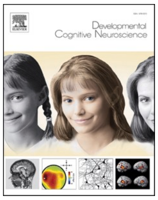
Developmental Cognitive Neuroscience 57 (2022) 101131

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

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



# Adolescent to young adult longitudinal development across 8 years for matching emotional stimuli during functional magnetic resonance imaging

Nora C. Vetter [a,b,*,1], Juliane H. Fröhner [a,2], Klara Hoffmann [b], Lea L. Backhausen [a,b,3], Michael N. Smolka [a,4]

[a] *Department of Psychiatry and Neuroimaging Center, Technische Universität Dresden, Germany*
[b] *Department of Child and Adolescent Psychiatry, Faculty of Medicine of the Technische Universität Dresden, Germany*

## ARTICLE INFO

*Keywords:*
Adolescence
Longitudinal
Functional magnetic resonance imaging
Emotion
Distraction
Reliability

## ABSTRACT

We investigated development from adolescence to young adulthood of neural bottom-up and top-down processes using a functional magnetic resonance imaging task on emotional attention. We followed 249 participants from age 14–22 in up to four waves resulting in 687 total scans of a matching task in which participants decided whether two pictures were the same including distracting emotional or neutral scenes. We applied generalized additive mixed models and a reliability approach for longitudinal analysis. Reaction times and error rates decreased longitudinally. For top-down processing, we found a longitudinal increase for the bilateral inferior frontal gyrus (IFG) for negative stimuli and in the left IFG also for positive and neutral stimuli. For bottom-up activation in the bilateral amygdala, we found a relative stability for negative and neutral stimuli. For positive stimuli, there was an increase starting in the twenties. Results show ongoing behavioral and top-down prefrontal development relatively independent from emotional valence. Amygdala bottom-up activation remained stable except for positive stimuli. Current findings add to the sparse literature on longitudinal top-down and bottom-up development into young adulthood and emphasize the role of reliability. These findings might help to characterize healthy in contrast to dysfunctional development of emotional attention.

## 1. Introduction

During adolescence, the interplay of subcortical regions involved in bottom-up emotional attention such as the amygdala and prefrontal regions involved in top-down control still mature within hierarchical circuit-based changes that seem to continue into the 20 s (Casey et al., 2016, 2019; Shulman et al., 2016). This maturational imbalance may lead to psychopathological symptoms, which often emerge in adolescence and young adulthood (Paus et al., 2008). Longitudinal studies on the neural development of emotional attention that span adolescence and young adulthood in a complete longitudinal design are largely lacking. Therefore, we investigated the longitudinal development of bottom-up and top-down systems across adolescence until young adulthood of emotional attention and included inter-individual variability. We were focusing on mid-adolescence to young adulthood as a transitioning phase and investigated continued development into young adulthood.

Emotional attention requires bottom-up and top-down systems to draw attention toward or away from emotional stimuli (Corbetta and Shulman, 2002; Dolcos et al., 2020). When the emotional content is irrelevant, the individual either gets distracted and turns their attention toward the negative or positive distractor (bottom-up attention) or successfully keeps it on the primary task (top-down attention; Vuilleumier and Huang, 2009). In adolescence, the maturational imbalance may lead to prepotent bottom-up processing (Casey et al., 2008). Therefore, adolescents may easily get distracted and turn their attention towards an emotional distractor such as seeing a (former) friend while focusing on the current exam in the classroom, instead of inhibiting their automatic impulse and keeping their attention on the primary task by ignoring the emotional distractor (Vuilleumier and Huang, 2009).

---

[*] Correspondence to: Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, 01307 Dresden, Germany.
*E-mail address:* Nora.vetter@tu-dresden.de (N.C. Vetter).

[1] https://orcid.org/0000-0003-1477-5395

[2] https://orcid.org/0000-0002-8493-6396

[3] https://orcid.org/0000-0001-6482-1316

[4] https://orcid.org/0000-0001-5398-5569

https://doi.org/10.1016/j.dcn.2022.101131
Received 24 February 2022; Received in revised form 3 June 2022; Accepted 29 June 2022
Available online 3 July 2022

Adolescent neural development of emotional attention was investigated by few earlier cross-sectional studies. Regarding top-down regions, some studies suggested a decrease across adolescence until young adulthood in the prefrontal cortex (PFC) with higher activation in adolescents versus adults for distracting emotional facial expressions while evaluating features (nose width, gender; Deeley et al., 2008; Monk et al., 2003). For the bottom-up region amygdala other studies showed an inverted U-shaped trajectory for amygdala activation with higher activation in adolescents compared to children and/or adults (Guyer et al., 2008; Hare et al., 2008). A more recent study showed that cognitive control remains susceptible to negative emotional information into young adulthood. Young adults continued to react specifically to negatively valenced emotional information together with an ongoing development of prefrontal circuitries (Cohen et al., 2016).

Longitudinal studies remain sparse and covered only related socio-cognitive processes. One study did not find PFC changes but increasing activity in early adolescence in the temporal pole (for passive viewing of emotional faces; Pfeifer et al., 2011). Another study found a linear increase across adolescence for inhibitory control in prefrontal regions, while amygdala bottom-up processing decreased (Paulsen et al., 2015). While the right IFG showed relative stability, age comparisons revealed a decrease and the medial PFC showed a dip of activation in mid-adolescence for social cognition (Overgaauw et al., 2015). For inhibitory control, one study found a decrease in PFC activation in 123 participants including ages 9–26 years (Ordaz et al., 2013; antisaccade task) while another one found a PFC increase with age in 290 participants from childhood to young adulthood including ages 7–23 years (Cope et al., 2020; go/no-go task). Taken together, some previous longitudinal studies suggest changes in prefrontal top-down and in bottom-up regions as the amygdala across adolescence. However, emotional attention was not targeted in these studies and methodically, they either had relatively small samples (about 30–50 participants) and only two to three time points or used accelerated designs in larger cohorts, reducing their ability to precisely model trajectories.

Therefore, it still remains open how top-down and bottom-up resources in emotional attention continue to develop across adolescence, especially when using a complete longitudinal design. Here, we employed a well-classified emotional attention task (Pilhatsch et al., 2014; Vetter et al., 2015, 2017, 2018), which presents a pair of non-emotional abstract pictures and a pair of pictures containing either negative, positive or neutral emotional scenes. Participants are asked to indicate whether the pictures of one pair are the same, while ignoring the other pair. Importantly, the emotional content of the pictures to match or to ignore is irrelevant. Our previous analysis with 144 adolescents from age 14–16 showed an increase of anterior cingulate cortex and bilateral IFG activation while amygdala activation remained stable (Vetter et al., 2015) for negative and positive versus neutral stimuli across conditions. This might indicate an ongoing top-down maturation in mid-adolescence when faced with emotional distractors. The amygdala showed a higher activation for negative stimuli in another analysis of this project in 164 fourteen-year-olds with a family history of depression (Pilhatsch et al., 2014). In a related study, we found an increased activation for negative stimuli in the left anterior insula reaching into the IFG for adolescents with attention-deficit-hyperactivity disorder (ADHD; Vetter et al., 2018). Taken together, although we could characterize adolescent development of emotional attention from age 14 until age 16, in risk factors and developmental disorders, the extended neural development until young adulthood is still unclear.

The current study sought to investigate the longitudinal development of emotional attention covering age 14–22 in a large community-based sample of 249 typically developing adolescents. We applied longitudinal flexible generalized additive mixed models (GAMM, Wood, 2004, 2006). Based on earlier studies (Deeley et al., 2008; Monk et al., 2003; Paulsen et al., 2015) and our previous analyses, we aimed to investigate development in prefrontal top-down regions and the bottom-up region amygdala (Guyer et al., 2008; Pilhatsch et al., 2014; Vetter et al., 2015). Matching emotional, specifically negative stimuli elicited stronger distraction for behavioral (slower reaction times for negative versus neutral, indicating an attention capture effect) and neural findings (as seen in parts of this data set: Pilhatsch et al., 2014; Vetter et al., 2015, and other analyses: Vetter et al., 2018). For matching abstract stimuli (completely ignoring emotional stimuli) there was no attention capture effect, presumably caused by heightened difficulty of comparing abstract, scrambled pictures (indicated by higher error rates). Therefore, we considered matching abstract stimuli not suitable for analysis and chose the matching emotional condition (matching negative/ positive/ neutral trials) for our models. Since reliability is especially important for longitudinal functional magnetic resonance imaging (fMRI; Bennett and Miller, 2013; Elliott et al., 2020; Herting et al., 2018) and given findings of parts of this data set (Vetter et al., 2015, 2017), we sought to use reliable regions of interest (ROIs) and conducted a reliability analysis first.

We expected error rates and reaction times to further decrease with age (Vetter et al., 2015). We hypothesized that activation in reliable PFC top-down control ROIs will continue to develop, i.e. increase with age for emotional attention when matching emotional but not neutral pictures, since there is evidence for increasing PFC activity during adolescence (Crone and Dahl, 2012), also shown in our data from age 14–16 (Vetter et al., 2015). As a bottom-up process, we expected activation in the amygdala to remain stable for emotional attention when matching emotional and neutral pictures (Vetter et al., 2015).

## 2. Materials and methods

### 2.1. Participants

Data acquisition was part of the project “The adolescent brain” (for more details see Ripke et al., 2012). The ethics committee of the Technische Universität Dresden approved the study and it was performed according to the Declaration of Helsinki. We recruited 260 adolescents at baseline via school visits in the local area. Informed consent was signed by participants and one of their legal guardians for participants younger than 18 at each wave of data acquisition. Exclusion criteria were a bipolar disorder, schizophrenia or major neuro-developmental disorders such as autism, a premature birth, head trauma or a history of several neurological or medical disorders. The sample stems from rather well-educated households (as a proxy for socioeconomic status), with around 60% of the parents having obtained a university or college degree (see Table 1); 97% of the sample had a white European ethnicity. A urine test assured no use of illicit drugs at the day of assessment. Participants had normal or corrected to normal vision.

We invited participants to four scanning sessions at the age of 14, 16, 18, and 22 (see Supplement A for procedure of assessment). We had to exclude some participants due to excessive head movements (>3 mm/ degrees per slice), low behavioral performance during the fMRI task, i.e. more than 25% incorrect answers or mean reaction times (RT) higher than 3 SDs from the mean of their age samples (see Table S1 for an overview of exclusion reasons). 249 participants completed at least one assessment.

See Table 1 for the participant characteristics of the total sample and Fig. 1 for the sample distribution across the four waves including gender and the longitudinal study design. Please note that some participants could not be scanned before their 15th birthday, therefore, we assessed them only from the second wave on. Due to this reason and due to exclusion criteria (see Table S1 in the supplements) 202 participants resulted at TP1.

### 2.2. Stimuli, design and procedure

The ACES task is a perceptual discrimination task on attention related emotion processing based on Vuilleumier et al. (2001), on which

**Table 1**
Participant characteristics (687 total scans).

| Parameter | TP 1 (N = 202) | TP 2 (N = 187) | TP 3 (N = 170) | TP 4 (N = 128) |
|---|---|---|---|---|
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Age in years | 14.5 (0.3) | 16.6 (0.4) | 18.7 (0.6) | 22.1 (0.7) |
| Age range | 13.6–15 | 15.7–17.9 | 17.3–21.3 | 20.1–24.8 |
| % (n) of females | 50% (101) | 50% (93) | 48% (82) | 49% (63) |
| State anxiety[a] | 36.7 (6.4) | 37.3 (6.8) | 36.2 (7.1) | 33 (5.8) |
| Trait anxiety[a] | 36.5 (7.2) | 37.1 (8) | 36.4 (7.5) | 34 (7.4) |
| Demographics - *Data from the first wave is reported* | | | | |
| IQ[b] | 113.03 (10.86) | | | |
| Handedness | 181 right-handers,<br>20 left-handers<br>1 bimanual | | | |
| Maternal education[c] | 3.67 (1.65) | | | |
| Paternal education[c] | 3.52 (1.76) | | | |

*Note*: TP =Time point.

[a] We measured state and trait anxiety with the State Trait anxiety Inventory (STAI Spielberger, 1983). For technical reasons, missings resulted in the STAI: TP1 n = 1, TP2 n = 34, TP3 n = 2, TP4 n = 5.

[b] General cognitive ability estimated with the subtests Similarities, Block Design, Vocabulary, and Matrices from the Wechsler Intelligence Scale For Children (WISC-IV, German adaptation; Petermann and Petermann, 2007);

[c] Questions about highest level of education from ESPAD (Hibell et al., 1997) ranging from 1 = "Professional qualification, e.g. PhD" to 7 = "Did not attend school or completed primary school education only".

our group published four publications (Pilhatsch et al., 2014; Vetter et al., 2015, 2017, 2018). Participants had to decide whether a pair of visual target stimuli was equal while another pair was presented as a distractor. In each trial a pair of non-emotional pictures and a pair of pictures from one of three emotional categories (negative, positive, neutral) taken from the International Affective Picture System (IAPS; Lang et al., 2005) was shown. The selection of IAPS stimuli was balanced with respect to normed emotional valence and arousal ratings (see Supplement of Vetter et al., 2015: https://bit.ly/3rggS7V). We created non-emotional abstract control pictures by shredding the chosen IAPS pictures beyond recognition using picture manipulation software (www.gimp.org). Because the development of emotion recognition in facial stimuli (Vetter et al., 2013, 2014) might interfere with the development of emotional attention targeted here, we used emotional scenes rather than faces. We arranged one pair horizontally and the other vertically around a fixation cross (Fig. 2). Participants had to attend to the horizontal or vertical pair for a given trial as indicated by a task cue (double-arrow) and report whether the two items of the pair were the same (which was the case in 50%).

In half of the trials participants had to match emotional pictures (matching emotional) and in the other half shredded pictures (matching abstract, ignoring emotional). Positioning of IAPS or shredded pairs was random. Altogether, there were six different trial types: Matching emotional (negative, positive, neutral) while ignoring abstract (shredded control stimuli) and ignoring emotional (negative, positive, neutral) while matching abstract (shredded control stimuli). The matching emotional and matching abstract conditions were presented counterbalanced. In total, there were twenty trials per condition, pseudo-randomly interleaved by jittered inter-stimulus intervals. One trial consisted of the following phases: After a task cue (1 s), two picture pairs were shown for 1 s on the next screen. During the presentation of the picture screen and the following 1.5 s, the participant responded via button press (maximum time to answer 2.5 s). After the picture screen jittered inter-stimulus intervals were employed (mean: 5000 ms, range: 3000–7000 ms) presenting a fixation cross. Dependent on the reaction time the jitter was flexibly adapted resulting in a mean trial length of seven seconds. The functional run lasted about 14 min and included 120 trials. Behavioral data were collected by ResponseGrips (©NordicNeuroLab) with a button on a grip in each hand. The presentation of the task and the recording of the behavioral responses was conducted with Presentation® software (version 11.1, Neurobehavioral Systems, Inc., Albany, CA). Inside the scanner, participants practiced the task prior to the scanning session with stimuli not used in the experiment. In our analyses, we focused on trials with negative pictures. As performance measures, we extracted RT and percentage of errors computed as proportion of incorrect and missed trials relative to total of matching negative, positive, or neutral trials (N = 20 each).

## 2.3. *Functional imaging*

### 2.3.1. *Image acquisition*

Scanning was performed with a 3 T whole-body MR tomograph (Magnetom TRIO, Siemens, Erlangen, Germany) with a 12-channel head coil during the first two time points and a 32-channel head coil during the last two time points (with an additional scanner software update from VB 15 to VB 17a). The following sequence parameters were used for each of the acquisition waves. For functional imaging a standard Echo Planar Imaging (EPI) sequence was used (repetition time, TR/ echo time, TE: 2410/25 ms; flip angle: 80°). FMRI scans contained 42 transversal slices, tilted up 30° clockwise from the anterior commissure–posterior commissure line. Voxel size was 3×3x3mm (thickness of 2 mm with 1 mm gap, field of view (FOV) of 192×192 mm, in-plane resolution of 64 × 64 pixels). Furthermore, a 3D T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) image data set was acquired (TR/TE = 1900/2.26 ms, FOV = 256×256 mm, 176 slices, 1×1x1mm voxel size, flip angle = 9°). During functional imaging the task images were presented via a head-coil-mounted display system based on LCD technology (NordicNeuroLab AS, Bergen, Norway).

## 2.4. *Analysis of fMRI data*

### 2.4.1. *Preprocessing*

We preprocessed and statistically analyzed functional images using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK). First, functional images were slice-time corrected by using the middle slice as reference and then realigned to the first image by 6-degree rigid spatial transformation. Scans were spatially normalized to the standard space defined by the Montreal Neurological Institute (MNI) EPI template and smoothed with a Gaussian kernel of 8 mm at full-width half maximum.

### 2.4.2. *Statistical analysis*

*2.4.2.1. First-level analysis.* We computed in the first-level analysis, within a general linear model (GLM) a fixed effects analysis for each participant within each voxel of the brain. The different conditions (attention and emotional valence) were modeled as six regressors of interest (matching (1) negative, (2) positive and (3) neutral, as well as matching abstract while ignoring (4) negative, (5) positive, and (6) neutral). These regressors of interest were modeled at the point as stick functions convolved with a canonical hemodynamic response function. Trials with missing or wrong responses were modeled as a separate regressor, integrated as a covariate of no interest. Thus, only correct answers were analyzed. Controlling for movement parameters, the six subject-specific movement regressors (three rotation and three translation parameters), derived from the rigid-body realignment, were also included as covariates of no interest. Each regressor of this within-subject GLM served as a regressor in a multiple regression analysis. A high-pass filter with a cut-off of 128 s was used to remove low frequency physiological noise (Henson, 2006). Moreover, an AR(1) model was applied for the residual temporal autocorrelation (Henson, 2006).

*2.4.2.2. Reliability analyses.* We continued with an exploratory

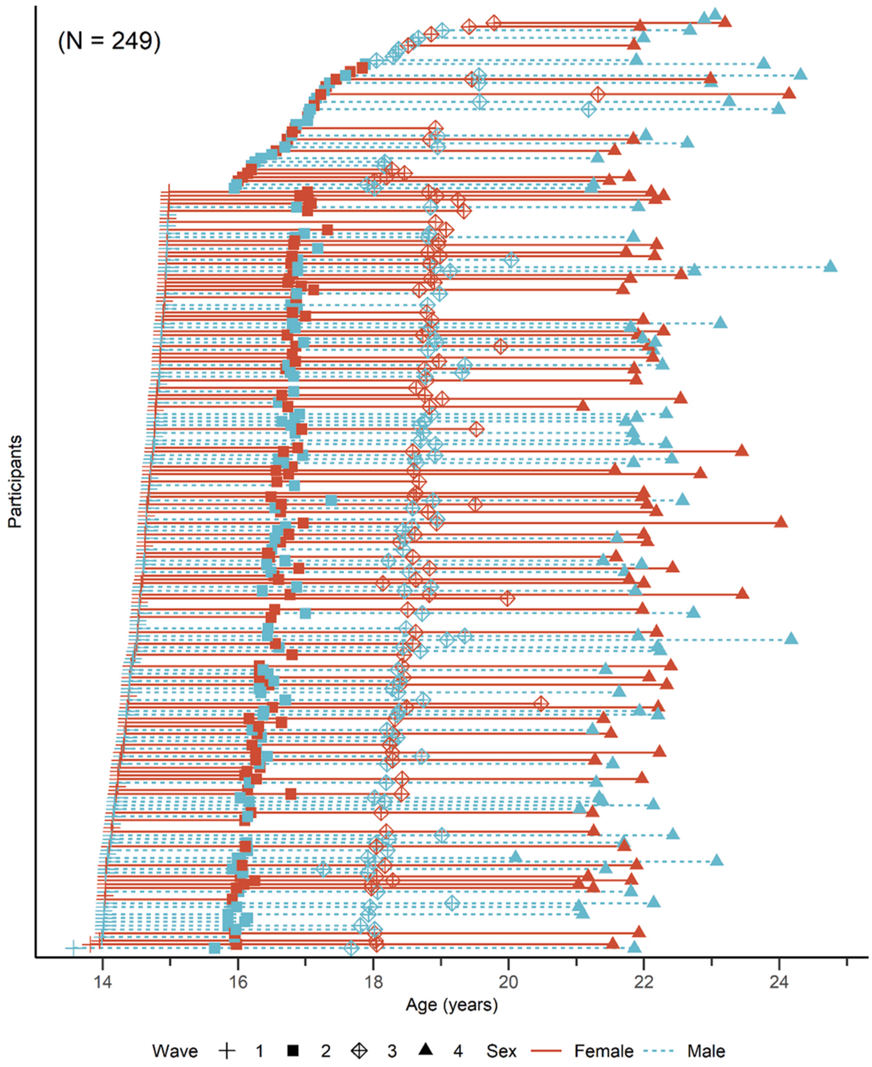


**Fig. 1.** Sample distribution across the four waves including sex.

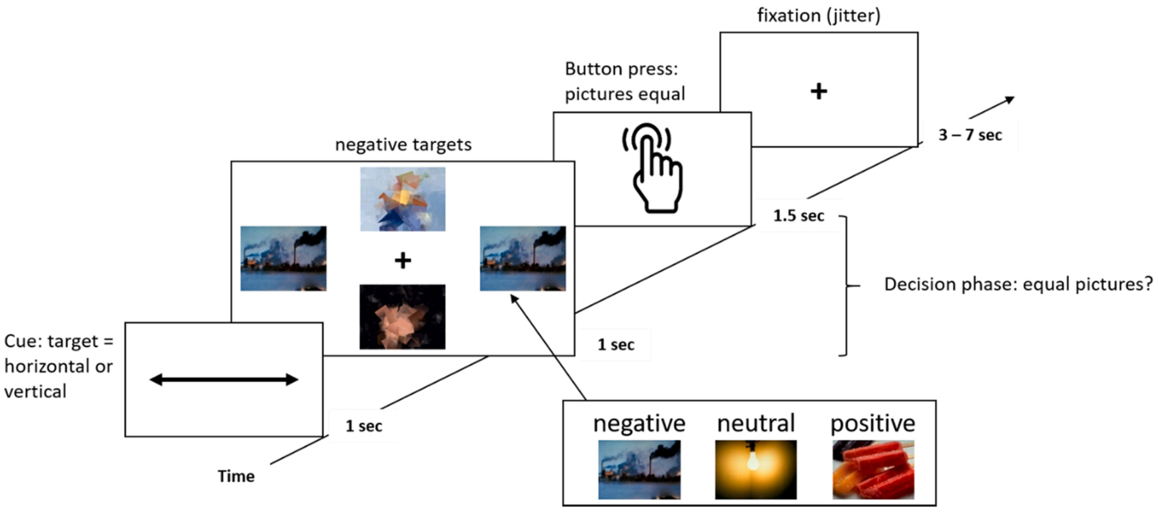


**Fig. 2.** Example trial for the matching emotional condition. See Fig. S1 for the ignoring emotion condition.

reliability analysis using data of the 81 participants who completed all four waves to follow up on our previous paper and the important issue of reliability (Vetter et al., 2015, 2017). We used the *fmreli* toolbox (Fröhner et al., 2019; https://github.com/nkroemer/reliability) to compute voxel-wise intraclass coefficients between the acquisition waves (ICC; Shrout and Fleiss, 1979) using the ICC(3,1) as in our previous reliability analyses (Vetter et al., 2017). We computed the average ICC across all age comparisons and set the threshold to at least moderately reliable voxels (ICC>0.4). We chose a moderate threshold given that intraindividual changes and interindividual variability in those might limit the upper bound of ICC values. In addition, we used the split-half feature of *fmreli* to compare longitudinal reliability with reliability within one session (Fröhner et al., 2019).

We first analyzed reliability of matching emotional versus neutral stimuli as this was the focus of our previous analyses (Vetter et al., 2015, 2017, 2018). On the whole-brain level in our current n = 81 sample this contrast elicited activation in the superior frontal gyrus (overlapping with the anatomical IFG), the middle and posterior cingulate gyrus and the inferior occipital gyrus. Since there were no reliable voxels > 0.4 in the contrast matching emotional versus neutral for longitudinal as well as split-half reliability, we considered the contrast as not suitable for an analysis of individual differences (see https://neurovault.org/collections/LKPJWWEA/ and supplementary Tables S5 and S6).

We therefore focused on the contrast all stimuli (negative/positive/neutral) versus implsuppicit baseline, which elicits broad whole-brain activation. We decided to include neutral stimuli to show the development in longitudinal models also for neutral stimuli. In terms of longitudinal reliability, we found moderately reliable voxels in a bilateral frontal cluster among other occipital and temporal clusters (Fig. 3; and for ICC and main-effect maps see our repository https://neurovault.org/collections/LKPJWWEA/). The bilateral frontal cluster covers medial frontal regions overlapping with the anatomical IFG in the WFU-PickAtl<as with the Talairach Daemon (TD) Brodman atlas (Lancaster et al., 2000; Maldjian et al., 2003, 2004). In our previous analysis, we found increasing IFG activity from age 14–16 and interpreted this as increasing top-down control (Vetter et al., 2015). Therefore, we chose the bilateral reliable cluster, which partly included the bilateral anatomical IFG as ROIs (left IFGreli, right IFGreli). Please note that split-half reliabilities for this contrast were high (see https://neurovault.org/collections/LKPJWWEA/ and Table S4). We decided to use longitudinal reliability and the moderate threshold, because we were interested to find a reliable marker of individual change in top-down control, which can further be associated with other variables of interest.

We aimed at investigating the interplay of top-down and bottom-up regions in longitudinal adolescent development. The amygdala as a bottom-up region also plays an important role for the task (see activation for matching negative stimuli) and its activation for negative stimuli might serve as a potential neural marker for psychopathology in family history of depression (Pilhatsch et al., 2014) or ADHD (Vetter et al., 2018). Similar to our previous publications (Pilhatsch et al., 2014; Vetter et al., 2015, 2018), we defined the amygdala as a left and right anatomical ROI with the WFU-pickatlas using the TD Atlas.

However, it has to be noted, that the reliability analyses did not show any voxels in the amygdala with an ICC above 0.4 and thus the results have to be interpreted with caution (See Supplement C for reliability values). We extracted percent signal change from the bilateral IFGreli and amygdala ROIs using rfxplot (Gläscher, 2009).

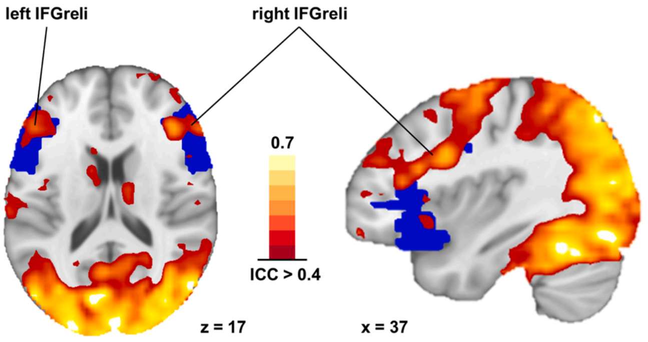


**Fig. 3.** Reliable IFG ROIs (left and right IFGreli) for the “all stimuli (negative, positive, neutral) versus implicit baseline” contrast exceeding a threshold of an average ICC across all comparisons > 0.4. The IFGreli ROIs overlap with the anatomical IFG of the WFU-PickAtlas with the Talairach Daemon (TD) Brodman atlas (blue).

*2.4.2.3. GAMM analyses depicting age trajectories.* GAMM fit a combination of unknown basis smooth functions of predictor variables to best predict the outcome variable. GAMM allow the inclusion of participants with missing data, i.e. the whole sample with n = 249 participants. We used GAMM to characterize behavioral and fMRI trajectories. Thus, we report GAMM with age as predictor and the following outcome variables: percentage of errors, RT, signal from the left and right IFGreli ROIs as top-down activation and signal from the left and right amygdala ROIs as bottom-up activation. We used the emotional conditions (negative versus implicit baseline, positive versus implicit baseline) and as a contrasting condition also neutral versus implicit baseline. As GAMMs do not require a priori knowledge about the inherent form of the data, they flexibly allow for different shapes of nonlinear growth. We used R version 3.6.1 with the mgcv package version 1.8–33 (Wood, 2004, 2006) and the packages itsadug version 2.3 and ggplot2 version 3.3.2 for visualization of smooth curves. The k parameters, which specify the number of basis smooth functions to build the curves and thus influence potential over-fitting, were set to four following results from k.check (mgcv package) analyses of the models. All models included a random intercept per participant and shrinkage versions of cubic regression splines as smooth terms.

Additionally, we computed difference scores for each outcome variable for participants with data at age 14 and 22 (N = 104). To characterize the difference as decreasing, stable or increasing, we computed individual SD over all measurements per emotional valence and averaged them. The average SDs were: 1) for the left IFGreli: SD= 0.09 for negative, SD= 0.08 each for positive and neutral stimuli 2) for the right IFGreli: SD= 0.08 for negative, SD= 0.07 for positive and SD = 0.08 for neutral stimuli 3) for the left amygdala SD= 0.14 for negative, SD= 0.13 each for positive and neutral stimuli 4) for the right amygdala: SD= 0.13 for negative, SD= 0.12 each for positive and neutral stimuli. Differences with an absolute value lower than 1 average SD were categorized as stable (following the approach of Narvacan et al., 2017). Accordingly, we categorized significant differences greater than ± 1 mean SD as de- or increasing.

## 2.5. Behavioral analysis

### 2.5.1. Reliability: ICC analysis

With SPSS for Windows (version 18) we computed intraclass coefficients between the acquisition waves using the ICC(3,1) (ICC; Shrout and Fleiss, 1979) for RT and percentage of errors for matching stimuli (negative, positive, neutral) for the n = 81 of the whole data set. In addition, we computed ICCs of difference scores (negtaive or positive minus neutral) for RTs and percentrage of errors.

### 2.5.2. Individual trajectories: GAMM analyses

Similarly as for imaging data, we report GAMM with age as predictor and the outcome variables percentage of errors and RT for matching negative, positive and neutral stimuli for all 249 participants including those with missing data. Additionally, we computed difference scores for each outcome variable for participants with data at age 14 and 22 (n = 104). To characterize the difference as decreasing, stable or increasing, we computed individual SD over all measurements per emotional valence and averaged them. Differences with an absolute value lower than the average SD were categorized as stable (following the approach of Narvacan et al., 2017). The average SDs were: 1) for RT: SD= 43.55 for negative, SD= 43.4 for positive and SD = 38.83 for neutral stimuli 2) for percentage of errors: SD= 0.04 each for negative, positive and neutral stimuli. Differences with an absolute value lower than 1 average SD were categorized as stable (following the approach of Narvacan et al., 2017). Accordingly, we categorized significant differences greater than ± 1 mean SD as de- or increasing.

## 3. Results

### 3.1. Behavioral results

#### 3.1.1. Reliability: ICC analysis

RTs were moderately reliable for attending negative (average ICC across comparisons = 0.47), positive (ICC = 0.45), and neutral (ICC = 0.47) stimuli. Percentage of errors showed poor reliability for attending negative (ICC = 0.21), positive (ICC = 0.2), and neutral (ICC = 0.2) stimuli (for values of matching abstract stimuli see supplement D). Reliabilities of difference scores for percentage of errors were even lower (matching negative minus neutral: ICC = 0.14, matching positive minus neutral: ICC = 0.12). In addition, ICCs for RT differences for matching negative or positive minus neutral stimuli could not be estimated due to negative covariances. For ANOVAs please see supplement B “Behavioral data (ANOVAS) for the n = 202 sample and the n = 81 sample”.

#### 3.1.2. Individual trajectories: GAMM analyses

*3.1.2.1. Reaction times.* Individual trajectories show that reaction times decreased over time (see Fig. 4 and Table 2 for a summary of GAMM results). On the individual level from age 14–22 (see Fig. 4), the majority, around 2/3rd (66.3% for negative, 65.4% each for positive and neutral) of the participants showed decreasing RT, while RT remained stable for around a third (30.8% for negative, 31.7% for positive, 32.7 for neutral) and increased for only a low percentage under 5% (2.9% each for negative and positive, 4.8% for neutral), (for values of matching abstract stimuli see supplement E).

*3.1.2.2. Percentage of errors.* Individual trajectories show that the percentage of errors decreased over time (see Fig. 5 and Table 2 for a summary of GAMM results). On the individual level from age 14–22 (see Fig. 5), the majority, around 2/3rd (58.7% for negative, 63.5% for positive, and 70.2% for neutral) of the participants showed stable percentage of errors, while percentage of errors decreased for around a third (31.7% for negative, 29.8% for positive, 26% for neutral) and increased for only a low percentage under 10% (9.6% for negative, 6.7% for positive, 3.8% for neutral), (for values of matching abstract stimuli see supplement E).

**Table 2**
Summary of GAMM results for matching emotional stimuli.

| | edf | Ref.df | *F* | *p* |
|---|---|---|---|---|
| Percentage of errors | | | | |
| Negative | 1.57 | 3 | 12.18 | < 0.001 |
| Positive | 1.64 | 3 | 11.14 | < 0.001 |
| Neutral | 1.80 | 3 | 17.67 | < 0.001 |
| Reaction times | | | | |
| Negative | 2.04 | 3 | 51.34 | < 0.001 |
| Positive | 2.13 | 3 | 59.84 | < 0.001 |
| Neutral | 2.09 | 3 | 56.07 | < 0.001 |
| IFGreli left | | | | |
| Negative | 1.93 | 3 | 7.08 | < 0.001 |
| Positive | 1.98 | 3 | 11.74 | < 0.001 |
| Neutral | 1.51 | 3 | 2.45 | .009 |
| IFGreli right | | | | |
| Negative | 1.60 | 3 | 2.165 | .017 |
| Positive | 0.80 | 3 | 0.65 | .105 |
| Neutral | 1.43 | 3 | 0.88 | .153 |
| Amygdala left | | | | |
| Negative | 1.74 | 3 | 1.63 | .055 |
| Positive | 2.04 | 3 | 4.02 | .002 |
| Neutral | 1.39 | 3 | 0.77 | .182 |
| Amygdala right | | | | |
| Negative | 1.75 | 3 | 1.53 | .066 |
| Positive | 2.41 | 3 | 6.69 | < 0.001 |
| Neutral | $5.95e^{-07}$ | 3 | 0 | .651 |

*Note*: IFG = inferior frontal gyrus. The values refer to the smooth including age S (age).

### 3.2. Imaging results

#### 3.2.1. Individual trajectories: GAMM analyses inferior frontal gyrus

On average, the signal in the left and right IFGreli increased over time (see Figs. 6 and 7 and Table 2). On the individual level from age 14–22 (see Figs. 6 and 7), the signal for attending negative, neutral or positive decreased for a low amount, in around 1/5th of participants (negative left: 13.6%, right: 20.2%; neutral left: 19.2%, right: 20.2%; positive left: 8.7%, right: 21.2%). For the left IFGreli 39.8% remained stable and 46.6% showed an increase for attending negative; 46.2% remained stable while 34.6% showed an increase for attending neutral.

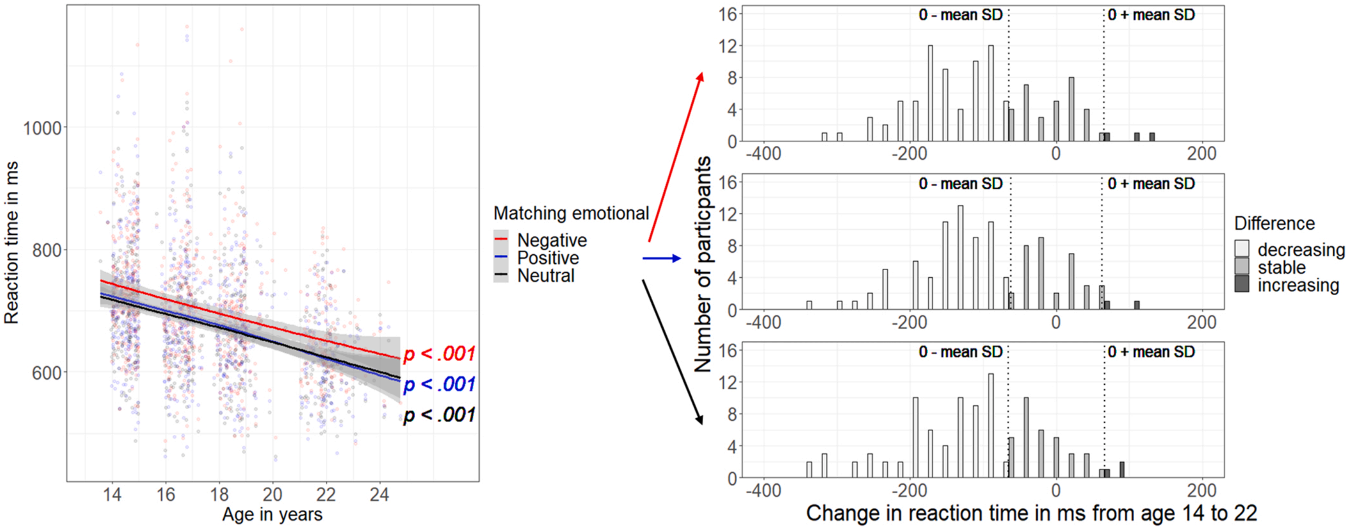


**Fig. 4.** Developmental trajectories and change-scores for reaction times. Left: GAMMs, confidence interval: 95%, one dot representing one participant. Right: Histograms of differences from age 14–22 of 104 participants that completed measurements at age 14 and 22 (see methods). We categorized significant differences greater than ± 1 mean SD (see methods) as de- or increasing. To increase visibility of single dots, we slightly jittered their position for percentage errors (position. jitter function of ggplot).

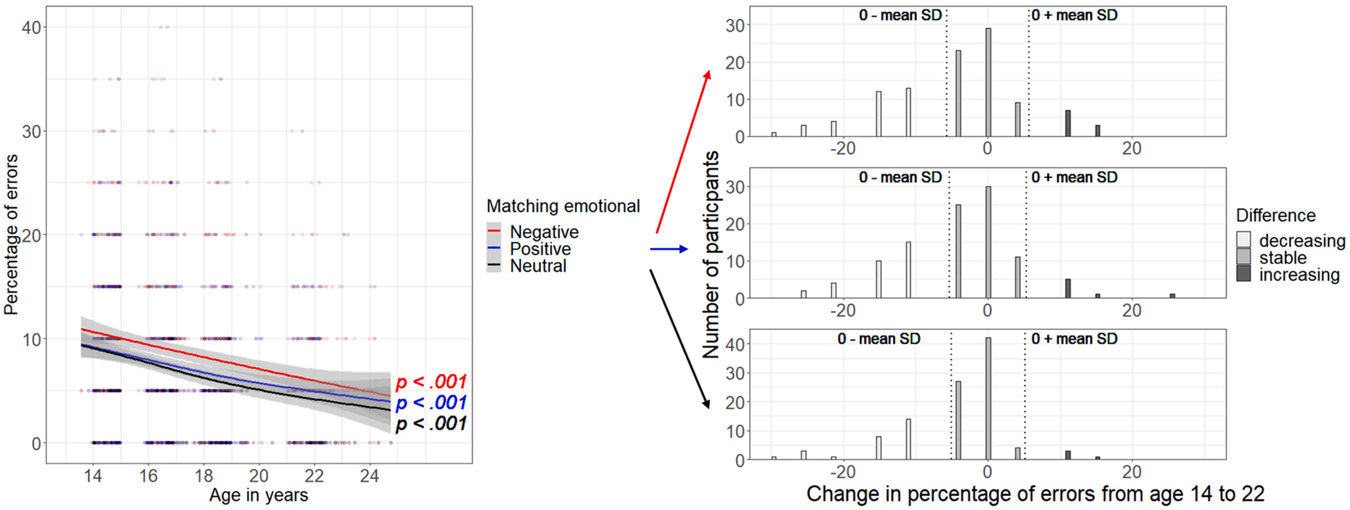


**Fig. 5.** Developmental trajectories and change scores for percentage of errors. Left: GAMMs, confidence interval: 95%, one dot representing one participant. Right: Histograms of differences from age 14–22 of 104 participants that completed measurements at age 14 and 22 (see methods). We categorized significant differences greater than ± 1 mean SD (see methods) as de- or increasing. To increase visibility of single dots, we slightly jittered their position for percentage errors (position. jitter function of ggplot).

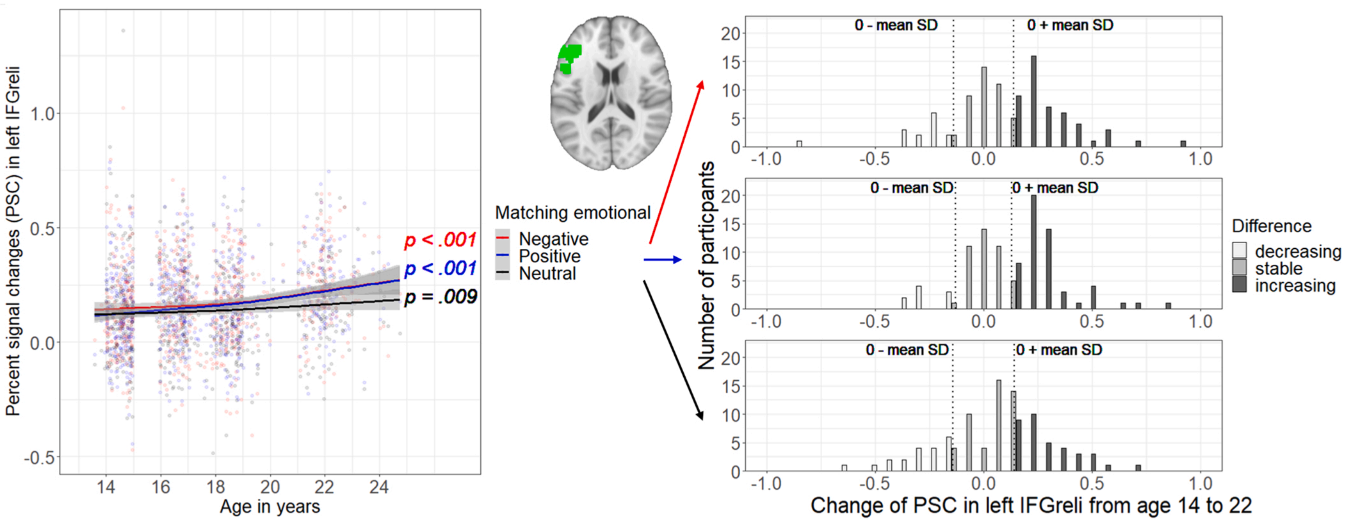


**Fig. 6.** Developmental trajectories and change scores for percent signal changes (PSC) in the left reliable inferior frontal gyrus ROI (IFGreli). Left: GAMMs of PSCs, confidence interval: 95%, one dot representing one participant. Right: Histograms of difference in PSC from age 14–22 of 104 participants that completed measurements at age 14 and 22 (see methods). We categorized significant differences greater than ± 1 mean SD (see methods) as de- or increasing.

For attending positive in the left IFGreli 40.4% remained stable and 51% showed an increase. For the right IFGreli 42.3% remained stable while 37.5% showed an increase for attending negative. For attending positive in the right IFGreli 47.1% remained stable and 31.71% showed an increase and for attending neutral 59.6% remained stable and 20.2% increased. For values of matching abstract stimuli see supplement E.

### 3.2.2. *Individual trajectories: GAMM analyses amygdala*

Of note, the left and the right amygdala ROI showed an average ICC lower than 0.3 and results have to be dealt with caution (see Table S2). On average, the signal in the left and right amygdala did not change for neutral stimuli (see Figs. 8, 9 and Table 2) but changed for attending positive stimuli (with a $p < .001$ for the right and $p = .002$ for the left amygdala). The trajectory for positive stimuli seemed to decrease slightly until about age 18 and then increased until age 24. Although a similar trajectory seems to result for negative stimuli, the trajectory change was only significant at trend (left:.005 and right:.066). On the individual level, around 50% of the participants remained stable (negative left 48.1%, negative right: 49.5%; positive left 51%, positive right: 48.1%; neutral left 56.7%, neutral right: 47.6%), around 30% showed an increase from age 14–22, except for left neutral left with only 22.1% (negative left 28.8%, negative right: 29.1%; positive left 30.8%, positive right: 36.5%; neutral right: 29.1%) and overall only around 20% decreased: (negative left 23.1%, negative right: 21.4%; positive left 18.3%, positive right: 15.4%; neutral left: 21.2%, neutral right: 23.3%). For values of matching abstract stimuli see supplement E.

## 4. Discussion

This longitudinal study investigated for the first time adolescent development of neural bottom-up and top-down processes underlying emotional attention when matching emotional stimuli from age 14–22 in a large sample of n=249 participants with up to four time points in a complete longitudinal design. The task is to match the two pictures irrespective of emotional valence. Thus, the emotional content of the pictures is working as a distractor here. Behavioral results indicate an

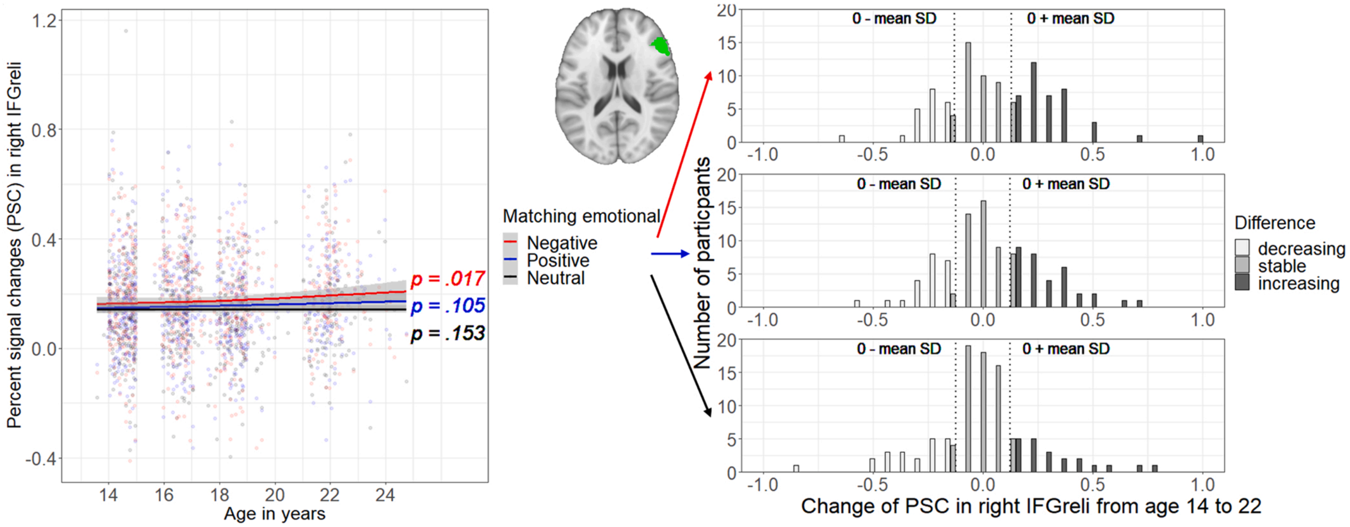


**Fig. 7.** Developmental trajectories and change scores for percent signal changes (PSC) in the right reliable inferior frontal gyrus ROI (IFGreli). Left: GAMMs of PSCs, confidence interval: 95%, one dot representing one participant. Right: Histograms of difference in PSC from age 14–22 of 104 participants that completed measurements at age 14 and 22 (see methods). We categorized significant differences greater than ± 1 mean SD (see methods) as de- or increasing.

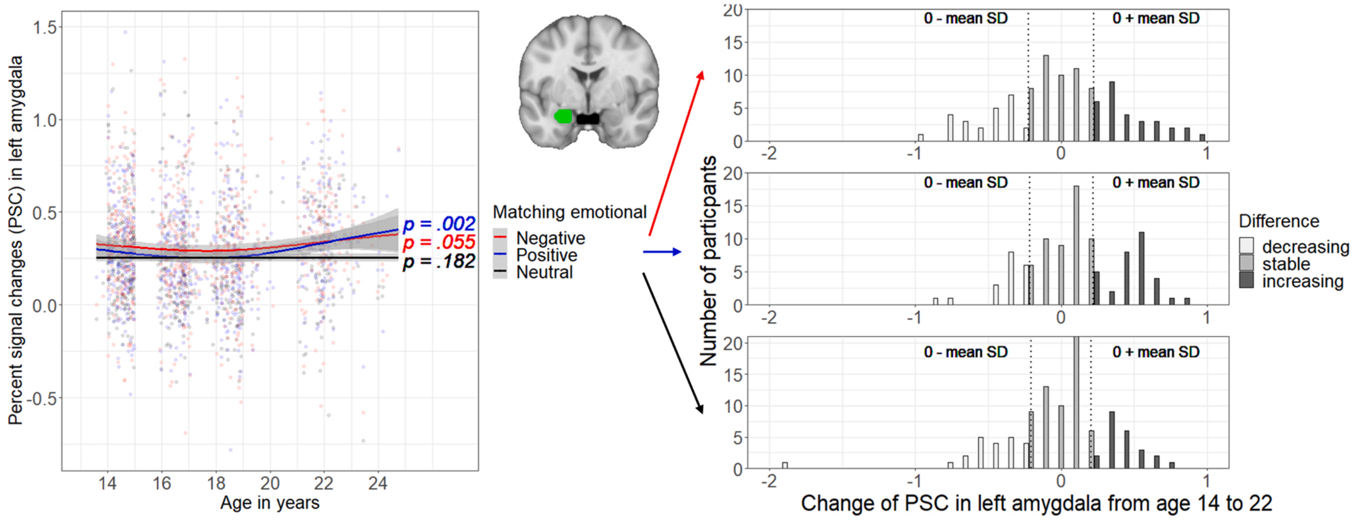


**Fig. 8.** Developmental trajectories and change scores for percent signal changes (PSC) in the left anatomical Amygdala ROI. Left: GAMMs of PSCs, confidence interval: 95%, one dot representing one participant. Right: Histograms of difference in PSC from age 14–22 of 104 participants that completed measurements at age 14 and 22 (see methods). We categorized significant differences greater than ± 1 mean SD (see methods) as de- or increasing.

expected improvement of performance (i.e. decreasing RT and percentage of errors) from age 14–22 for all conditions including matching negative, positive and neutral.

Partly confirming our hypothesis, for top-down processing, we found an increase for emotional but also for neutral stimuli in the left reliable IFG and only for negative stimuli for the right reliable IFG. In line with our hypothesis, for bottom-up activation, we found a relative stability for negative and neutral stimuli in the bilateral amygdala. In contrast, unexpectedly, we found an increase starting in the twenties for positive stimuli in the bilateral amygdala.

### *4.1. Behavioral development*

Behaviorally, RT GAMM trajectories showed a decrease for emotional attention, which was reflected by two third of the participants decreasing in individual difference scores. Still, there was some inter-individual variance in difference scores: around one third of participants did not show an improvement in RT. For percentage of errors, we also found an overall decreasing trajectory in GAMMs. Here, larger inter-individual variance in difference scores arose: only one third also decreased and about two third of the participants showed a stable individual difference score, presumably due to lower variance in difference scores. Overall, the behavioral longitudinal improvement shown in GAMMs for both RT and percentage of errors is in line with a previous analysis of this task from age 14–16 (Vetter et al., 2015) and studies of adolescent emotional attention (Cohen-Gilbert and Thomas, 2013; Grose-Fifer et al., 2013). Interestingly, at age 18 and 22 participants made more errors when matching negative versus neutral and at age 22 also positive versus neutral pictures (see Figure S3 in the supplements), which was not seen at ages 14 or 16. For negative stimuli this mirrors the attention capture effect in RT (see Figure S2 in the supplements).

### *4.2. Neural development*

#### *4.2.1. Inferior frontal gyrus*

We found an increasing GAMM trajectory in the left reliable IFG ROI

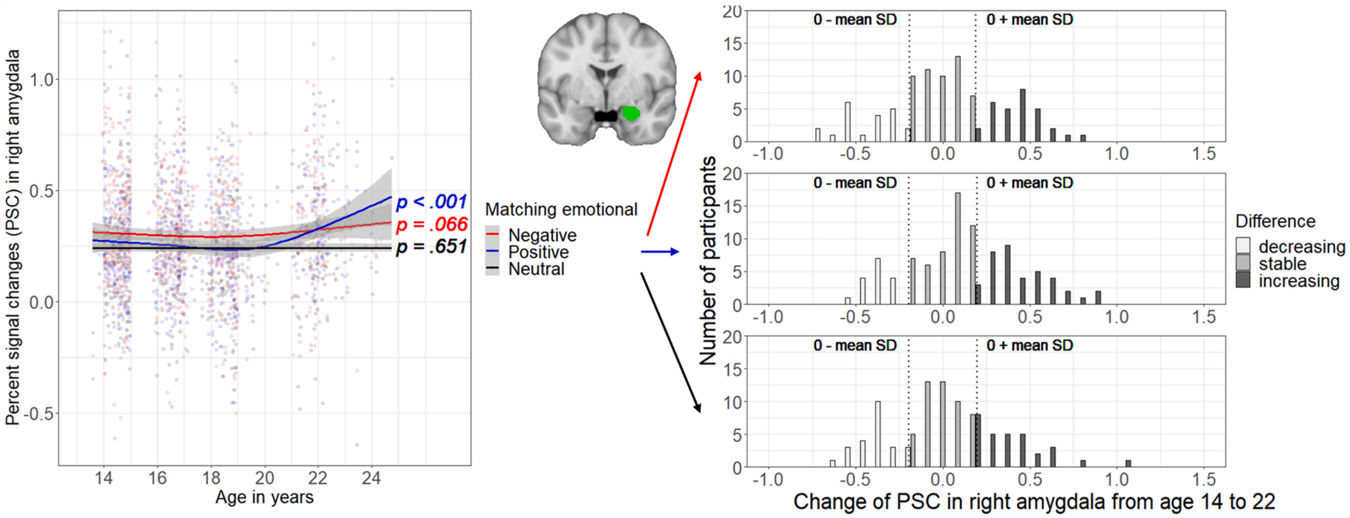


**Fig. 9.** Developmental trajectories and histograms of change scores for percent signal changes (PSC) in the right anatomical Amygdala ROI. Left: GAMMs of PSCs, confidence interval: 95%, one dot representing one participant. Right: Histograms of difference in PSC from age 14–22 of 104 participants that completed measurements at age 14 and 22 (see methods). We categorized significant differences greater than ± 1 mean SD (see methods) as de- or increasing.

for matching negative, positive and neutral stimuli from age 14–22, which is reflected by 35–51% of participants having an increasing individual difference score. We also found this GAMM increase in the right reliable IFG but only for negative stimuli corresponding to 38% of individual difference scores. Of note, there was substantial inter-individual variability in individual difference scores, i.e. not all participants show this increase, which is in accordance with inter-individual differences in anatomical prefrontal maturation (Mills et al., 2021). Current findings thus demonstrate a significant neural development between 14 and 22 years in cognitive control of emotions (Cromheeke and Mueller, 2014) but also neutral stimuli (in the left IFG) by an increased activation of top-down regions. This is partly consistent with our previous results (Vetter et al., 2015). From a broader perspective, adolescent top-down control seems to increase according to previous cross-sectional studies (Deeley et al., 2008; Monk et al., 2003) and longitudinal findings in reward tasks (Paulsen et al., 2015) or inhibitory control (Cope et al., 2020). However, there have been contradictory findings in longitudinal studies with PFC decreases or no PFC changes (Ordaz et al., 2013; Overgaauw et al., 2015; Pfeifer et al., 2011). Comparisons with these studies are difficult due to either the small samples (n = 38), different age rages (e.g. 10–13 years), tasks (theory of mind task, passive viewing of emotional faces, antisaccade task) or prefrontal clusters. Clearly, more longitudinal studies are required. Overall, the increased top-down control is in accordance with ongoing anatomical prefrontal maturation across adolescence (Gogtay et al., 2004; Tamnes et al., 2017). It is also in line with the imbalance model of adolescent brain development that posits an ongoing maturation in the PFC within hierarchical circuit-based changes underlying emotion regulation and extending into the 20 s (Casey et al., 2019). Beyond adolescence, also early adulthood is a sensitive phase for processing emotional cues (Bos et al., 2020; Casey et al., 2019).

#### 4.2.2. Amygdala

Importantly, the mean reliability across time points for amygdala activation was poor (mean ICC right/left=0.21/0.29, see Table S2). Therefore trajectories have to be interpreted cautiously with this background. For the amygdala we found a relative stability for neutral and negative stimuli (although significant at trend-level), while at about age 22 there was an increase for positive stimuli. There were substantial inter-individual differences with about 50% of participant showing stability and 22–30% an increase, while around 20% decreased. The increasing trajectory of amygdala activation for positive stimuli (and at trend for negative) at about age 22 is a nice addition to a previous longitudinal study (Pfeifer et al., 2011) that found an increase in right amygdala activation from age 10–13 for sad faces. This age span, however, was not included in our study. In contrast to our findings are a previous longitudinal (Paulsen et al., 2015) and two previous cross-sectional studies (Guyer et al., 2008; Hare et al., 2008). A ventral amygdala decrease from age 10–22 was found in a reward task for loss trials (Paulsen et al., 2015). The two cross-sectional studies indicated heightened amygdala activation for negative stimuli or stimuli with different valences in adolescents compared to children and/or adults, suggesting an inverted U-shaped function (Guyer et al., 2008; Hare et al., 2008) or a continued decrease from childhood across adolescence for negative scenes (Decety et al., 2012). The amygdala seems to be important for our task since it is activated specifically for negative emotional stimuli and sensitive for (emerging) psychopathology (Pilhatsch et al., 2014; Vetter et al., 2018). However, the developmental trajectory for negative stimuli did not change significantly, only at-trend level. In contrast, the trajectory increased for positive stimuli at about age 22. This is surprising, given previous findings of an enhanced amygdala activity in adolescents versus adults during encoding of positive IAPS pictures (Vasa et al., 2011) and a behavioral positivity bias in older adults (Carstensen and Mikels, 2005). The enhanced amygdala activation at young adulthood is partly in line with the behavioral results of more errors for positive versus neutral stimuli at age 22 (see Figure S3 in the supplements). Future longitudinal studies that investigate emotional attention of positive, negative and neutral stimuli both behaviorally and in fMRI might shed more light on this finding.

Taken together, our behavioral findings regarding error rates and the neural findings in the PFC and amygdala point towards a continued sensitivity in young adulthood to negative and positive emotional information together with an ongoing prefrontal and amygdala development. Similarly, a recent study found that cognitive control remains influenced by negative emotional information into young adulthood as prefrontal circuitries continue to develop (Cohen et al., 2016).

### 4.3. Methodical aspect: reliability

Current findings on reliability extend our previous reliability analysis on age 14–16 (Vetter et al., 2017) with more time points and demonstrate behavioral as well as fMRI whole brain and ROI-specific reliability across adolescence into young adulthood for an emotional attention task. The mean behavioral reliability across all four time points was poor (percentage of errors) to fair (RT) with even lower behavioral reliability for difference scores, in accordance with our analysis from age 14–16 (Vetter et al., 2017). Regarding fMRI, we could not use the initial difference contrast (Vetter et al., 2015) due to low reliability. This low reliability might be due to the expected ongoing behavioral development across adolescence. However, the contrast matching positive and matching negative versus neutral pictures had also low split-half reliabilities. Recent work demonstrated lower reliability of difference contrasts both in behavioral (Hedge et al., 2018) as well as in fMRI data (Elliott et al., 2020; Infantolino et al., 2018), also specificially in adolescent fMRI (Baranger et al., 2021). Infantolino et al. (2018) conclude that the low reliability in difference contrasts results of too much shared variance respectively too less unique variance for the single conditions. Thus, substracting leaves almost pure error variance which is uncorrelated between measurements and therefore unreliable per definition (Infantolino et al., 2018). For the final single contrasts 'matching negative/positive/neutral minus implicit baseline', the mean reliability across all four time points in the IFG was at least above fair in accordance with our previous study (Vetter et al., 2017). This amount of IFG reliability is also in line with an adult (Gee et al., 2015) and an adolescent study on emotional processing (Overgaauw et al., 2015). The latter found IFG activation at baseline correlated with activation 2 years later indicating some degree of reliability (Overgaauw et al., 2015). Further, the overall amount of reliability is similar as in a recent meta-analysis of reliability in short time intervals in adults (Elliott et al., 2020) and other adolescent studies examining the PFC (Baranger et al., 2021; Koolschijn et al., 2011; Ordaz et al., 2013; van den Bulk et al., 2013). The mean reliability across all four time points for amygdala activation was poor with an ICC of 0.2 (see supplement C) which is in accordance with an adolescent study (van den Bulk et al., 2013) and adult emotion studies (Gee et al., 2015; Johnstone et al., 2005; Plichta et al., 2012).

### *4.4. Limitations*

The current study has several strengths such as the four time points and the many total scans as well as the complete longitudinal design. As one limitation, we used a liberal 'fair' threshold for defining reliable clusters and the amygdala ROIs even did not pass this threshold. Future studies might test reliability within short time intervals in adults first before applying paradigms to developing age groups in longitudinal designs (Herting et al., 2018). An overall limitation of the current task is that matching abstract stimuli did not induce the expected attention capture effect, therefore our group developed an alternative task (Marxen et al., 2021). The unexpected findings, especially of the late increase in amygdala activation might be driven by participants > 20 years that contain fewer data sets. Our conclusions are limited to a specific age range of mid-adolescence until young adulthood. Future studies could cover the full spectrum of development such that the trajectories of top-down and bottom-up resources can be characterized from childhood to adulthood.

## 5. Conclusion

To conclude, top-down control seems to improve from age 14–22 as indicated by improving performance during a perceptual discrimination task on emotional attention. Additionally, participants showed an increase in IFG activity, indicating again the recruitment of cognitive control resources. An increase of amygdala activity at about age 22 for positive stimuli points towards a rather stable bottom-up system during adolescence and a later change in adulthood for positive stimuli, whereby the low reliability has to be considered. The study emphasizes the role of reliability for longitudinal fMRI studies and offers a relatively novel approach for the identification of ROIs. Results might help to better characterize healthy in contrast to dysfunctional development since these networks play a critical role in the manifestation of affective disorders. For future research on adolescent bottom-up and top-down development that target the current brain development models 1) there should be more longitudinal studies ideally using complete longitudinal designs; 2) reliability of the task and ROIs should be tested beforehand; 3) the development across late childhood to young (or mid-) adulthood should be targeted at with large samples; 4) connectivity of the top-down and bottom-up networks could be focused; and 5) the potential different top-down/bottom-up development depending on specific cognitive-emotional processes (using different emotion, reward or cognitive tasks) could be aimed at.

## Ethics approval

All participants and additionally for participants younger than 18 years one of their legal guardians provided written informed consent at each wave of data acquisition. The ethics committee of the Technische Universität Dresden approved the study (EK 235092007) and it was performed according to the Declaration of Helsinki.

## CRediT authorship contribution statement

**Nora C. Vetter:** Conceptualization; Writing – original draft, Formal analysis, Investigation, Funding acquisition. **Juliane H. Fröhner:** Writing – review & editing, Investigation, Visualization, Formal analysis. **Klara Hoffmann:** Formal analysis. **Lea L. Backhausen:** Formal analysis, Visualization. **Michael N. Smolka:** Conceptualization of project, Funding acquisition, Supervision, Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data Availability

The data that support the findings of this study are available upon reasonable request.

## Acknowledgements

This research was supported by Bundesministerium für Bildung und Forschung, Grant Numbers: 01EV0711, AERIAL 01EE1406A, 01EE1406B; Deutsche Forschungsgemeinschaft, DFG Grant Numbers: 178833530 (SFB 940), 290210763 (VE 892/2-1); Faculty of Medicine at the Technische Universität Dresden, MeDDrive Grant. We would like to thank Maximilian Mahr for his assistance in data analysis and especially the participants and their families for their enduring commitment.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.dcn.2022.101131.

## References

Baranger, D.A.A., Lindenmuth, M., Nance, M., Guyer, A.E., Keenan, K., Hipwell, A.E., Shaw, D.S., Forbes, E.E., 2021. The longitudinal stability of fMRI activation during reward processing in adolescents and young adults. NeuroImage 232, 117872. https://doi.org/10.1016/j.neuroimage.2021.117872.

Bennett, C.M., Miller, M.B., 2013. fMRI reliability: Influences of task and experimental design. Cogn., Affect. Behav. Neurosci. 13 (4), 690–702.

Bos, D.J., Dreyfuss, M., Tottenham, N., Hare, T.A., Galván, A., Casey, B.J., Jones, R.M., 2020. Distinct and similar patterns of emotional development in adolescents and young adults. Dev. Psychobiol. 62 (5), 591–599. https://doi.org/10.1002/dev.21942.

Carstensen, L.L., Mikels, J.A., 2005. At the intersection of emotion and cognition: aging and the positivity effect. Curr. Dir. Psychol. Sci. 14 (3), 117–121. https://doi.org/10.1111/j.0963-7214.2005.00348.x.

Casey, B.J., Jones, R.M., Hare, T.A., 2008. Adolesc. Brain 1124 (1), 111–126. https://doi.org/10.1196/annals.1440.010. https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1196/annals.1440.010.

Casey, B.J., Galván, A., Somerville, L.H., 2016. Beyond simple models of adolescence to an integrated circuit-based account: a commentary. Dev. Cogn. Neurosci. 17, 128–130. https://doi.org/10.1016/j.dcn.2015.12.006.

Casey, B.J., Heller, A.S., Gee, D.G., Cohen, A.O., 2019. Development of the emotional brain. Neurosci. Lett. 693, 29–34. https://doi.org/10.1016/j.neulet.2017.11.055.

Cohen, A.O., Breiner, K., Steinberg, L., Bonnie, R.J., Scott, E.S., Taylor-Thompson, K., Rudolph, M.D., Chein, J., Richeson, J.A., Heller, A.S., 2016. When is an adolescent an adult? Assessing cognitive control in emotional and nonemotional contexts. Psychol. Sci. 27 (4), 549–562.

Cohen-Gilbert, J.E., Thomas, K.M., 2013. Inhibitory control during emotional distraction across adolescence and early adulthood. Child Dev. 84 (6), 1954–1966.

Cope, L.M., Hardee, J.E., Martz, M.E., Zucker, R.A., Nichols, T.E., Heitzeg, M.M., 2020. Developmental maturation of inhibitory control circuitry in a high-risk sample: a longitudinal fMRI study. Dev. Cogn. Neurosci. 43, 100781 https://doi.org/10.1016/j.dcn.2020.100781.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3 (3), 201–215. https://doi.org/10.1038/nrn755.

Cromheeke, S., Mueller, S.C., 2014. Probing emotional influences on cognitive control: an ALE meta-analysis of cognition emotion interactions. Brain Struct. Funct. 219, 995–1008. https://doi.org/10.1007/s00429-013-0549-z.

Crone, E.A., Dahl, R.E., 2012. Understanding adolescence as a period of social-affective engagement and goal flexibility. Nat. Rev. Neurosci. 13, 636–650. https://doi.org/10.1038/nrn3313.

Decety, J., Michalska, K.J., Kinzler, K.D., 2012. The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. Cereb. Cortex 22 (1), 209–220. https://doi.org/10.1093/cercor/bhr111.

Deeley, Q., Daly, E.M., Azuma, R., Surguladze, S., Giampietro, V., Brammer, M.J., 2008. Changes in male brain responses to emotional faces from adolescence to middle age. Neuroimage 40, 389–397. https://doi.org/10.1016/j.neuroimage.2007.11.023.

van den Bulk, B.G., Koolschijn, P.C.M., Meens, P.H.F., van Lang, N.D.J., van der Wee, N.J.A., Rombouts, S.A., Vermeiren, R.R., Crone, E.A., 2013. How stable is activation in the amygdala and prefrontal cortex in adolescence? a study of emotional face processing across three measurements. Dev. Cogn. Neurosci. 4, 65–76.

Dolcos, F., Katsumi, Y., Moore, M., Berggren, N., de Gelder, B., Derakshan, N., Hamm, A.O., Koster, E.H.W., Ladouceur, C.D., Okon-Singer, H., Pegna, A.J., Richter, T., Schweizer, S., Van den Stock, J., Ventura-Bort, C., Weymar, M., Dolcos, S., 2020. Neural correlates of emotion-attention interactions: From perception, learning, and memory to social cognition, individual differences, and training interventions. Neurosci. Biobehav. Rev. 108, 559–601. https://doi.org/10.1016/j.neubiorev.2019.08.017.

Elliott, M.L., Knodt, A.R., Ireland, D., Morris, M.L., Poulton, R., Ramrakha, S., Sison, M.L., Moffitt, T.E., Caspi, A., Hariri, A.R., 2020. What Is the test-retest reliability of common task-functional mri measures? new empirical evidence and a meta-analysis. Psychol. Sci. 31 (7), 792–806. https://doi.org/10.1177/0956797620916786.

Fröhner, J.H., Teckentrup, V., Smolka, M.N., Kroemer, N.B., 2019. Addressing the reliability fallacy in fMRI: similar group effects may arise from unreliable individual effects. NeuroImage 195, 174–189. https://doi.org/10.1016/j.neuroimage.2019.03.053.

Gee, D.G., McEwen, S.C., Forsyth, J.K., Haut, K.M., Bearden, C.E., Addington, J., Goodyear, B., Cadenhead, K.S., Mirzakhanian, H., Cornblatt, B.A., 2015. Reliability of an fMRI paradigm for emotional processing in a multisite longitudinal study. Hum. Brain Mapp. 36 (7), 2558–2579.

Gläscher, J., 2009. Visualization of group inference data in functional neuroimaging. Neuroinformatics 7 (1), 73–82.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., et al., 2004. Dynamic mapping of human cortical development during childhoodthrough early adulthood. Proc. Natl. Acad. Sci. U.S.A. 101, 8174–8179.

Grose-Fifer, J., Rodrigues, A., Hoover, S., Zottoli, T., 2013. Attentional capture by emotional faces in adolescence. Adv. Cogn. Psychol. 9, 81–91. https://doi.org/10.2478/v10053-008-0134-9.

Guyer, A.E., Monk, C.S., McClure-Tone, E.B., Nelson, E.E., Roberson-Nay, R., Adler, A.D., Fromm, S.J., Leibenluft, E., Pine, D.S., Ernst, M., 2008. A developmental examination of amygdala response to facial expressions. J. Cogn. Neurosci. 20 (9), 1565–1582.

Hare, T.A., Tottenham, N., Galvan, A., Voss, H.U., Glover, G.H., Casey, B.J., 2008. Biological substrates of emotional reactivity and regulation in adolescence during an emotional go-nogo task. Biol. Psychiatry 63 (10), 927–934. https://doi.org/10.1016/j.biopsych.2008.03.015015.

Hedge, C., Powell, G., Sumner, P., 2018. The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. Behav. Res. Methods 50 (3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1.

Henson, R., 2006. In: R.S.J.Frackowiak, K.J.F., Function, C.F.B. (Eds.), Analysis of fMRI Timeseries: Linear Time-Invariant Models, Event-related fMRI and Optimal Experimental Design. Elsevier Books.

Herting, M.M., Gautam, P., Chen, Z., Mezher, A., Vetter, N.C., 2018. Test-retest reliability of longitudinal task-based fMRI: Implications for developmental studies. Dev. Cogn. Neurosci. 33, 17–26. https://doi.org/10.1016/j.dcn.2017.07.001.

Hibell, B., Andersson, B., Bjarnason, T., Kokkevi, A., Morgan, M., & Narusk, A. (1997). The 1995 ESPAD report. Alcohol Drug Use Stud. In., 26,.

Infantolino, Z.P., Luking, K.R., Sauder, C.L., Curtin, J.J., Hajcak, G., 2018. Robust is not necessarily reliable: from within-subjects fMRI contrasts to between-subjects comparisons. NeuroImage 173, 146–152. https://doi.org/10.1016/j.neuroimage.2018.02.024.

Johnstone, T., Somerville, L.H., Alexander, A.L., Oakes, T.R., Davidson, R.J., Kalin, N.H., Whalen, P.J., 2005. Stability of amygdala BOLD response to fearful faces over multiple scan sessions. NeuroImage 25 (4), 1112–1123.

Koolschijn, P.C.M.P., Schel, M. a, de Rooij, M., Rombouts, S. a R.B., Crone, E.A., 2011. A three-year longitudinal functional magnetic resonance imaging study of performance monitoring and test-retest reliability from childhood to early adulthood. J. Neurosci.: Off. J. Soc. Neurosci. 31 (11), 4204–4212.

Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T., 2000. Automated Talairach Atlas labels for functional brain mapping. Hum. Brain Mapp. 10 (3), 120–131. https://doi.org/10.1002/1097-0193(200007)10:3<120::AID-HBM30>3.0.CO;2-8.

Lang, P.J., Bradley, Cuthberg, 2005. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical Report.*

Maldjian, J.A., Laurienti, P.J., Burdette, J.H., 2004. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. NeuroImage 21 (1), 450–455. https://doi.org/10.1016/j.neuroimage.2003.09.032.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. NeuroImage 19 (3), 1233–1239. https://doi.org/10.1016/s1053-8119(03)00169-1.

Marxen, M., Jacob, M.J., Hellrung, L., Riedel, P., Smolka, M.N., 2021. Questioning the role of amygdala and insula in an attentional capture by emotional stimuli task. Hum. Brain Mapp. 42 (5), 1257–1267. https://doi.org/10.1002/hbm.25290.

Mills, K.L., Siegmund, K.D., Tamnes, C.K., Ferschmann, L., Wierenga, L.M., Bos, M.G.N., Luna, B., Li, C., Herting, M.M., 2021. Inter-individual variability in structural brain development from late childhood to young adulthood. NeuroImage 242, 118450. https://doi.org/10.1016/j.neuroimage.2021.118450.

Monk, C.S., McClure, E.B., Nelson, E.E., Zarahn, E., Bilder, R.M., Leibenluft, E., Charney, D.S., Ernst, M., Pine, D.S., 2003. Adolescent immaturity in attention-related brain engagement to emotional facial expressions. NeuroImage 20 (1), 420–428.

Narvacan, K., Treit, S., Camicioli, R., Martin, W., Beaulieu, C., 2017. Evolution of deep gray matter volume across the human lifespan. Hum. Brain Mapp. 38 (8), 3771–3790. https://doi.org/10.1002/hbm.23604.

Ordaz, S.J., Foran, W., Velanova, K., Luna, B., 2013. Longitudinal growth curves of brain function underlying inhibitory control through adolescence. J. Neurosci.: Off. J. Soc. Neurosci. 33 (46), 18109–18124.

Overgaauw, S., van Duijvenvoorde, A.C.K., Gunther Moor, B., Crone, E.A., 2015. A longitudinal analysis of neural regions involved in reading the mind in the eyes. Soc. Cogn. Affect. Neurosci. 10 (5), 619–627.

Paulsen, D.J., Hallquist, M.N., Geier, C.F., Luna, B., 2015. Effects of incentives, age, and behavior on brain activation during inhibitory control: a longitudinal fMRI study. Dev. Cogn. Neurosci. 11, 105–115.

Paus, T., Keshavan, M., Giedd, J.N., 2008. Why do many psychiatric disorders emerge during adolescence? Nat. Rev. Neurosci. 9, 947–957. https://doi.org/10.1038/nrn2513.

Petermann, F., Petermann, U., 2007. Hamburg Wechsler Intelligenztest für Kinder –IV (HAWIK-IV). Huber.

Pfeifer, J.H., Masten, C.L., Moore III, W.E., Oswald, T.M., Mazziotta, J.C., Iacoboni, M., Dapretto, M., 2011. Entering adolescence: resistance to peer influence, risky behavior, and neural changes in emotion reactivity. Neuron 69 (5), 1029–1036.

Pilhatsch, M., Vetter, N.C., Hübner, T., Ripke, S., Müller, K.U., Marxen, M., Rodehacke, S., Mennigen, E., Schmidt, D., Kroemer, N.B., Smolka, M.N., 2014. Amygdala-function perturbations in healthy mid-adolescents with familial liability for depression. J. Am. Acad. Child Adolesc. Psychiatry Vol. 53 (Issue 5), 559–568.

Plichta, M.M., Schwarz, A.J., Grimm, O., Morgen, K., Mier, D., Haddad, L., Meyer-Lindenberg, A., 2012. Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery. NeuroImage 60 (3), 1746–1758.

Ripke, S., Hübner, T., Mennigen, E., Müller, K.U., Rodehacke, S., Schmidt, D., 2012. Reward processing and intertemporal decision making in adults and adolescents: The role of impulsivity and decision consistency. Brain. Res. 1478, 36–47.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86 (2), 420–428.

Shulman, E.P., Smith, A.R., Silva, K., Icenogle, G., Duell, N., Chein, J., Steinberg, L., 2016. The dual systems model: Review, reappraisal, and reaffirmation. Dev. Cogn. Neurosci. 17, 103–117. https://doi.org/10.1016/j.dcn.2015.12.010.

Spielberger, C.D., 1983. Manual for the State-Trait Anxiety Inventory STAI (Form Y) ("Self-Evaluation Questionnaire"). 〈http://ubir.buffalo.edu/xmlui/handle/10477/1873〉.

Tamnes, C.K., Herting, M.M., Goddings, A.-L., Meuwese, R., Blakemore, S.-J., Dahl, R.E., Güroğlu, B., Raznahan, A., Sowell, E.R., Crone, E.A., Mills, K.L., 2017. Development of the cerebral cortex across adolescence: A multisample study of interrelated longitudinal changes in cortical volume, surface area and thickness. J. Neurosci. 3302–3316. https://doi.org/10.1523/JNEUROSCI.3302-16.2017.

Vasa, R.A., Pine, D.S., Thorn, J.M., Nelson, T.E., Spinelli, S., Nelson, E., Maheu, F.S., Ernst, M., Bruck, M., Mostofsky, S.H., 2011. Enhanced right amygdala activity in adolescents during encoding of positively valenced pictures. Dev. Cogn. Neurosci. 1 (1), 88–99.

Vetter, N.C., Altgassen, M., Phillips, L., Mahy, C.E.V., Kliegel, M., 2013p. Development of affective theory of mind across adolescence: Disentangling the role of executive functions. Dev. Neuropsychol. 38 (2), 114–125. https://doi.org/10.1080/87565641.2012.733786.

Vetter, N.C., Weigelt, S., Döhnel, K., Smolka, M.N., Kliegel, M., 2014. Ongoing neural development of affective theory of mind in adolescence. Soc. Cogn. Affect. Neurosci. 9 (7), 1022–1029.

Vetter, N.C., Pilhatsch, M., Weigelt, S., Ripke, S., Smolka, M.N., 2015. Mid-adolescent neurocognitive development of ignoring and attending emotional stimuli. Dev. Cogn. Neurosci. Vol. 14, 23–31.

Vetter, N.C., Steding, J., Jurk, S., Ripke, S., Mennigen, E., Smolka, M.N., 2017. Reliability in adolescent fMRI within two years—a comparison of three tasks. Sci. Rep. Vol. 7 (1), 2287.

Vetter, N.C., Buse, J., Backhausen, L.L., Rubia, K., Smolka, M.N., Roessner, V., 2018. Anterior insula hyperactivation in ADHD when faced with distracting negative stimuli. Hum. Brain Mapp. Vol. 39 (7), 2972–2986.

Vuilleumier, P., Huang, Y.-M., 2009. Emotional attention: uncovering the mechanisms of affective biases in perception. Curr. Dir. Psychol. Sci. 18 (3), 148–152.

Vuilleumier, P., Armony, J.L., Driver, J., Dolan, R.J., 2001. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30 (3), 829–841. https://doi.org/10.1016/S0896-6273(01)00328-2.

Wood, S.N., 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. J. Am. Stat. Assoc. 99 (467), 673–686. https://doi.org/10.1198/016214504000000980.

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC. ⟨https://doi.org/10.1201/9781420010404⟩.