bar diagrams (M±SE).
The wide distribution confirms functional connectivity studies reporting a brain wide spreading of nodes contributing to age prediction (Dosenbach et al., 2010; Keulers et al., 2012; Rudolph et al., 2017; Tia et al., 2016). The wide distribution also agrees with reports that teen-age brain maturation takes place particularly in long-range connections, leading to improved whole-brain network integration (Dosenbach et al., 2007; Grayson and Fair, 2017; Kelly et al., 2009; Luna et al., 2015; Stevens, 2016; Vogel et al., 2010). Our own research showed that increased long range functional connectivity strength at this age occurred on a much larger scale than age-related differences in task-induced activation (Keulers et al., 2012). Based on such reports one would expect that all functional systems and networks increase their efficacy, internally as well as in interactions with one another. Consequently, voxels with age-specific activity modulation are found throughout all functional systems. It is the brain as a whole that becomes more optimized for information processing, rather than that the observed age-specific differences are driven by the maturation of areas specifically involved in certain neurocognitive functions.

Because we analyzed data from two phases of the same task in which every trial consisted of a metacognitive decision making and a cognitive encoding phase, the question could be raised whether the generalization between these two tasks can be explained by this specific design of the paradigm. If the cognitive encoding phase contains some of the brain activation that is evoked by the preceding metacognitive decision making phase, this shared brain activation might have driven the generalization of age classification. However, every precaution was taken to avoid such a contamination. First, the task was constructed as a slow event-related design with on average 14 s between the onsets of the 2 task phases. Second, additional temporal jittering of the trial-to-trial variances have however shown to be consistent (Jimura and Pol-
2007. Distinct brain networks for adaptive and stable task control in humans. Proc. Natl. Acad. Sci. U.S.A. 104, 11073–11078.

Dosenbach, N.U.F., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., et al., 2010. Prediction of individual brain maturity using fMRI. Science 329, 1358–1361.

Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cognit. Sci. 14, 172–179.

Duncan, J., Owen, A.M., 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci. 23, 475–483.

Dunn, L.M., Dunn, L.M., Schlitzch, J.E.P.T., 2005. Peabody Picture Vocabulary Test-III-NL. Harcourt Court Publishers.

Ernst, M., 2014. The triadic model perspective for the study of adolescent motivated behavior. Brain Cogn. 89, 104–111.

Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U., Church, J.A., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2009. Functional brain networks develop from a "local to distributed" organization. PLoS Comput. Biol. 5, 1–14.

Friston, K.J., Williams, S., Howard, R.D., Frackowiak, R.S.J., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.

Grayson, D.S., Fair, D.A., 2017. Development of large-scale functional networks from birth to adulthood: a guide to the neuroimaging literature. Neuroimage 160, 15–31.

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., Rapoport, J.L., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. Nat. Neurosci. 2, 861–863.

Giorgio, A., Watkins, K.E., Chadwick, M., James, S., Wium-Idend, L., Douaud, G., De Stefano, N., Matthews, P.M., Smith, S.M., Johannsen, B., James, A.C., 2010. Longitudinal changes in grey and white matter during adolescence. Neuroimage 49, 94–103.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Vaituzis, A.C., Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl. Acad. Sci. U.S.A. 101, 8174–8179.

Hart, H., Chantiluke, K., Fair, D.A., Power, J.D., Miezin, F.M., 2009. Percent signal change for fMRI calculations. Retrieved 23th. August.

Mayer, A.R., Franco, A.R., Ling, J., Canive, J.M., 2007. Assesment and quantification of head motion in neuropsychiatri functional imaging research as applied to multitask cognitive performance. Neuropsychol. Rev. 17, 103–117.

Scribah, D., Levin, D.J., Menon, V., 2008. A critical role for the right fronto-insula cortex in switching between central-executive and default-mode networks. Proc. Natl. Acad. Sci. U.S.A. 105, 12569–12574.

Stevens, M.C., 2016. The contributions of resting state and task-based functional connectivity studies to our understanding of adolescent brain network maturation. Neurosci. Biobehav. Rev. 70, 15–32.

Stiers, P., Falbo, L., Goulas, A., van Gog, T., de Bruijn, A., 2016. Inference of mental state from fMRI data by classifying mental states across individuals. Psychol. Sci. 20, 362–375.

Tia, L., Ma, L., Wang, L., 2016. Alterations of functional connectivities from early to middle adulthood: clues from multivariate pattern analysis of fMRI data. Neuroimage 129, 389–400.

Van Dijk, K.R.A., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59, 431–441.

Van Dijck, R.A., Huizenga, H.M., Somerville, L.H., Delgado, M.R., Powers, A., Weeda, W.D., Casey, B.J., Weber, E.U., Figner, B., 2015. Neural correlates of expected and unexpected reward in the human brain. Proc. Natl. Acad. Sci. U.S.A. 112, 1454–1459.

Vogel, A.C., Paus, T., 2010. Growth of white matter in the adolescent brain: myelin or axon. Brain Cogn. 72, 26–35.

Petersen, S.E., 2011. Development of anterior cingulate functional connectivity from late childhood through early adulthood. Cerebr. Cortex 19, 640–657.

Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood: a longitudinal MRI study. Nat. Neurosci. 2, 861–863.

Weeda, W.D., Casey, B.J., Weber, E.U., Figner, B., 2015. Neural correlates of expected and unexpected reward in the human brain. Proc. Natl. Acad. Sci. U.S.A. 112, 1454–1459.