The adolescent brain : unraveling the neural mechanisms of cognitive and affective development

Peters, S.

The Adolescent Brain

Unraveling the neural mechanisms
of cognitive and affective development

Cover design by: Vanessa Mendozzi
Printed by: GVO drukkers & vormgevers B.V.
ISBN: 978-90-6464-955-4

The Adolescent Brain

Unraveling the neural mechanisms of cognitive and affective development

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 27 januari 2016
klokke 16.15

door

Sabine Peters
geboren te Leiderdorp
in 1987

Promotor:
Prof.dr. E.A. Crone

Co-promotors:
Dr. P.M.C.P. Koolschijn (Universiteit van Amsterdam)
Dr. A.C.K. Van Duijvenvoorde

Promotiecommissie:
Prof.dr. P.W. Van den Broek
Prof.dr. I.H.A. Franken (Erasmus Universiteit Rotterdam)
Dr. M. Huizinga (Vrije Universiteit Amsterdam)

Contents

Chapter 1

General Introduction

Adolescence

Adolescence is an important developmental stage marking the transition from childhood to adulthood. The exact time period of adolescence is not well-defined, but it is generally thought to begin with the onset of puberty (around age 9-12 years; Marshall & Tanner, 1969, 1970). Puberty is characterized by a sudden surge in sex hormones, most markedly the hormone testosterone for boys and estradiol for girls. This surge in hormones is associated with the development of secondary sex characteristics, such as body hair growth and voice lowering for boys, and breast development and the first menstruation for girls (Petersen, Crockett, Richards, & Boxer, 1988). The end of adolescence, i.e. the entering of 'adulthood', is less clear and seems to be mostly culturally defined. For instance, in Western cultures adulthood is generally thought to begin when a child becomes relatively independent from parents (Choudhury, 2010).

Aside from drastic changes in physical characteristics, adolescence is also known as a period of marked behavioral changes. During adolescence, there is a gradual development of cognitive control or 'executive functions', i.e., the ability to control and regulate one's own actions, but at the same time, adolescents show increased sensitivity to affective and social rewards (Galvan, 2013). For instance, adolescents are prone to increased impulsivity and risk taking behaviors, such as substance use, reckless driving, unprotected sex, and violent behaviors (Steinberg et al., 2008). Health-risk behaviors are the leading cause of mortality amongst persons 10-24 years old in the United States (Kann et al., 2014; Sells & Blum, 1996). The increased incidence of risk taking in adolescence and the associated heightened mortality rate is a problem for society, and a growing amount of research attention is directed towards the mechanisms behind cognitive control and risk taking in adolescence. The goal of this thesis was to investigate both cognitive and affective aspects of development in adolescence, combining the use of neuroimaging methods, behavioral measures and hormonal assessments.

Brain development in adolescence

An increasing amount of research on adolescent cognitive control and risk taking has focused on the dynamic changes in brain development during this period. One of the main methods to examine brain development is by using magnetic resonance imaging (MRI). With MRI scanners, it is possible to create images of an individuals' brain in a safe and noninvasive manner, thus making it suitable for investigating healthy and underage participants. Although MRI provides images of the structure of the brain, the emergence of functional MRI (fMRI) has ensured that we can also study the functioning of the brain, i.e. which brain areas are engaged while performing a certain task. In addition, functional connectivity between regions can for instance be studied using resting state fMRI, to determine which regions are interconnected when a person is at rest.

Several key studies have shown that the brain continues to develop for a longer period than previously thought, with structural development continuing until around the early twenties (Giedd, 2004; Koolschijn & Crone, 2013). Structural brain development has been studied in both

gray matter (the brain cells, i.e. neurons) and white matter (the connections between neurons). Research has shown that white matter increases in a roughly linear fashion between childhood and adulthood (Giedd, 2004; Giorgio et al., 2010). Gray matter, on the other hand, follows an inverse U-shaped developmental trajectory, with cortical thickness increasing in childhood and decreasing in adolescence until the levels stabilize in adulthood (Giedd, 2004; Gogtay et al., 2004; Tamnes et al., 2010). Importantly, the development of gray matter does not occur at the same speed in each brain region. Development appears to be slowest in the prefrontal and parietal cortex (Giedd et al., 2009), the regions that are especially associated with cognitive control functions (Diamond, 2013; Niendam et al., 2012).

Models of brain development in adolescence

The results from these neuroimaging studies have been used to explain both the increasing capacity for cognitive control from childhood to adulthood, as well as the adolescent-specific rise in risk taking. These and other studies have led to the formulation of models for brain development that attempted to explain the increased incidence of risk taking behavior in adolescence. These models (e.g. Ernst, Pine, & Hardin, 2006; Somerville & Casey, 2010; Steinberg, 2008) have used different names but here I used the term 'imbalance models' of adolescent development (see Figure 1).

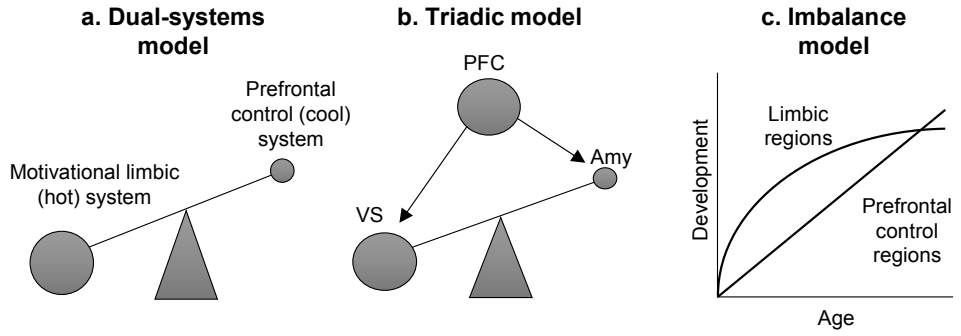


Figure 1: Adapted from Casey et al. (2015). Models of adolescent brain development. Abbreviations: Amy = amygdala, PFC = prefrontal cortex, VS = ventral striatum

The models are based on the findings that structural brain development in 'cognitive' regions (e.g. frontal and parietal cortex) is relatively delayed compared to 'affective' brain regions (e.g. striatum, amygdala). Specifically, it appears that affective brain regions mature relatively quickly, with some studies showing an adolescent peak in activity (Braams, Peters, Peper, Güroğlu, & Crone, 2014; Braams, van Duijvenvoorde, Peper, & Crone, 2015; Galvan et al., 2006; Van Leijenhorst et al., 2010). Cognitive brain regions in the frontal and parietal cortex on the other

hand, develop relatively more slowly with a gradual increase throughout adolescence. The hypothesis in imbalance models is that in adolescence, cognitive regions are not mature enough to keep relatively mature or even hyperactive affective regions under control, leading to increased risk taking behaviors. Naturally, this is a simplified account of brain development, but an intuitively appealing account that has inspired numerous other studies.

Towards new models of adolescent brain development

More recently, several authors (Casey, 2015; Crone & Dahl, 2012; Johnson, 2011; Pfeifer & Allen, 2012) have voiced the need for adjustments to imbalance accounts of adolescent development. One of the main reasons for this was that findings on development of functional brain activity were not always consistent with an imbalance model of adolescent development. For instance, although many studies found an increase with age in activity in the prefrontal cortex during cognitive tasks (as predicted by imbalance models), numerous studies also reported decreases in activity with increasing age. Both increases and decreases in activity with advancing age have been interpreted as reflecting increasing maturity (Pfeifer & Allen, 2012). That is, decreased activity in children and adolescents relative to adults has been interpreted as reflecting the immaturity of underlying brain structure and/or an inability to recruit these regions, whereas the studies showing increased activity in children and adolescents interpreted this as possibly less efficient and less focal activity compared to adults. It is a problem that this makes imbalance models virtually unfalsifiable (Pfeifer & Allen, 2012). In addition, many prior studies have collapsed across age groups (e.g. 8-12, 13-17 years), which increases power but decreases sensitivity to pinpoint exact moments of change, and very few studies have attempted to disentangle whether neural changes with development can be ascribed to age, or rather to differences in task performance or strategy use.

Other authors therefore argued instead for a different model emphasizing an increased flexibility of recruitment of control regions in adolescence, depending on motivational salience (Crone & Dahl, 2012). In other words, the hypothesis is that children and adolescents are in fact capable of recruiting brain regions for cognitive control, but often under different circumstances than adults. For instance, when motivation for a certain task is high (e.g., learning to play a new computer game, planning a birthday party), frontoparietal regions can be recruited by adolescents, and under very salient circumstances possibly even more strongly than in adults. Support for the idea that frontoparietal recruitment is not only a matter of increasing age comes from several sources. First, studies on relatively complex forms of executive functions such as performance monitoring paradigms, which rely on multiple executive functions such as working memory, inhibition and switching, have shown that the frontoparietal network is not exactly 'offline' in younger children. Instead, it appears that this network is activated in different situations for children than in adults. That is, in a task in which participants are instructed to learn rules based on negative and positive feedback, children do show less activity in frontal and parie-

tal regions after receiving negative feedback compared to adults, consistent with the idea that the control network is still 'offline' in younger children. However, younger children showed more activity than adults in these same regions after receiving positive feedback (van den Bos, Güroğlu, van den Bulk, Rombouts, & Crone, 2009; van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008). This argues against a simple model of frontal immaturity in childhood and adolescence. In addition, several studies showed that frontoparietal activity is perhaps related more to performance (Booth et al., 2004; Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Koolschijn, Schel, de Rooij, Rombouts, & Crone, 2011) and strategy differences (Andersen, Visser, Crone, Koolschijn, & Raijmakers, 2014) rather than age per se. These findings also support the notion that the frontoparietal network can be recruited provided that performance and/or strategy use are more adult-like.

In new frameworks for adolescent brain development, several authors have argued that the field of developmental cognitive neuroscience should move towards a more comprehensive view of the developmental brain which takes into account the complex interactions between different brain regions (Casey, 2015; Johnson, 2011). One way to test this is by using connectivity analyses rather than focusing on brain activity in isolated regions. Assessing connectivity is especially important in the time period of adolescence, given the prediction derived from models of adolescent brain development that 'control regions' and 'affective regions' are unbalanced.

To date relatively few studies have tested the longitudinal development of the frontoparietal control network compared to a subcortical affective network in adolescents. In this thesis, I investigated both cognitive and affective aspects of development in a large sample of normally developing adolescents. The main questions addressed in this thesis were 1) how brain regions in the frontoparietal network develop through adolescence, and 2) how connections between affective and cognitive brain regions influence risk taking behavior.

Approach – The Braintime Study

The studies described in this thesis were all part of a larger study, the 'Braintime' project. In this project, the overarching goal was to investigate normative development of cognitive, social and affective domains in relation to brain development, hormones and genes. To this end, data was collected from 299 normally developing participants ranging from 8 to 25 years at the first time point, who were all contacted again two years later for a follow-up measurement. Of the initial 299 participants, 286 were willing to participate a second time and for 254 of them it was possible to collect a second fMRI scan (32 were excluded due to braces), resulting in a total number of 553 MRI scans collected at two time points.

Outline of current dissertation

In this thesis, I report the results from the Braintime study in which I investigated both cognitive and affective aspects of adolescent brain development. I examined this using a combination of

functional task-based MRI, resting state fMRI and structural MRI, as well as behavioral measures and hormonal assessments.

The first part of this thesis **(Chapters 2-7)** is devoted to cognitive aspects of adolescent development. I start in **Chapter 2** with an overview of the literature on the development of cognitive control and specifically cognitive flexibility in childhood and adolescence. This theoretical chapter puts forth several important questions that are addressed in this thesis. In **Chapter 3**, I present a novel experimental paradigm for a feedback learning task, which I validated in an adult sample. The task relies on multiple executive functions (inhibition, working memory and switching) and is thus a good measure to study neural reactions during cognitive control in general, and for neural reactions to positive and negative feedback in particular. The goal was to investigate whether brain regions in the frontoparietal network were mostly sensitive to the valence of feedback, or rather to the informative value of feedback. In **Chapter 4**, I describe a comprehensive study on the neural development of feedback learning in a large child and adolescent sample. In this study, my goal was to investigate normative age-related changes in neural reactions in the frontoparietal network to positive and negative feedback based on a cross-sectional sample. The question addressed in **Chapter 5** is whether age-related differences in neural activity can truly be ascribed to age or rather to differences in strategy use from childhood to early adulthood. This is an important question given that prior results on age-related changes in brain function may not actually be due to maturational processes, complicating the interpretation of earlier research. To test this question, hidden Markov models were used to detect strategies at an individual level. In **Chapter 6** I investigated longitudinal rather than cross-sectional patterns of change in the frontoparietal network. The same participants from Chapter 4 and 5 were followed-up two years later. With this paper, I assessed within-person changes and thus more accurately investigated developmental trajectories. In addition, I explored which factors contribute to developmental change over and beyond age: performance, working memory and/or cortical thickness. The final chapter in the section on cognitive development is **Chapter 7**. In this chapter, I tested whether performance and neural activity during the feedback learning task, were predictive of "real-world" learning as measured by school performance indices (reading and mathematics performance).

The next chapters **(Chapters 8-9)** concerned developmental changes in the affective domain, particularly risk taking behavior, and how this relates to brain development. Alcohol consumption was used as a real-life index of risk taking behavior in adolescence. In **Chapter 8**, I investigated the relationship between amygdala-prefrontal functional connectivity during resting state and alcohol use. It has been hypothesized that increased connectivity between subcortical and cortical regions protects against impulsive behaviors. In addition, I tested whether testosterone played a role in the relation between brain connectivity and alcohol use. Sex hormones such as testosterone rise quickly during adolescence and may explain the propensity to increased risk taking behaviors in this developmental period. Finally, in **Chapter 9**, a longitudinal follow-up

study investigating alcohol use and functional brain connectivity is described. My main goal was to investigate whether increased alcohol use follows from decreased connectivity between the amygdala and the orbitofrontal cortex (i.e., decreased connectivity influencing a person's propensity to consume alcohol), or whether instead, alcohol use precedes amygdala-orbitofrontal connectivity (i.e., a 'damaging' effect of alcohol consumption on connectivity).

Chapter 2

Development of cognitive flexibility: A literature review

This chapter is based on:

Peters, S., & Crone, E.A. (2014). Cognitive flexibility in childhood and adolescence. In J. A. Grange, & G. Houghton, Task Switching and Cognitive Control. Oxford University Press.

Abstract

Cognitive flexibility is an important aspect of executive functioning, which is essential for successful adaptation to changing environmental demands. Cognitive flexibility is a complex concept that has been measured with many different paradigms. Conceptually, a distinction can be made between instructed flexibility (measured with task switching paradigms) and adaptive flexibility (measured with performance monitoring paradigms). Both types of flexibility are still developing across childhood and adolescence. In this chapter, we describe the behavioral and neural development of these two types of cognitive flexibility. For instructed flexibility, results indicate that performance is generally better in adults compare to children. FMRI-studies point to increased neural activity in adults compared to children in frontal and parietal areas, the basal ganglia and the thalamus. Adaptive flexibility also shows an age-related increase in performance. With development, neural activity patterns show a specialization of frontal and parietal areas for processing different types of feedback, which carry different informative value for future behavior. This indicates that instructed and adaptive flexibility are separable processes with different developmental trajectories. Future studies should focus on the interaction between instructed and adaptive flexibility across development, and include longitudinal, genetic and network analyses in their methods.

Executive functions

An important characteristic of human cognition is that people are constantly looking for ways to improve themselves. Rather than sticking to the status quo, people use flexibility and creativity to successfully adapt to an ever-changing environment. For instance, children in schools have to get used to switching between different subjects for each class. Instead of continuing to speak French during Spanish class, children tend to switch quite easily between classes. In situations like these, most people will recognize when they need to change their approach to adapt to a new environment. For successful adaptation to a changing environment, executive functions are of paramount importance. Executive functions refer to the ability to behave in goal-directed actions in new and challenging situations, and to overcome automatic thoughts and behaviors (Garon, Bryson, & Smith, 2008). Executive functions are thought to consist of at least 3 sub processes: (1) working memory, (2) inhibition and (3) cognitive flexibility (Huizinga, Dolan, & van der Molen, 2006; Miyake et al., 2000), although the exact number of sub processes and their structure remains a matter of debate.

Cognitive flexibility

In this chapter we will focus on cognitive flexibility. Cognitive flexibility is defined as the ability to adapt behavior to changing environmental demands. It is therefore essential for adaptive cognition, creative problem solving and 'out of the box' thinking. Without cognitive flexibility, we would be stuck in rigid behavioral patterns and we would be having difficulty adapting to new situations. Cognitive flexibility is, because of its inherently adaptive nature, arguably the most important component of executive functioning, as well as the component that shows the most profound changes across child and adolescent development. However, cognitive flexibility cannot be seen as an entirely distinct executive function, because it relies heavily on two other executive functions: working memory and inhibition (Huizinga et al., 2006). Working memory is important for flexibility because it is necessary to remember your goals and keep relevant information on-line. Inhibition is important because for flexible behavior, automatic responses need to be inhibited in order to choose a response that is better suited to current demands.

Development of cognitive flexibility

Cognitive flexibility undergoes profound changes across child and adolescent development (Cragg & Chevalier, 2012). The development of cognitive flexibility is of considerable importance for successful school performance. For instance, cognitive flexibility scores are related to math scores and phonemic awareness in preschoolers (Blair & Razza, 2007) and task switching performance was found to be related to reading and math scores in a longitudinal study in preschool and primary school (Bull, Espy, & Wiebe, 2008). Furthermore, interventions aiming to improve executive functioning skills (including flexibility) improve school readiness in young children (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008). Additionally, impairments in cognitive

flexibility are associated with a number of psychiatric disorders, such as autism spectrum disorder (Corbett, Constantine, Hendren, Rocke, & Ozonoff, 2009), obsessive compulsive disorder (Chamberlain, Blackwell, Fineberg, Robbins, & Sahakian, 2005) and anorexia and bulimia nervosa (Tchanturia et al., 2012). Together, these studies highlight that cognitive flexibility is a cornerstone of cognitive development.

Measuring cognitive flexibility: Adaptive and instructed flexibility

Cognitive flexibility has been studied with many different experimental paradigms (Ionescu, 2012). Many researchers have argued for a better conceptual understanding of a complex and multifaceted construct such as cognitive flexibility (Cragg & Chevalier, 2012; Morton, 2010). Conceptually, a distinction can be made between instructed flexibility and adaptive flexibility. For instructed flexibility, individuals need to adapt behavior based on changing task rules, such as in task switching paradigms. Adaptive flexibility, on the other hand, requires an individual to infer rules based on feedback about prior behavior, such as in performance monitoring paradigms.

In this chapter, we describe the development of these two different types of cognitive flexibility: instructed flexibility and adaptive flexibility, by focusing on task switching and performance monitoring paradigms. Our focus will be on the development of school-aged children (6-17 years), a developmental group which can be directly compared with adults using the same paradigms (Cragg & Chevalier, 2012). Even though most of the research to date on the development of cognitive flexibility has focused on preschool years (Blair, Zelazo, & Greenberg, 2005), school-aged development is marked by continuing improvement of many types of cognitive flexibility. A better understanding of these developmental improvements will hopefully contribute to improving education and interventions.

In the search for a more mechanistic understanding of subprocesses involved in cognitive flexibility, researchers have adopted a new approach to complement behavioral research. That is to say, the processes involved in cognitive flexibility and their developmental trajectory have been investigated from a neurocognitive perspective. This approach has proven highly valuable in situations where there are no overt behavioral indices (such as in the case of feedback processing; Crone, Zanolie, Van Leijenhorst, Westenberg, & Rombouts, 2008) or when different subprocesses may result in the same behavioral outcome (Morton, 2010). Therefore, this chapter will focus on the development of both behavioral and neural processes involved in instructed and adaptive cognitive flexibility. We will first provide a brief overview of the neural mechanisms of cognitive flexibility in adults. Next, we describe several studies that investigated the neurocognitive development of cognitive flexibility. These studies are subdivided into instructed flexibility (measured with task switching paradigms) and adaptive flexibility (measured with performance monitoring paradigms).

Neural mechanisms of cognitive flexibility in adults

Using functional magnetic resonance imaging (fMRI), researchers have been able to identify the neural areas that are important for cognitive flexibility. Most studies in this area have focused on instructed flexibility, by using task switching paradigms. FMRI studies of task switching in adults have generally compared neural activity during switch trials with activation during non-switch trials. The underlying rationale is that this subtraction paradigm allows the researcher to isolate the activity that is specific for switch trials. With this approach, it was found that task switching activates an extensive brain network including activity mostly in the lateral and medial prefrontal cortex (PFC), but also in the parietal lobes, cerebellum and other subcortical areas (Monsell, 2003).

A recent meta-analysis analyzed data from different types of cognitive flexibility paradigms to investigate in more detail the specific contributions of these brain areas (Kim, Cilles, Johnson, & Gold, 2012). The main aim was to distinguish brain areas that are domain-general (activate across different types of flexibility) and which areas are specialized for different types of flexibility. Three different types of paradigms were distinguished by the authors: two types of instructed flexibility (perceptual attention switching and response switching) and adaptive flexibility (referred to as 'context-based switching' by the authors). In response switching paradigms, participants are trained to make certain stimulus-response mappings (e.g. press left for circles, right for squares). After a switch, the correct stimulus-response mapping is changed (e.g. press left for squares, right for circles). In attention switching paradigms, participants respond with a key press to different dimensions of the same stimuli. For instance, prior to the switch participants could be asked to respond based on the color of the stimuli. After a switch, participants still see the same stimuli, but now they have to attend to a new dimension (e.g., shape) and ignore the previously correct dimension. In adaptive flexibility paradigms, participants are required to switch between task rules or cognitive 'sets'. The need to switch is inferred from performance feedback.

Kim et al. (2012) found a widespread network for all three switching types that consisted of medial and lateral PFC, but also parietal, temporal and occipital areas, the caudate nucleus and the thalamus were found. This was consistent with previous meta-analyses which did not distinguish between different switching types (Buchsbaum, Greer, Chang, & Berman, 2005; Derrfuss, Brass, Neumann, & von Cramon, 2005; Wager, Jonides, & Reading, 2004). To investigate areas that are domain-general in nature, the researchers investigated which areas were activated for all three types of switching. Regions in the inferior frontal junction and posterior parietal cortex were found. The researchers therefore argue that that these areas are probably involved in representing and updating task sets, because these processes are essential in all three types of switching (Miyake et al., 2000). Subsequently, they examined which areas were specifically activated for the three switching paradigms. The results revealed that in addition to domain-general regions associated with switching, there were also differences in brain activity for the three switching paradigms. For attention switching, compared to the other types of switching, preferen-

tial activity was found in several caudal PFC regions, posterior parietal cortex and inferior temporal areas. For response switching, preferential activity was found in DLPFC, several frontal gyri and posterior parietal cortex. Adaptive flexibility tasks resulted in preferential activation in rostral PFC regions, middle frontal gyrus, cuneus, inferior temporal cortex, occipital areas and subcortical structures. In summary, the findings of this meta-analysis indicated that there are several overlapping regions that are important for different types of task switching, but there are also separate neural networks underlying different types of instructedF and adaptive flexibility. This illustrates that task switching is not a unitary process and highlights the importance of studying different types of switching.

In the next paragraphs, the behavioral and neural development of cognitive flexibility will be further described, based on the distinction between instructed and adaptive flexibility.
It is well-known that at birth, the structure of the brain is not yet fully developed. Both gray matter (consisting of neurons) and white matter (structural connections between areas) are still maturing. This has been investigated with post-mortem studies and, more recently, with structural neuroimaging techniques. During child and adolescent development, linear increases have been observed in white matter connectivity (Barnea-Goraly et al., 2005; Giedd et al., 1999). For gray matter development, a nonlinear pattern has been found (Casey, Giedd, & Thomas, 2000): For the first two postnatal years of development, gray matter volume increases. After that period, gray matter decreases again with different brain areas differing in the rate at which they mature. The lateral PFC is one of the last areas to develop in adolescence (Casey et al., 2000) with estimates that gray matter in the lateral PFC reaches adult levels in the early twenties (Giedd, 2004). Similar results were found for the parietal cortex and the caudate (Giedd, 2004). It therefore seems that the areas that are important for task switching in adults undergo prolonged structural development during childhood and adolescence. This observation, combined with the results of behavioral studies in which increased switch costs for children were found, led a number of researchers to hypothesize that the development of cognitive flexibility may be related to brain maturation. Several studies have investigated the neural development of instructed flexibility, using two versions of task switching paradigms (response switching and attention switching) that were previously briefly described in the section on neuroimaging studies in adults.

Development of instructed flexibility

Behavioral development of instructed flexibility

The task switching paradigm has proven to be a very valuable tool to investigate the development of instructed cognitive flexibility. Task switching paradigms have been developed in different varieties. For example, a distinction can be made between the alternating-runs variant of the task switching paradigm, in which participants are instructed to switch after every n^{th} trial, and the cue-based switching paradigm, in which a switch is indicated by a cue. In both paradigms, the

demand to switch between tasks can result in 'mixing costs' (increased reaction times when comparing a task block with switches to a task block without switches), and 'switch costs' (increased reaction times when comparing switch trials within a switch block to non-switch trials within a switch block) (Monsell, 2003).

Developmental comparisons have indicated higher switch costs for children compared to adults, both in alternating run paradigms (e.g. Huizinga et al., 2006) and cued switching paradigms (e.g. Crone, Bunge, van der Molen, & Ridderinkhof, 2006). Developmental improvements are most pronounced between ages 7 and 12, although improvements are still observed in adolescence (Huizinga et al., 2006). In a comprehensive review, Cragg and Chevalier (2012) summarized many different processes which may contribute to these developmental differences. First, they mention that developmental differences may result from an improvement in goal setting. For example, in the alternating run paradigm it is necessary to maintain the task set in mind (e.g., Task A, Task A, Task B, Task B, Task A, Task A) and in cued task switching it is necessary to 'translate' the cue into the correct task representation. The finding of higher mixing costs (i.e., trials in a switch block compared to trials in a non-switch block) in children relative to adults has been interpreted as developmental differences in goal setting (Chevalier & Blaye, 2009). Second, developmental differences may be related to overcoming the previously relevant rule. For example, Cepeda, Kramer, and Gonzalez de Sather (2001) showed that increasing the time between trials reduced switch costs in adults but not in children, suggesting that children experience larger interference from previous task set activation. Third, developmental differences may be related to interference from overlapping stimulus response mappings. For example, Crone et al. (2006) showed that it is relatively more difficult for children to switch tasks when the switch trial is associated with the same response hand as on the previous trial compared to when the switch trial is associated with the opposite response hand, most likely because they experience more interference from the previous task-response mapping. Taken together, different underlying processes can account for developmental changes in switching, but exactly how and when these processes take place is not well understood. One way to entangle these developmental processes is by studying the neural mechanisms involved in task switching.

Neurocognitive development of instructed flexibility

In response switching paradigms, participants are instructed to switch between different stimulus-response mappings. Rubia et al. (2006) used a response switching task in which adolescents (10-17 years, all males), and adults (20-43 years, all males) were presented with a grid divided into four squares, with either a horizontal or vertical bidirectional arrow in the middle. On each trial, a red dot appeared in one of the quadrants. When the arrow was horizontal, participants had to judge whether the dot was on the left or right side of the grid. When the arrow was vertical, participants had to judge if the dot was in the upper or lower half of the grid. Behavioral results revealed enhanced switch performance in adults compared to adolescents. When compar-

ing switch trials to non-switch trials, adults showed more activation than adolescents in right inferior PFC, left parietal cortex, anterior cingulate cortex and putamen. The same areas were also found to be related to age in a region of interest and whole-brain regression analysis, further confirming these findings. The most robust age correlations were found in the parietal cortex and basal ganglia, which the authors took to suggest that maturation of these areas is most important for the development of task switching capacity.

Researchers from the same laboratory performed a second study using a similar paradigm with a sample of 63 participants between 13-38 years old (38 males) (Christakou et al., 2009). The behavioral results showed that there were switch costs for both adults and adolescents, but no significant age differences were found. Despite the absence of behavioral differences, with increasing age there was more activity in the bilateral inferior frontal cortex, the anterior cingulate cortex, the caudate nucleus, putamen, thalamus and inferior parietal lobe during switching. This was similar to the findings reported in a previous study using only male participants (Rubia et al., 2006).

A third response switching study was performed by Crone et al. (2006). Participants (17 children 8-12 years, 20 adolescents 13-17 years, 20 adults 18-24 years) were taught to associate three different cues with a pre-specified stimulus-response mapping. Cues could be followed by bivalent stimuli (when depending on the rule, the same stimuli could be associated with different response mappings) or univalent stimuli (stimuli that were always associated with the same response). Switching performance was better for adults and adolescents than for children. The pre-SMA was activated in all groups for rule switching compared to repetition. Interestingly, children also recruited this region more for bivalent repetition trials than for univalent repetition trials, indicating this region is not only active for rule switching but also for rule representation. This suggests a less mature activity pattern in pre-SMA for children compared to adolescents and adults.

A potential confound in task switching studies is the performance difference that is often found between children and adults. Thomas et al. (2011) addressed this issue by using an algorithm that adjusts task difficulty to ensure a similar task performance for all age groups. Participants (21 adolescents 10-16 years and 21 adults 22-40 years) were accustomed to a certain stimulus-response mapping (press '1' for 'X', press '2' for 'O'). When cued with a background color change, they were required to inhibit the prepotent response and replace it with an alternate response (press '3'). Despite the fact that the task was designed to equalize accuracy, there were still significant differences in accuracy between adolescents and adults. The fMRI results indicated that when comparing switch to non-switch trials, adults showed more activity in frontal and parietal areas (bilateral precentral gyri, precuneus, inferior and superior parietal lobule, paracentral lobule). To summarize, the developmental fMRI studies investigating instructed flexibility with response switching paradigms generally find increased activation in frontal and parietal

areas as well as the basal ganglia in adults compared to children or adolescents during switch trials.

Developmental fMRI studies of attention switching

A different aspect of instructed flexibility is investigated with attention switching paradigms. In these tasks, participants are required to switch attention between different perceptual aspects of stimuli. Casey et al. (2004) were the first researchers to investigate attention switching in a developmental fMRI study. Adults (18-23 years, $N = 7$) and children (7-11 years, $N = 7$) performed a task in which they were instructed to judge which of three objects was unique. The unique attribute could be either color or shape, and was alternated on each trial without cueing. Performance on these switch blocks was compared with control blocks, in which the relevant attribute was always color or always shape. Adults performed better on this task in terms of accuracy and reaction time than children. However, no switch costs were found, as indicated by similar behavioral performance on the switch blocks versus the control blocks. A possible explanation is that the task does not require deliberate top-down attention switching, but the unique stimulus 'pops out', regardless of the dimension that was being attended to in the previous trial. Neural activity during switch trials compared to non-switch trials resulted in activation in the caudate nucleus in both adults and children. Adults showed more activation during switches than children in the superior frontal gyrus, superior parietal cortex and middle temporal gyrus. However, the task seems to require a minimal amount of cognitive control because the 'pop-out' effect of the unique stimulus is probably sufficient for participants to successfully perform the task.

Using a more traditional experimental paradigm for attention switching, Morton, Bosma and Ansari (2009) instructed participants (14 children 11-13 years and 13 adults 19-25 years) to perform the Dimensional Change Card Sorting task. In this task, participants were instructed to sort cards according to changing stimulus dimensions (e.g. red and blue boats and flowers that had to be sorted by either color or shape). Before each trial, participants were cued with the correct sorting rule. When comparing switch and non-switch trials, adults showed more activation in the left posterior parietal cortex and the right thalamus compared to children. Children showed more activation than adults in the right superior frontal sulcus. Behavioral performance was similar for children and adults, suggesting that the age differences could not be explained in terms of performance differences. The authors hypothesized that the parietal cortex is less efficient in children than in adults, leading to compensatory recruitment of the frontal sulcus area. The increased recruitment of superior parietal cortex with development concurs with the findings of Casey et al. (2004). An important difference is that Morton et al. (2009) found a frontoparietal network associated with switching, whereas this network was not found in the Casey et al. (2004) study. Possibly, this is due to the limited cognitive control demands in the experimental task of Casey et al. (2004).

The attention switching findings partly correspond to the results from response switching paradigms (Christakou et al., 2009; Crone, Bunge, et al., 2006; Rubia et al., 2006). That is to say, using both approaches, a frontoparietal/subcortical network associated with switching was found, that was generally more pronounced in adults than children. However, the exact loci of activation were slightly different, indicating that response switching and attention switching may rely on different underlying processes. In line with this hypothesis, similar findings were reported in the direct comparison of these two types of switching in the meta-analysis on adult studies by Kim et al. (2012). An unexpected result was that Morton et al. (2009) found increased activation in a frontal area in children relative to adults. Possibly, the dimensional task switching paradigm used in the Morton et al. (2009) study requires additional processes that are not present in response shifting. This may cause strategy differences between children and adults, which in turn can result in different patterns of neural activity. Children may, for instance, resort to more effortful strategies for updating working memory, which explains the increase in frontal activity. Together, these findings demonstrate that specific task demands can result in different patterns of neural activation across development. In the next paragraph, age differences in task specific demands will be further investigated in the development of adaptive flexibility.

Development of adaptive flexibility

Behavioral development of adaptive flexibility

A second type of cognitive flexibility is adaptive flexibility, which is switching based on performance feedback. Rather than switching based on an arbitrary cue, performance feedback indicates whether a switch is necessary in order to maintain optimal performance. This method of studying cognitive flexibility may more closely resemble real-life situations.

The task that has been the most widely used to study performance-based switching is the Wisconsin Card Sorting Task (WCST; Milner, 1963). In the WCST, participants are instructed to sort cards with pictures following different sorting rules (based on number, color and shape). After sorting, positive or negative performance feedback is provided, indicating whether the correct rule was applied. The rule can change without warning; therefore, the switch cue is the unexpected negative feedback after a certain number of correct sorts. When this happens, participants have to use positive and negative feedback to find the correct rule again. Performance on the WCST improves dramatically in school-aged children, and reaches adult levels around the age of 12 years (Crone, Ridderinkhof, Worm, Somsen, & van der Molen, 2004). Interestingly, some authors have suggested that performance of children is similar to that of patients with damage to the lateral prefrontal cortex (Chelune & Baer, 1986). Similar to prefrontal patients, children have difficulty switching to a new rule and perseverate in the old rule; even if feedback indicates this is no longer correct.

The WCST has been a promising starting point for experimental research into the development of adaptive cognitive flexibility (performance-based switching). The WCST is a more complex task than the simple task switching paradigm, but it may therefore also more adequately reflect the complex nature of flexibility in daily life. Several different subprocesses are necessary for successful performance-based switching. It is essential to investigate these different processes in isolation to come to a full understanding of the development of performance-based switching.

One of the complications in the study of performance-based switching is that positive and negative feedback can have different meanings in different stages of an experimental task. For instance, Barcelo and Knight (2002) analyzed error types in the WCST and distinguished efficient errors (errors during the search for a new rule) and random errors (mistakes in a series of correct responses, due to distraction or memory errors). Along the same vein, positive feedback can also have different meanings. Zanolie, Van Leijenhorst, Rombouts and Crone (2008) distinguished first positive feedback (the first positive feedback when the new correct rule is found) and correct responses (when participants are continuing to use the correct response). First positive feedback provides new information that is valuable for learning, whereas correct positive feedback merely confirms what the participant already knows. Because of these differences in informative value, it is important to distinguish different feedback types, since they may rely on different neural processes.

A second complication in the study of performance-based switching is that behavioral outcomes cannot be measured until the next trial occurs. FMRI provides an excellent tool to distinguish the processes underlying performance-based switching. An advantage of fMRI studies is that we can focus on the effects of learning signals (i.e., feedback) in the brain, instead of on the subsequent behavioral response. To date, several neuroimaging studies have investigated the neural underpinnings of feedback processing in adults.

The previously mentioned meta-analysis on different types of task switcshing in adults already indicated that there are both domain-general areas that are activated for all types of switching, as well as specific activations for adaptive flexibility (Kim et al., 2012). Most of the studies analyzed in this meta-analysis on adaptive flexibility used paradigms based on the WCST, but also included semantic classification tasks, letter matching tasks, visual detection tasks and intra-extradimensional switching tasks. The focus in these studies was usually on the differences in activity for switch and non-switch trials. Other studies have specifically investigated activations during the presentation of positive and negative performance feedback. This approach is also crucial for our understanding of adaptive flexibility because it allows researchers to investigate the neural reactions to an adaptive learning signal.

FMRI results in adults have indicated that for feedback processing, a network similar to the task switching network is activated, including DLPFC, medial PFC (particularly SMA/ACC), parietal areas and the caudate (Holroyd et al., 2004; van Veen, Holroyd, Cohen, Stenger, & Carter, 2004; Zanolie et al., 2008). These areas have typically been found when comparing activity for

negative feedback with activity for positive feedback. This contrast is similar to the comparison between switch and non-switch trials, because negative feedback by definition indicates a 'switch' in behavior is necessary for optimal performance. However, not all studies have replicated this negative feedback network. For instance, Nieuwenhuis, Slagter, von Geusau, Heslenfeld and Holroyd (2005) failed to find differential activity to negative feedback in the medial prefrontal cortex using a time-estimation paradigm. Possibly, both positive and negative feedback signals are informative for adaptive learning, especially when both can be used to improve performance. Further research is necessary to unravel the exact contributions of the cognitive control network to feedback processing. In the next section, an overview of the developmental fMRI studies on adaptive flexibility will be provided.

Neurocognitive development of adaptive flexibility

The first neuroimaging study to investigate adaptive flexibility in children used a WCST-like paradigm (Crone et al., 2008). The task was child-friendly and instead of the extra-dimensional switches that are required in the WCST, there were three different rules based on only one dimension (location). Participants (17 children 8-11y, 20 adolescents 14-15y and 20 adults 18-24y) were instructed to 'help a dog find its way home' (i.e., find the correct location out of four possible options). Different positive and negative feedback types were distinguished. Adults showed more activation in ACC after negative feedback which indicated a switch (and thus violated expectations) and in the DLPFC after error-related negative feedback. In contrast to adults, children did not differentiate between the different negative feedback types in the DLPFC and ACC. The results (see Figure 1) demonstrate that, with development, there is an age-related specialization in the ACC and DLPFC in their contribution to feedback monitoring. This led to the hypothesis that children have difficulty recognizing the informative value for learning in the different feedback types. Together these results indicate that over the course of development, ACC and DLPFC regions become more specialized for different types of feedback.

One of the difficulties with the Crone et al. (2008) study is that negative feedback was unexpected and therefore may have been more salient than positive feedback. Van Duijvenvoorde, Zanolie, Rombouts, Raijmakers and Crone (2008) examined this question by focusing on learning from positive and negative feedback. Participants (18 children 8-9y, 19 early adolescents 11-13y, 18 adults 18-25y) performed a task in which positive and negative feedback were used to find one of two correct sorting rules (sort by color or sort by shape). Two stimuli were always presented in pairs of trials: a guess trial and a repetition trial. On guess trials, the correct sorting rule (e.g. color or shape) was cued, and the participant had to infer the correct answer based on positive or negative feedback (with an expected value of 50% correct). On the subsequent trial, the participant was expected to apply the correct rule based on information retrieved from the guess trial. The behavioral results showed that adults were better at using performance feedback than adolescents, who in turn performed better than children. Additional-

ly, children performed relatively more inaccurately after negative feedback than after positive feedback. This valence difference was reflected in neural activation patterns. For adults, DLPFC and superior parietal cortex were more active after negative feedback, whereas in children (8-9-years), these areas were activated more after positive feedback. The adolescents did not differentiate between positive and negative feedback, suggesting the turning point was around this age (11-13-years). For both adolescents and adults, but not for children, the SMA was more active after negative than positive feedback. Thus, SMA involvement during feedback learning seems to reach adults levels before the DLPFC.

Finally, in a third paradigm feedback processing was studied using a probabilistic learning task. In such paradigms, feedback is correct on approximately 70-80 percent of the trials. These tasks are more complex than tasks that use determinative feedback (which is by definition 100 percent valid) because it requires attending to long term goals. In a study by van den Bos, Guroglu, van den Bulk, Rombouts and Crone (2009) participants (18 children 8-11y, 27 adolescents 13-16y and 22 adults 18-22y) were instructed to respond to one of two pictures. One picture resulted in positive feedback on 70-80 percent of the trials and the other picture resulted in positive feedback on 20-30 percent of the trials. Children, adolescents and adults performed equally well on the task. Interestingly, when exploring alternative rules (the stimuli that do not result in the highest probability of positive feedback), children showed more activity in DLPFC and dACC after positive feedback, whereas adults showed more activation in these areas after negative feedback. Because these differences were only found when exploring alternative rules, this suggests that the developmental difference is not simply a valence effect (van Duijvenvoorde et al., 2008) but the informative value for learning that shows a different neural developmental trajectory.

Taken together, the studies that investigated the development of adaptive flexibility have demonstrated that with increasing age, several brain areas in the cognitive control network become more specialized for different types of performance feedback. Activity in prefrontal areas (DLPFC, ACC) is more differentiated in adults for different types of positive and negative feedback, which suggests that with development, these areas become more sensitive to the informative value of feedback. The question for future research is how the development takes place in learning from positive to negative, and learning from non-informative to informative feedback.

Overall summary

The neural circuits underlying the development of instructed flexibility and adaptive flexibility were reviewed in this chapter. The development of instructed flexibility has been investigated with two types of task switching: response switching and attention switching. For both paradigms, better behavioral performance was generally found for adults compared to children (but see Christakou et al. (2009) and Morton et al. (2009) for exceptions). Concurrent with these behavioral differences, increased activity for adults compared to children for switching was found in

frontal and parietal areas, the basal ganglia and the thalamus. Morton et al. (2009) additionally found increased activity for children in the right superior frontal sulcus, suggesting possible strategy differences between adults and children. Second, for adaptive flexibility a different developmental pattern was found. Performance-monitoring paradigms have found an age-related specialization of frontal and parietal areas for processing different types of feedback (non-informative versus informative, positive versus negative). These findings indicate that instructed flexibility and adaptive flexibility are separable processes that both contribute to cognitive flexibility. How and when these processes interact is an important question for future research.

Future directions

As highlighted in this chapter, over the past decade several researchers have attempted to understand the neural development underlying the development of cognitive flexibility. Although the current body of literature has provided valuable insights into cognitive flexibility, several advancements can be made which will allow us to draw more firm conclusions about the development of cognitive flexibility. We recommend three directions for future research:

Longitudinal studies

To date, only one study has investigated the neurocognitive development of cognitive flexibility with a longitudinal design (Koolschijn et al., 2011). Included in this study on performance monitoring were the same participants that participated in the Crone et al (2008) study (N at follow up after 3.5 years: 10 adults, 12 adolescents and 10 children). The results indicated that the neural networks involved during the WCST-like task changed over the 3.5 year interval: more activity was found in DLPFC, superior parietal cortex and ACC. Intriguingly, these changes were strongly related to performance improvements, more so than to age per se. This suggests that in previous cross-sectional studies, which often found performance differences between age groups, performance differences can possibly partly explain neural differences between age groups.

The findings of this study highlight the importance of longitudinal designs in developmental psychology and developmental cognitive neuroscience. Longitudinal studies have important advantages over cross-sectional studies (where different age groups are compared as a proxy for development). With longitudinal designs, it is possible to truly investigate development within participants. In cross-sectional designs, within-subject development may be masked if between-subject variability in developmental trajectories is relatively large. Cohort effects (generational differences between age groups other than age) are not a confounding factor in longitudinal studies. Given the important behavioral improvements within individuals over time, longitudinal designs will be of critical importance for future research into the development of cognitive flexibility.

Network analyses

Several studies in adults have indicated that, for cognitive flexibility tasks, it is not only important *which* areas are activated, but also how these areas interact with each other (Madden et al., 2010; Stelzel, Basten, & Fiebach, 2011; van Schouwenburg, den Ouden, & Cools, 2010). Interaction between different brain areas can be described as *functional connectivity*. With functional connectivity analyses, it is possible to investigate temporal correlations between neural signals in different brain areas. It is assumed that a temporal correlation means that these regions work in concert. Research into the development of functional connectivity is an important future venture that undoubtedly will provide us with further insights into the neural development of cognitive flexibility (Morton, 2010; Stevens, 2009).

One developmental study that has investigated functional connectivity in adaptive flexibility was performed by Van den Bos, Cohen, Kahnt and Crone (2012). They re-analyzed the data from a study on probabilistic feedback learning (van den Bos et al., 2009). For all trials, the 'prediction error' was calculated using a simple reinforcement learning model. When feedback for a stimulus was better than expected, a positive prediction error was generated by the model for that stimulus, which in turn increases the decision weight for that stimulus (i.e., the chance that stimulus will be selected on the next trial is larger). When feedback was worse than expected, a negative prediction error was generated, resulting in a lower decision weight. The researchers found that the representation of the prediction error resulted in robust activation in the striatum, an effect that was similar across age groups. Interestingly, a functional connectivity analysis showed that with increasing age, the connectivity between activity in the striatum and medial PFC increased. The strength of this connection was also related to the tendency to change expectations after negative feedback. These findings demonstrate that focusing on network analyses can provide additional information over focusing on activity in single brain areas.

Genetics

Finally, recent studies have indicated that there are profound genetic influences on executive functioning. This has been found with both quantitative and molecular genetics studies (Morton, 2010). In quantitative genetic studies, family members who share different percentages of their genes are compared in executive functioning skills. For instance, monozygotic twins share 100 percent of their genes, whereas dizygotic twins and siblings share 50 percent of their genes. A comparison of executive functioning in the different family members can provide an estimate of the amount of variation that is explained by genes or environmental influences. With this approach, it was found that executive functions have a surprisingly high heritability factor (Friedman et al., 2008). This means that monozygotic twins who grow up in completely different environments, will still be very comparable in terms of their executive functioning.

Another approach to investigate genetic influences on executive functioning is to examine specific genetic variations in a particular gene. People who have one variation of the gene can

be compared in executive functioning skills to people with another variation of the gene. With this approach, genetic differences that influence dopamine functioning have already been found in adults for learning from errors (Klein et al., 2007) and task switching (Colzato, Waszak, Nieuwenhuis, Posthuma, & Hommel, 2010; van Holstein et al., 2011) and in children for feedback processing (Althaus et al., 2009). Although few studies have investigated genetic influences on the development of cognitive flexibility, the current body of literature provides strong indications that these influences are likely to be found in future studies.

Conclusion

Executive functions in general, and cognitive flexibility in particular, are important aspects of child development. Cognitive flexibility has been studied extensively in children, which has led to important insights on its developmental origins. Two main types of cognitive flexibility were described in this chapter: instructed and adaptive flexibility. Maturation of brain areas involved in cognitive flexibility tasks seems to play an important role in the development of both types of flexibility. Despite the great progress in this area of research, there are also still many open questions. Future studies should focus on combined approaches using different types of cognitive flexibility in the same individuals, to investigate whether they rely on similar neural mechanisms. Additionally, developmental theories have often been inferred from studies on cross-sectional samples. For truly developmental models, longitudinal studies are essential.

To further advance our knowledge, functional connectivity within the cognitive flexibility network should be investigated, as well as possible influences of genetic variations on the development of cognitive flexibility. We expect that, together, these new insights will allow us to better track the developmental trajectory of cognitive flexibility from childhood, through adolescence into adulthood. This knowledge will hopefully help to create better learning environments across childhood development.

Chapter 3

Neural activity for feedback learning in adults

This chapter is based on:

Peters, S., Koolschijn, P.M.C.P., Raijmakers, M.E.J., Schrijver, M.S., Overgaauw, S. & Crone, E.A. Feedback learning: Neural reactions to valence and informative value depend on individual differences in strategy use (submitted for publication, 2014).

Abstract

Learning from feedback is an important aspect of flexible behavior, but how neural regions contribute to complex feedback learning is not yet understood. Previous studies found that a widespread brain network is activated for feedback learning, including the dorsolateral prefrontal cortex (DLPFC), pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC) and superior parietal cortex (SPC). In this fMRI study, participants performed a multiple-trial learning task where a distinction was made between a rule learning phase and a rule application phase, and between positive and negative learning, so that we could investigate which areas are sensitive to learning value and valence of feedback. We studied feedback processing in a more complex learning environment, where both positive and negative feedback were similarly informative for learning. This way, we addressed the issue that in previous studies, positive and negative feedback often differed in informative value. Additionally, finite mixture modeling on task performance allowed us to distinguish two different latent learning strategies, which allowed us to further pinpoint individual differences in neural activation. The results indicated that the DLPFC, pre-SMA and SPC were all sensitive to learning value of feedback (feedback during the learning phase compared to feedback during the application phase), with all regions, except for left SPC, showing more activation after negative learning compared to positive learning. Additionally, individual differences in strategy use were related to the degree to which areas in the feedback learning network were recruited, such that high performers showed faster learning and increased neural activation to learning signals in general. Together, these findings highlight the need for a more mechanistic understanding of complex learning, for which a modeling approach of strategy types proved to be specifically valuable.

Introduction

One of the key elements of successful learning is the ability to use performance feedback in order to adjust future behavior. Feedback can be either positive, signaling a continuation of current behavior, or negative, indicating the need for a behavioral adjustment (Holroyd & Coles, 2002). Learning from both positive and negative feedback signals is essential for successful adaptation to a changing environment. Even though much progress is made in understanding neural responses to single-trial positive and negative feedback, much less is known about how these neural regions respond to learning in a complex environment across multiple trials. In this study, we investigated the neural basis of positive and negative feedback learning in a complex repeated learning task. Second, we employed hidden Markov models to categorize learners in strategy types, in order to relate neural activation after feedback to latent strategy differences.

Previous research has indicated that a widespread network of brain regions is involved in feedback processing, including the dorsal anterior cingulate cortex (dACC)/pre-supplementary motor area (pre-SMA) (Holroyd et al., 2004; Mars et al., 2005; Monchi, Petrides, Petre, Worsley, & Dagher, 2001; Ullsperger & von Cramon, 2003), (dorso)lateral prefrontal cortex (DLPFC) (Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; Konishi et al., 2002; Lie, Specht, Marshall, & Fink, 2006; van den Bos et al., 2009; van Duijvenvoorde et al., 2008; van Veen et al., 2004; Zanolie et al., 2008), caudate nucleus (Monchi et al., 2001; Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006) and superior parietal cortex (SPC) (van Duijvenvoorde et al., 2008; Zanolie et al., 2008). Many of these studies reported more activation after negative feedback compared to positive feedback. For example, a paradigm that has frequently been used to investigate performance monitoring and feedback learning across multiple trials is the Wisconsin Card Sorting Task (WCST; Milner, 1963). In the WCST, participants are instructed to sort cards according to different sorting rules (based on number, color or shape), followed by positive or negative performance feedback. In addition, the task encompasses occasional unexpected rule switches, signaled by negative feedback. These studies, and related switching studies, reported specifically more DLPFC and pre-SMA/ACC activity after negative than positive feedback (Dove et al., 2000; Konishi et al., 2002; Lie et al., 2006; Zanolie et al., 2008). Together, these studies led to the conclusion that this network is important for expectation violation and performance adjustment.

Despite these consistent findings, a second set of studies has reported more DLPFC activity after positive feedback (van den Bos et al., 2009; van Veen et al., 2004), which does not fit well with the hypothesis of either expectation violation or performance adjustment. For example, in a probabilistic learning task, the DLPFC and SPC were more active after positive feedback when participants were applying the correct rule, compared to negative feedback (van den Bos et al., 2009). One proposed explanation for these conflicting findings is that the DLPFC and SPC respond to informative value of feedback rather than valence. In probabilistic designs, negative

feedback after applying the correct rule does not indicate that a switch in behavior is necessary. The previous experimental paradigms thus differed in the informative value that was provided by positive and negative feedback; in the latter probabilistic paradigm (van den Bos et al., 2009) positive feedback could be more informative than negative feedback. One way to unravel these inconsistent findings is by studying feedback processing in a more complex but deterministic learning environment, in which both positive and negative feedback are important for learning but might differ in informative value between trials.

Prior studies have shown that collapsing performance and using simple performance indices as correlates for neural activation fails to take into account much of the variance in behavior. In previous behavioral studies it was found that individual differences in learning performance are partially due to categorically different strategies that participants apply (Johansen & Palmeri, 2002; Schmittmann, Visser, & Raijmakers, 2006; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). These strategies can be characterized based on the degree to which participants take into account the informative value of feedback. These prior studies hypothesized that neural engagement differs depending on the efficiency of strategy use (Schmittmann, van der Maas, & Raijmakers, 2012), and having a better index of these strategy groups may allow us to further understand how neural activation in DLPFC, pre-SMA and SPC is important for complex learning. Thus, there is a high need for a more mechanistic understanding of how individuals learn in complex environments; therefore, we made use of a modeling approach to dissociate low-strategy and high-strategy learners.

In the current study, we used a higher-order rule-learning paradigm, which distinguished different feedback types based on early and late learning. This idea relates to prior research in which early and late learning were compared (Brovelli, Laksiri, Nazarian, Meunier, & Boussaoud, 2008; Eliassen et al., 2012). For example, Eliassen et al. (2012) investigated which neural areas are sensitive to the learning value of feedback by comparing activity during feedback on first trials of a learning sequence with feedback during later trials. Results showed that several brain regions were uniquely activated by first trial feedback, including the striatum, DLPFC and parietal and temporal areas. Here, we developed a learning task in which participants had to categorize stimuli in one of three locations. This paradigm allowed us to study neural responses to the informative value of positive and negative feedback in a learning phase of the task. In addition, it allowed us to model behavior according to different strategies.

First, we predicted that neural responses in DLPFC, SPC, and pre-SMA/ACC to both positive and negative feedback would be larger when learning compared to when applying correct sorting rules. This hypothesis was based on a prior study that found that for novel higher-order rule-learning, the DLPFC and pre-SMA were more active for early than late learning, consistent with the idea that these areas are sensitive to informative value for learning (Boettiger & D'Esposito, 2005). Second, we addressed the question whether negative feedback would result in additional activation in DLPFC and pre-SMA/ACC during learning compared to positive feed-

back, because of their presumed role in expectation violation and performance adjustment (Holroyd et al., 2004). Third, we hypothesized that low-strategy learners would be less sensitive to the informative value of feedback, leading to less activation in the feedback learning network during learning, when compared to high-strategy learners.

Methods

Participants

Thirty-two healthy young adults were included in the study (16 women, 18-25 years old, *M* = 21.39, *SD* = 1.88). Participants were paid volunteers, recruited through local advertisements or through a university course credits system. All participants were right-handed and reported no history of neurological or psychiatric disorders or (past or current) use of psychotropic medication. Informed consent was obtained and the study was approved by the Internal Review Board at the Leiden University Medical Center. All anatomical scans were reviewed and cleared by a radiologist. One female participant (not included in the demographics above) was excluded from the analysis due to excessive movement in the MRI scanner (> 5mm). IQ was estimated with two subtests of the WAIS-IIIR (Similarities and Block Design). All estimated IQ-scores were within the normal range (*M* = 113.05, *SD* = 9.13, range = 100–130).

Feedback Learning Task

Participants performed a feedback learning task in the MRI scanner. They were presented with three empty boxes under which three possible stimuli were presented one by one, for multiple times in a random order (Figure 1).

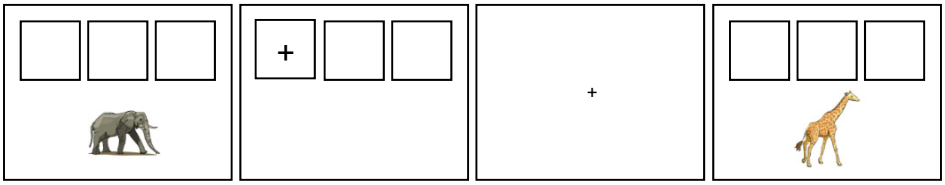


Figure 1: Sorting task, display of task sequence. Participants sorted stimuli on one of three locations and used performance feedback to find the correct location for three stimuli that all correspond to one of the boxes.

The participants were informed that each of the three stimuli belonged in one of the three boxes. By trial-and-error and using performance feedback, the correct location for all three stimuli could be found. Performance feedback was presented in the form of a plus-sign for positive feedback and a minus-sign for negative feedback. After applying the correct location twice for each stimulus, or after twelve trials, a new sequence with new stimuli was presented. New stimuli were chosen to prevent participants from trying to find a relation between sequences. In total, fifteen sequences were presented, which resulted in a maximum of 180 trials. Prior to scanning, participants practiced three sequences similar to the experimental task to ensure they understood the task. The task was divided in two blocks of eight and seven sequences each. The stimuli were presented in a pseudorandom order, with a maximum of two identical stimuli in a row. Stimuli were 250x250 pixels clip-art images found through Google Image search. Interstimulus intervals were jittered using OptSeq (Dale, 1999; see also http://surfer.nmr.mgh.harvard.edu/optseq), with intervals varying between 0-6 seconds.

Feedback types

For each of the three stimuli, a distinction was made between trials in the learning phase and the application phase. The learning phase was defined as those trials when participants had not yet found the correct location of the stimulus, and were guessing or reasoning to find the correct location. The application phase was when each stimulus was already sorted correctly at least once. Feedback in the learning phase was only scored as 'learning' based on subsequent responses for the same stimulus, i.e. choosing the same response (after positive feedback) or choosing a different response (after negative feedback). We were only interested in trials that actually resulted in learning; therefore trials in the learning phase that did not result in learning were excluded from further analysis (M = 4.2%, SD = 3.7% of the trials). The four feedback types were defined as follows:

Learning phase. (a) Positive learning: The first correct response for a stimulus, if followed by a correct response the next time that stimulus appears. (b) Negative learning: An incorrect response for a stimulus that had not yet been sorted correctly. The next time that stimulus appeared, the participant did not perseverate but chose another option.

Application phase. Choosing the correct location for a stimulus that was sorted correctly before.

Learning performance

Learning rate was calculated for each participant as a measure of learning performance. This was defined as the percentage of trials in the learning phase, where feedback was successfully used on the next trial.

Strategy performance

A second way in which performance relations were studied was by looking at the efficiency of strategy use. In the experimental task, several strategies differing in efficiency could be employed to learn the correct stimulus-location relations. The most optimal strategy accounts not only for feedback on the former responses of the same stimulus but also for other stimuli. This is based on two insights into the learning task: first, a location can only be associated with one stimulus (hence, choose a location that is not already correctly associated with another stimulus), and second, the probability of a correct response depends on the number of stimulus candidates for the chosen location (hence, choose a location for which negative feedback was given with another stimulus).

For each trial the efficiency of a response was defined as follows, in ascending order of the complexity of reasoning that is violated (note that for strategy analysis the non-learning trials were included): (a) Mistake: repeating a previously made error (no reasoning required to know this is an incorrect response) or making an error after an earlier correct response for the same stimulus. (b) Inefficient: when the location of a picture is known, the participant fails to deduce this location is not the correct location of another picture. (c) Suboptimal: when participants learned that a picture does not belong in a certain location (and if that is the only knowledge), the optimal decision is to place another stimulus in that location, to ensure a 50 percent chance of being correct. Failing to use this strategy is scored as a suboptimal decision. (d) Optimal: all other cases.

Strategy groups

We expected latent groups of participants that employed different learning strategies differing in the complexity of reasoning. We aimed to test: (a) how many strategies we could distinguish, and (b) which strategy each participant employed. Based on time series of trials that were coded in terms of the efficiency of responses (4-valued multinomial data), latent strategy performance groups were determined using finite mixtures of log-linear regression models (Leisch, 2004; McLachlan & Peel, 2000). More advanced hidden Markov models were applied as well (Visser, 2011), but these analyses resulted in less optimal models. Finite-mixture models assume that there are discrete components each of which produces series of responses according to a dedicated probability distribution, which characterizes a learning strategy. In this case, each component is a log-linear regression model, with trial number of the sub series as a possible predictor variable. The responses of each participant are assumed to be generated according to one component, but component membership is a latent variable, i.e., it cannot be observed directly from the data. To fit a mixture model to data, the number of components must be specified. The problem of selecting this parameter is known as 'model selection'. It can be resolved by using the minimum of the Bayesian information criterion (BIC), which implements a trade-off between log-likelihood of the model and the number of free parameters. The model with the lowest BIC is the optimal model.

We fitted mixture models with one, two or three components to our data; components were defined with and without trial number as a predictor variable. After establishing the optimal model, we assigned each individual to the component for which his/her data was most likely, according to the posterior probabilities of the data given the model (Visser, 2011). We used the depmixS4 software (Visser & Speekenbrink, 2010) for the subsequent analyses.

FMRI Data Acquisition

Scans were acquired with a standard whole-head coil on a Philips 3.0 Tesla MRI scanner at the Leiden University Medical Center. The functional scans were acquired using a T2*-weighted echo-planar imaging (EPI) during two functional runs with a variable number of volumes per subject, because the length of the task varied per subject. The first two volumes were discarded to allow for equilibration of T1 saturation effects (TR = 2.2 sec, TE = 30 ms, sequential acquisition, 38 slices of 2.75 mm, field of view 220 mm, 80x80 matrix, in-plane resolution 2.75 mm). A high-resolution 3D T1-FFE scan for anatomical reference was obtained (TR = 9.760 ms; TE = 4.59 ms, flip angle = 8 degrees, 140 slices, 0.875 × 0.875 × 1.2 mm^3 voxels, FOV = 224 × 168 × 177 mm^3). Head motion was restricted using a pillow and foam inserts that surrounded the head. The experimental task was projected on a screen that was viewed through a mirror.

FMRI Data Analysis

All data was analyzed with SPM8 (Wellcome Department of Cognitive Neurology, London). Images were corrected for differences in timing of slice acquisition and rigid body motion. Structural and functional volumes were spatially normalized to T1 templates. Translational movement parameters never exceeded 1 voxel (< 3 mm) in any direction for any participant or scan. The normalization algorithm used a 12-parameter affine transformation together with a nonlinear transformation involving cosine basis functions and resampled the volumes to 3 mm cubic voxels. Templates were based on the MNI305 stereotaxic space (Cocosco, Kollokian, Kwan, & Evans, 1997), an approximation of Talairach space (Talairach & Tourneaux, 1988). Functional volumes were spatially smoothed with an 8mm FWHM isotropic Gaussian kernel. Statistical analyses were performed on individual participant's data using the general linear model in SPM8. The fMRI time series data were modeled by a series of events convolved with a canonical hemodynamic response function. The feedback of each trial was modeled as an event of interest. The trial functions were used as covariates in a general linear model; along with a basic set of cosine functions that high-pass filtered the data, and a covariate for session effects. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used in pair-wise contrasts. The resulting contrast images, computed on a subject-by-subject basis, were submitted to group analyses. Task related responses were considered significant if they consisted of at least 10 contiguous voxels that exceeded a stringent threshold $p < .05$ (FWE-corrected).

Region-of-interest Analysis

Region-of-interest (ROI) analyses were performed with the Marsbar toolbox in SPM8 (Brett, Anton, Valabregue, & Poline, 2002). ROIs that spanned several functional brain regions (in our case, all ROIs) were subdivided by sequentially masking the functional ROI with each of several anatomical Marsbar ROIs. The contrast used to generate functional ROIs was based on the general contrast activation > fixation, (FDR-corrected, $p < .05$, 10 contiguous voxels) across all participants. For all ROI analyses, effects were considered significant at $\alpha = 0.008$ based on Bonferroni correction for six ROIs (unless reported otherwise).

Results

Behavioral results

Performance on the task was sufficient for all participants. On average, participants needed 135 (range 125-147, *SD* = 5.94) trials to complete the task (out of a maximum number of 180 trials). In general, participants had a high learning rate (*M* = 95.82%, *SD* = 3.71%).

Similar to learning rate, strategy performance was high (*M* = 94.84%, *SD* = 4.64% optimal trials). Strategy performance correlated strongly with learning rate ($r = 0.86$, $p < .001$). Suboptimal decisions (*M* = 1.57%, *SD* = 2.17%), inefficient decisions (*M* = 2.14%, *SD* = 1.81%) and mistakes (*M* = 1.45%, *SD* = 1.53%) were relatively rare. No sex differences were found for learning rate or strategy performance.

Strategy performance group

The optimal mixture model, i.e. the model with the lowest BIC, was one with two performance groups, referred to as the 'low-strategy learners' and the 'high-strategy learners' (see Table 1 for the fit statistics of the fitted models). The high-strategy learners have stable response probabilities over sub trials with very high probability of optimal responses. The low-strategy-learners have response probabilities that depend on sub trial number, such that suboptimal and inefficient responses mainly occur during the first few trials. Figure 2 shows the observed response probabilities for each performance group. The two groups did not differ in age, sex, or estimated IQ. As expected, low-strategy and high-strategy performers differed in learning rate ($t = 4.60$, $p < .001$), such that high-strategy learners had a higher learning rate than low-strategy learners, for both positive learning ($t = 2.54$, $p = .017$) and negative learning ($t = 4.68$, $p = .001$). To investigate possible differences in the relative contributions of positive and negative learning to the overall learning rate, the ratio of positive learning compared to negative learning rate was also calculated. This analysis showed that low-strategy performers learned relatively more from positive feedback than high-strategy performers ($t = 2.99$, $p = .014$) and high-strategy performers learned relatively more from negative feedback than low-strategy performers ($t = 3.59$, $p = .001$).

Table 1: Fit statistics of latent logistic regression models. The number of observations is 4297. Loglike is the Log-likelihood of the model; df is the number of freely estimated parameters in the model; BIC is the Bayesian Information Criterion. Models consist of 1, 2, or 3 components. Each component is a logistic regression function with an intercept (1) or an intercept + trial number as a predictor variable (1b). The asterisk indicates the parsimonious, best fitting model (with the lowest BIC).

Model	loglike	df	BIC
1	-1134.5	3	2294.1
1b	-1073.7	6	2197.6
1, 1	-1094.7	7	2248.0
1, 1b*	-1041.6	10	2166.8
1b, 1b	-1032.9	13	2174.5
1, 1, 1	-1088.7	11	2269.4
1, 1, 1b	-1041.4	14	2200.0

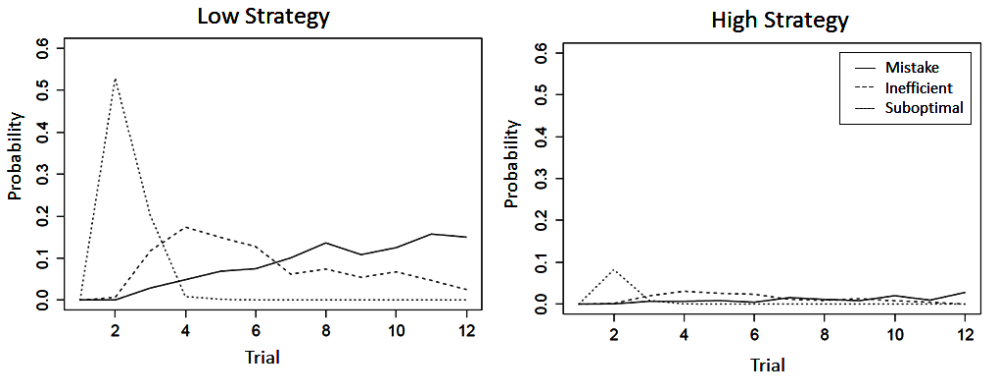


Figure 2: These figures display the probability of making a response at each trial for Low-strategy participants (left figure) and High-strategy participants (right figure). The different lines indicate different types of trials: mistakes, inefficient trials, sub-optimal trials. The remaining type of trials (optimal trials), are not shown.

Whole brain analyses

Learning versus application

To test which areas were responsive to learning value, we focused on the contrast learning (both positive and negative) > application. The regions that were more active for the learning > application contrast are presented in Figure 3, and included bilateral lateral PFC, pre-SMA and bilateral parietal cortex.

Positive learning vs. application

To test the hypothesis that the feedback learning network can also be activated by positive feedback, provided it is high in informative value for learning, the contrast positive learning > application was calculated. In both conditions positive feedback was provided, which ensures the only difference between conditions is informative value for learning. This contrast also resulted in robust activation in bilateral lateral PFC, pre-SMA and bilateral parietal cortex. As can be seen in Figure 3, there was high overlap in brain regions that were recruited for positive learning vs. application, and negative learning vs. application (see overlap in Figure 3). In the next section, positive and negative learning are tested against each other.

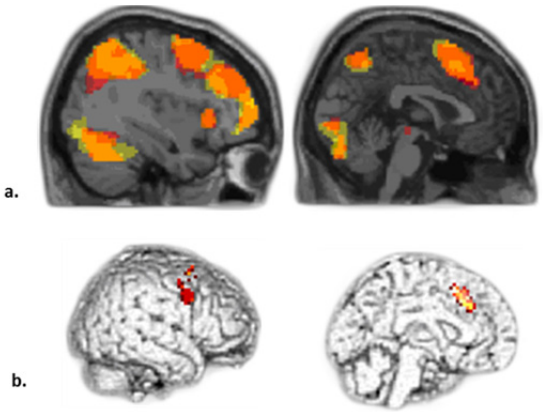


Figure 3: (a): Negative learning compared to application (red), positive learning compared to application (yellow), and overlap (orange). Results are FWE-corrected, $p < .05$, 10 contiguous voxels. The left image shows a lateral slice (x: 37 y: 8 z: 32) and the right image a medial slice (x: 1 y: 8 z: 12). (b): Negative learning > positive learning, FWE-corrected, $p < .05$, 10 contiguous voxels.

Positive learning vs. negative learning

To test which areas were sensitive to feedback valence, we focused on the contrast negative learning > positive learning and vice versa. We chose to compare positive and negative learning (instead of all negative > all positive feedback) to ensure that the only difference between the conditions was valence, and value for learning was comparable. The results demonstrated more activity for negative learning compared to positive learning in the right DLPFC, pre-SMA, anterior and middle cingulum. The reverse contrast, positive compared to negative learning resulted in activation in the bilateral supramarginal gyrus and bilateral superior temporal cortex (see Table 2).

Table 2: Areas that are activated for the contrast negative feedback > positive feedback during learning, FWE-corrected, 10 contiguous voxels.

Area of activation	x	y	z
R superior medial frontal gyrus	9	21	42
L superior medial frontal gyrus	-6	30	33
R superior frontal gyrus	21	3	54
R opercular inferior frontal gyrus	42	9	36
R precentral gyrus	36	0	48

Learning vs. application in low-strategy and high-strategy learners

Neural differences between high-strategy and low-strategy learners were investigated with a whole brain contrast comparing learning > application in the two groups. Uncorrected results are reported because the relatively small group of low-strategy learners did not survive FWE or FDR correction. A two-sample *t*-test (uncorrected, $p < .001$, 10 contiguous voxels) revealed that high-strategy learners showed more activation in right DLPFC and pre-SMA compared to low-strategy learners for the contrast learning > application (see Figure 4, and Table 3).

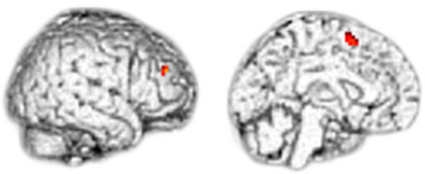


Figure 4: Areas that were more active for the contrast learning > application in high-strategy learners compared to low-strategy learners, $p < .001$, uncorrected, 10 contiguous voxels.

Table 3: Areas that are more activated for the contrast learning > application in high-strategy learners compared to low-strategy learners, p < .001, uncorrected, 10 contiguous voxels.

Area of activation	x	y	z
L opercular inferior frontal gyrus	-45	9	12
L opercular inferior frontal gyrus	-57	6	6
L inferior temporal gyrus	-39	-21	-21
L opercular inferior frontal gyrus	-33	6	21
L supplementary motor area	-3	12	57
R middle frontal gyrus	33	42	27

Region of Interest (ROI) analysis

The analyses above illustrate that a network of areas is involved in feedback learning. To further characterize sensitivity to valence and to investigate relations with performance, we performed a ROI analysis on five a priori defined regions based on the contrast all conditions > fixation. Two separate regions were found in the right DLPFC, resulting in a total of six ROIs (left DLPFC (center-of-mass: x: -38, y: 14, z: 48), right DLPFC (x: 44, y: 34, z: 32), right superior DLPFC (x: 35, y: 10, z: 55), pre-SMA (x: -8, y: 11, z: 54), left SPC (x: -30, y: -64, z: 57) and right SPC (x: 32, y: -62, z: 57). All significant results survived Bonferroni correction unless otherwise specified. Figure 5 shows a representative subset of the results.

Areas sensitive to learning value

All areas that were included in the ROI analysis were sensitive to learning value. This was defined as relatively more activity for feedback during learning than during application. That is to say, there were main effects of feedback type for all six ROIs (all *p*s < .001). Follow up paired-samples t-test comparisons indicated that all ROIs were more active during positive learning relative to application (all *p*s < .001) and negative learning relative to application (all *p*s < .001).

Areas sensitive to feedback valence

Next, we compared negative learning and positive learning to investigate which areas were sensitive to feedback valence in the learning phase.

- *DLPFC.* Right DLPFC and right superior DLPFC were both more active after negative learning than positive learning (right DLPFC: $t(31) = 4.57$, $p < .001$, right superior DLPFC ($t(31) = 4.51$, $p < .001$). The left DLPFC revealed a similar pattern ($t(31) = 2.73$ $p = .010$[1]). A direct comparison of valence (positive and negative learning) for the left and right DLPFC revealed that the difference between negative learning and positive learning was larger in right than in left DLPFC (Interaction Region x Condition; $F(1,31) = 10.41$, $p = .003$).

[1] This effect does not survive Bonferroni correction

- *Pre-SMA.* The pre-SMA was more active after negative learning than positive learning ($t(31) = 4.26$, $p < .001$). A direct comparison of region and valence showed that the right DLPFC and pre-SMA has a similar pattern for positive and negative learning (Interaction Region x Condition; $F(1,31) = .32$, $p = .58$). The left DLPFC and pre-SMA, however, differed in neural activity pattern ($F(1.31) = 16.59$, $p < .001$), such that the difference between negative learning and positive learning was larger in pre-SMA than in left DLPFC.

- *SPC.* The right SPC was more active after negative than after positive learning ($t(31) = 3.29$, $p = .003$). Left SPC did not differentiate between positive and negative learning ($t(31) = 1.44$, $p = .160$) A comparison of region and valence confirmed that left and right SPC demonstrate a different pattern (Interaction Region x Condition; $F(1,31) = 8.65$, $p = .006$).

Taken together, all a priori selected ROIs were more active for positive and negative learning compared to application, and all ROIS, except for left SPC (and left DLPFC without multiple comparisons corrections) were additionally more sensitive to negative learning than to positive learning.

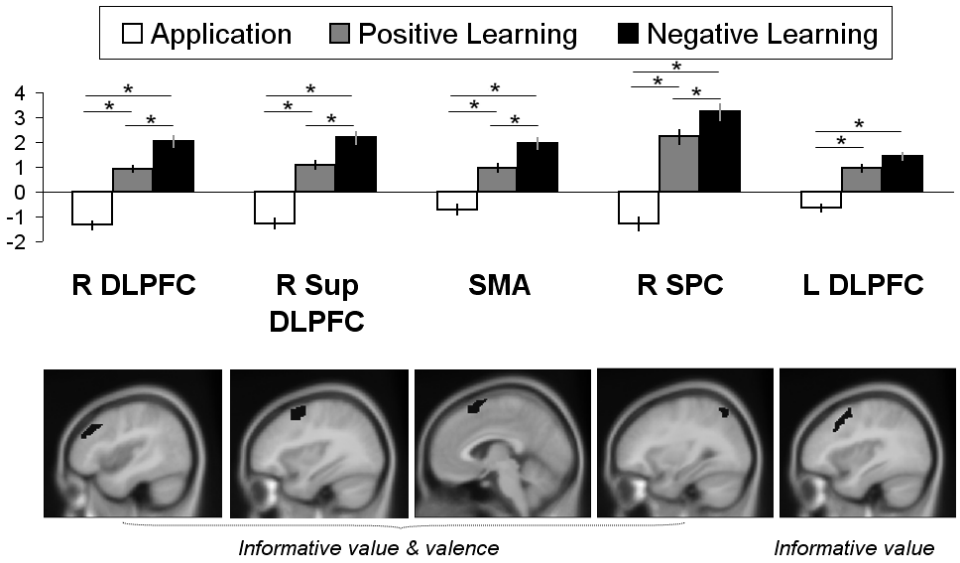


Figure 5. Region of Interest analysis Activation for application, positive learning, and negative learning. Abbreviations: L = left, R = right, DLPFC = dorsolateral prefrontal cortex, SMA = supplementary motor area, SPC = superior parietal cortex.

Correlations between neural activity and performance

Neural activity in the right superior DLPFC for learning > application was positively associated with learning rate in the task (r =.49, p = .005). A similar correlation was found for right SPC (r = .50, p = .004). Other regions did not show correlations with learning rate. For the strategy performance measures, we found that neural activation for learning > application correlated negatively with the percentage of mistakes in right SPC (r = -.36, p = .040[1]); and with the percentage of inefficient decisions in right superior DLPFC (r = -.46, p = 0.008), left DLPFC (r = -.37, p = .037[1]) and right SPC (r = -.49, p = .004). Taken together, right DLPFC and right SPC showed the most consistent relations to performance, but detailed strategy analyses confirmed that left DLPFC also contributed to efficient decisions in the task.

Discussion

Three main findings emerged from this study: (1) core areas in the feedback learning network (bilateral DLPFC, bilateral SPC and pre-SMA) were all sensitive to the learning value of feedback. Importantly, this effect was also found for positive feedback, but only when the positive feedback was informative for learning, (2) bilateral DLPFC, right SPC and pre-SMA were additionally more sensitive to negative feedback than positive feedback during learning, and (3) the degree to which areas in the feedback learning network were activated was dependent on learning performance and strategy use.

By comparing feedback during learning with feedback during application, it was found that a widespread network including DLPFC, pre-SMA and SPC was more active during learning. Our results are consistent with a prior single-trial learning study demonstrating a widespread network including the DLPFC and parietal areas which was sensitive to learning value (Eliassen et al., 2012). The current results also fit well with a prior study by Boettiger and D'Esposito (2005), which showed that complex learning of abstract stimuli that cannot easily be verbalized is associated with activation in the pre-SMA and DLPFC.

In prior studies, experimental designs were often developed such that positive feedback also revealed the rules for other stimuli, and with working memory demands that were often low (Dove et al., 2000; Lie et al., 2006; Zanolie et al., 2008). This study shows that positive feedback can result in similar neural activation as negative feedback when it is presented in a more complex learning setting that requires multiple trial learning. Prior studies have remained inconclusive about whether areas in the feedback learning network respond more to negative than to positive feedback. With stringent Bonferroni correction for multiple comparisons, three areas were found to be more active after negative feedback than positive feedback during learning (right DLPFC, pre-SMA and right SPC). This effect was also found in the left DLPFC, but only without multiple comparisons corrections.

The finding that the pre-SMA is more active after negative feedback is consistent with a number of other studies that have reported increased pre-SMA activation after negative feedback (Ozyurt, Rietze, & Thiel, 2012; Volz, Schubotz, & von Cramon, 2005). Second, the increased activation for negative versus positive feedback was larger for right DLPFC than for left DLPFC. These findings fit well with prior studies, in which it was suggested that the right DLPFC (together with the pre-SMA/ACC) is more important for performance adjustments than left DLPFC (Kerns et al., 2004), whereas the left DLPFC may be relatively more important for verbal working memory (Smith & Jonides, 1999; Wager & Smith, 2003). Notably, only right but not left SPFC showed sensitivity to valence. These findings suggest that the right hemisphere network in general is more important for performance adjustment.

The third question that was addressed in this study concerned individual differences in neural activity and learning performance. In the first set of general analyses, we found that learning rate was positively correlated with differential activation for learning compared to application in right DLPFC and, less strongly, in right SPC. Similar brain-behavior correlations in right DLPFC and right SPC were reported previously (Boettiger & D'Esposito, 2005; Koolschijn et al., 2011), and have been associated with a need for performance adjustment.

More importantly, we investigated whether differences in neural activity could be related to a categorization of strategy use. This method is valuable because it may detect latent differences in problem solving styles that cannot always be identified based on straightforward indices of behavior (Schmittmann et al., 2012). Indeed, participants differed in the strategies they employed during learning, and two different groups could be distinguished on the basis of strategy use. It was found that high-strategy learners showed more activity on a whole brain comparison of learning versus application in bilateral DLPFC and pre-SMA. Similarly, brain-behavior correlations were found for strategy performance in several regions within left and right DLPFC, demonstrating that especially these areas are related to individual differences in learning strategy (Andersen et al., 2014). Future studies should focus on deviant learning groups, and investigate whether the combined neuroimaging-behavioral modeling approach allows for the understanding of more variance in strategies.

In conclusion, with this study it was demonstrated that for complex feedback learning by reasoning, a network including DLPFC, pre-SMA and SPC is recruited that is sensitive to the learning value of both positive and negative feedback. Furthermore, all these regions, except for left SPFC, were additionally sensitive to valence. Detailed strategy categorizations confirmed that especially left and right DLPFC are associated with strategy differences in learning. These findings highlight the importance of distinguishing different subtypes of positive and negative feedback based on value for learning. In future research, the transition towards studying individual differences in strategy use and studying types of learning that more closely resemble real-life learning provides a promising way towards unraveling neural correlates for latent constructs of learning.

Chapter 4

Learning from positive and negative feedback across child and adolescent development

This chapter is based on:

Peters, S., Braams, B.R., Raijmakers, M.E.J., Koolschijn, P.C.M.P. & Crone, E.A (2014). The neural coding of feedback learning across child and adolescent development. Journal of Cognitive Neuroscience, 26, 1705-1720.

Abstract

The ability to learn from environmental cues is an important contributor to successful performance in a variety of settings, including school. Despite the progress in unraveling the neural correlates of cognitive control in childhood and adolescence, relatively little is known about how these brain regions contribute to learning. In this study, 268 participants aged 8-25 years performed a rule-learning task with performance feedback in a 3T MRI scanner. We examined the development of the frontoparietal network during feedback learning by exploring contributions of age and pubertal development. The prefrontal cortex showed more activation following negative compared to positive feedback with increasing age. In contrast, our data suggested that the parietal cortex demonstrated a switch from sensitivity to positive feedback in young children to negative feedback in adolescents and adults. These findings were interpreted in terms of separable contributions of the frontoparietal network in childhood, to more integrated functions in adulthood. Puberty (testosterone, estradiol and self-report) did not explain additional variance in neural activation patterns above age, suggesting that development of the frontoparietal network occurs relatively independently from hormonal development. This study presents novel insights into the development of learning, moving beyond a simple frontoparietal immaturity hypothesis.

Introduction

A crucial aspect of successful learning is the ability to process performance feedback and to adjust behavior on subsequent occasions, also referred to as feedback learning (Holroyd & Coles, 2002). Across development, children show marked improvements in feedback learning (Crone, Wendelken, Donohue, & Bunge, 2006; Eppinger, Mock, & Kray, 2009; Welsh, Pennington, & Groisser, 1991), which is evident from performance on neuropsychological tasks such as the Wisconsin Card Sorting Task and experimental switch tasks such as task-switching paradigms (Eppinger et al., 2009; van den Bos et al., 2009). Given the importance of learning in school settings, it is essential to unravel the mechanisms behind the development of feedback learning.

Prior studies of brain mechanisms for negative feedback processing in adults showed increased activation in dorsolateral prefrontal cortex (DLPFC), pre-supplementary motor area (pre-SMA)/anterior cingulate cortex (ACC), and superior parietal cortex (SPC) (Zanolie et al., 2008) after receiving negative feedback relative to positive feedback. In children, these areas showed less activation following negative feedback relative to positive feedback, compared to adults (Crone et al., 2008). This finding is consistent with evidence that these regions are still developing during adolescence, in terms of both structure (Raznahan et al., 2011; Shaw et al., 2008) and function (Klingberg, Forssberg, & Westerberg, 2002; Tamm, Menon, & Reiss, 2002).

Until recently, most theories on the development of cognitive control were based on the assumption that, with age, the frontoparietal network comes increasingly 'online' in a linear fashion (Bunge & Wright, 2007; Somerville, Jones, & Casey, 2010). However, recent reviews questioned whether cognitive development is associated with a linear increase in the frontoparietal network. (Crone & Dahl, 2012; Pfeifer & Allen, 2012). For example, prior studies on feedback learning indicated that children showed not only less activation after negative feedback compared to adults, but also more activation in the same regions in the frontoparietal network following positive feedback (van den Bos et al., 2009; van Duijvenvoorde et al., 2008). This suggests that there may be a qualitative shift with development in recruitment of brain areas for feedback learning.

The interpretation of prior studies is complicated due to differences in age group selection. Some studies compared 8-year-olds with 12-year-olds (van Duijvenvoorde et al., 2008) whereas others collapsed across 8-12-year-olds (Crone et al., 2008; Eppinger et al., 2009). Adolescent groups were selected from a wide age range from 13 to 16 years, which is problematic given that neural development continues in this period (van den Bos et al., 2009). Additionally, prior studies have not investigated the influence of pubertal maturation, which may explain additional variance in developmental change (Forbes & Dahl, 2010). For example, it was previously found that self-reported puberty scores explained additional variance over age in neural activity in a working memory study (Schweinsburg, Nagel, & Tapert, 2005). Others have suggested that brain regions supporting cognitive control develop relatively independent of pubertal influences

(Steinberg et al., 2008). A study including participants across the whole range of childhood and adolescence, focusing on both age and puberty effects, has not yet been performed.

In this study, we tested 268 participants between the ages of 8 and 25 to pinpoint developmental differences in feedback learning with increased precision. Participants performed a learning task in which stimuli were sorted at one of three possible locations, and choices were followed by negative or positive feedback. To control for informative value of feedback we distinguished between a learning phase and an application phase. These phases were based on a distinction between early feedback (informative for learning) and late feedback (uninformative for learning) (Eliassen et al., 2012). We formed two hypotheses: a) We predicted more activation in DLPFC, pre-SMA/ACC and SPC after feedback, both negative and positive, during learning compared to application of rules. We hypothesized that distinguishing between feedback informative for learning (learning phase) and uninformative feedback (application phase) is an ability that develops with age and contributes to performance. b) We predicted more activity in adults than children in DLPFC and SPC after negative feedback and more activity in children than adults in DLPFC and SPC after positive feedback (van Duijvenvoorde et al., 2008). In addition, we tested at what age activity patterns become adult-like and whether the transition between childhood and adulthood can be explained with both age and puberty as predictors (Casey, Jones, & Somerville, 2011).

Methods

Participants

The final sample (after exclusions) included 268 participants (138 female) between 8.01 and 25.95 years old (M = 14.22, SD = 3.63), who were recruited through local schools and advertisements. We selected age groups such that each age was represented by N > 20. Because of the relatively small number of 8- and 9-year-olds, they were combined into one age group, to ensure similar group sizes. See Table 1 for the final number of participants per age group and per sex. A chi square test indicated that the proportion of males to females was similar across age groups ($\chi 2(9) = 8.70$, $p = .465$). IQ was estimated with two subtests of the WAIS-III or WISC-III (Similarities and Block Design). All estimated IQ-scores were within the normal range (M = 110.25, SD = 10.62, range = 80-143). Adults received payment for participation and children and their parents received a present and travel reimbursement. None of the participants reported a history of neurological or psychiatric disorders or current use of psychotropic medication. This study was approved by the Internal Review Board at the Leiden University Medical Center and all participants (or participant's parents in the case of minors) provided written informed consent. All anatomical MRI scans were reviewed and cleared by a radiologist.

Table 1: Number of participants per group and per sex

	Female	Male	Total
8/9 years	20	9	29
10 years	11	12	23
11 years	13	14	27
12 years	19	11	30
13 years	16	20	36
14 years	10	17	27
15 years	10	11	21
16 years	11	9	20
17 years	12	11	23
18-25 years	16	16	32
Total	138	130	268

Exclusion criteria

Twenty-five participants were excluded from the analyses for the following reasons: Nineteen were excluded because movement in the MRI scanner exceeded 3.0 mm. Three participants were excluded because of damaged fMRI scans. Finally, three participants were excluded because they were extreme outliers on the total percentage of positive feedback (more than 3 times the inter-quartile range), thus indicating that they did not perform the task adequately.

Feedback Learning Task

Participants performed a child-friendly feedback learning task in the MRI scanner. On each trial, they saw a screen with three empty squares, under which a stimulus was presented (one of three possible stimuli) (see Figure 1). The participants were told that each stimulus belonged in one of the three squares and their task was to sort the stimuli into the correct square. Performance feedback was a plus-sign ('+') for a correct sort (positive feedback) and a minus-sign ('-') for an incorrect sort (negative feedback). The stimuli were presented in a pseudorandom order, with a maximum of two identical stimuli in a row. After 12 trials, or when the participant applied the correct location twice in total for each stimulus, the sequence ended and a new sequence was presented with three new stimuli. There were 15 sequences in total, resulting in a maximum of 180 trials. Before the MRI session, all participants practiced three sequences. The task was divided into two runs of eight and seven sequences respectively. Each trial started with a 500 ms fixation cross. Stimuli were presented for 2500 ms during which the response had to be given. Next, feedback was presented for 1000 ms. Inter-trial intervals were jittered based on OptSeq (Dale, 1999), with intervals varying between 0-6 seconds. The 500 ms fixation cross was presented for all trials and was not part of the ISI.

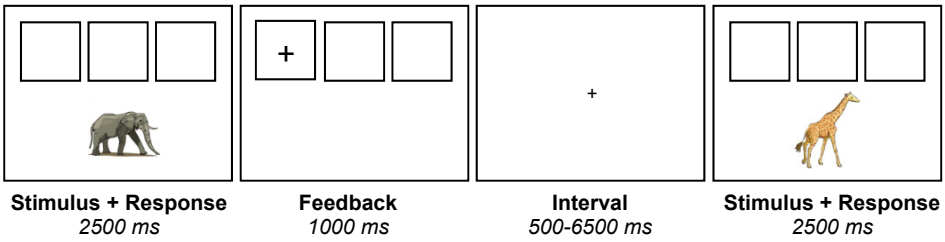


Figure 1: Display of task sequence.

Feedback types

For each sequence we made a distinction between a learning phase and an application phase. The learning phase was defined as those trials where participants had not yet found the correct location for the stimulus and were using trial-and-error or hypothesis testing to find the correct solution. The application phase was defined as those trials where a stimulus was sorted correctly on a preceding trial and which continued to be sorted correctly on subsequent trials. Some trials in the learning phase did not actually result in learning. That is, the feedback on these stimuli was not used correctly in a subsequent trial. These trials were excluded from the analysis (M = 3.65%, SD = 3.03% of all trials). Taken together, we defined the following three feedback types:

Learning phase. a) Positive Learning: refers to the sequence [CORRECT, correct] trials (upper case refers to the current trial, lower case refers to the preceding or subsequent stimulus): A first encountered correct feedback of a stimulus followed by a correct sort on the next trial of this stimulus. b) Negative Learning: refers to the sequence [ERROR, correct or error] trials: A first encountered error feedback of a stimulus followed by a choice for another location on the next trial of this stimulus.

Application phase. c) Application: refers to the sequence [correct, CORRECT] trials: A correct feedback of a stimulus preceded by a correct sort on an earlier trial for that stimulus.

Learning performance

To measure performance we calculated the 'learning rate' for each participant. This was defined as the percentage of trials in the learning phase where feedback was successfully used on the next trial. For this purpose we divided the number of trials scored as Positive Learning or Negative Learning, by the total number of trials during the learning phase.

Pubertal Development Measures

Pubertal Development Scale

To assess pubertal status, the Pubertal Development Scale (Petersen et al., 1988) was completed by participants under 18 (participants from 18-25 years were given the maximum score for pubertal development). Physical development was reported on five questions on a 4-point scale. Girls reported body growth, body hair, breast development, skin changes and menarche; boys reported body growth, body hair, facial hair, skin changes and voice changes. Fourteen girls and seven boys as well as the adults (18-25 years) did not fill out the PDS list, leaving the total number of included PDS scores at 220. Total PDS score was calculated as the mean of the five questions. Mean PDS score was 2.55 for girls (SD = 0.91, range 1.00-4.00), 2.29 for boys (SD = 0.82, range 1.00-4.00) and 2.42 for boys and girls combined (SD = 0.87, range 1.00-4.00).

Sex steroid levels

In addition to self-report measures of puberty, we collected saliva measures to extract testosterone and estradiol levels (de Water, Braams, Crone, & Peper, 2013; Peper, Mandl, et al., 2013). Boys and girls collected saliva by passive drool at home, directly after waking up. Girls using contraceptives collected saliva on the last day of the stopping period (day 7) and post-menarchal girls collected saliva on the 7th day of the menstrual cycle, to control for hormonal fluctuations during the cycle. Girls using contraceptives without a stopping period, such as hormonal intrauterine devices were excluded from this study. The saliva samples were assayed for testosterone and estradiol levels at the Department of Clinical Chemistry of the Free University Medical Centre. The lower detection limit was 4 pmol/L for testosterone, and 0.1 pg/ml for estradiol.

Testosterone levels from saliva were determined by isotope dilution - online solid phase extraction liquid chromatography – tandem mass spectrometry (ID-XLC-MS/MS; (Peper, Mandl, et al., 2013). Intra-essay coefficient of variation (CV) was 11% and 4%, at 10 and 140 pmol/L, respectively and inter-assay CV was 8% and 5%, at 31 and 195 pmol/L, respectively (de Water, et al., 2013). Seventeen girls and fourteen boys did not succeed in collecting an appropriate quantity of saliva for analysis, leaving the total number of included hormonal samples at 237. Testosterone levels were highly skewed, thus a log-transformation of the scores was used for further calculations. Testosterone levels were highly correlated with self-report PDS in both boys (r = .70, p < .001) and girls (r = .64, p < .001). Salivary estradiol was determined using an enzyme linked immunosorbent assay (ELISA; DRG Instruments, Marburg, Germany). Inter-assay CV was 8% and 15% at 10 and 40 pg/L, respectively. For eleven girls and fifteen males estradiol could not be determined, leaving the number of included samples at 242. Estradiol levels were correlated with PDS in both boys (r = .23, p = .013) and girls (r = .24, p = .007).

MRI Data Acquisition

MRI scans were acquired with a standard whole-head coil on a Philips 3.0 Tesla MRI scanner. Functional scans were acquired during two runs with T2*-weighted echo-planar imaging (EPI). The first two volumes were discarded to allow for equilibration of T1 saturation effects. Volumes covered the whole brain (TR = 2.2 s, TE = 30 ms, sequential acquisition, 38 slices, slice thickness = 2.75 mm, Field of View (FOV) = 220 x 220 x 114.68 mm). A high-resolution 3D T1-FFE scan for anatomical reference was obtained after the experimental tasks (TR = 9.76 ms, TE = 4.59 ms, 140 slices, voxel size = 0.875 mm, FOV = 224 × 177 × 168 mm). The experimental task was projected on a screen that was viewed through a mirror. Before the MRI session, participants were accustomed to the MRI environment and sounds with a mock scanner.

FMRI Data Analysis

All data was analyzed with SPM8 (Wellcome Trust Centre for Neuroimaging, London). Images were corrected for slice timing acquisition and rigid body motion. Structural and functional volumes were spatially normalized to T1 templates. The normalization algorithm used a 12-parameter affine transform together with a nonlinear transformation involving cosine basis functions and resampled the volumes to 3 mm cubic voxels. Templates were based on the MNI305 stereotaxic space (Cocosco et al., 1997), an approximation of Talairach space (Talairach & Tourneaux, 1988). Functional volumes were spatially smoothed with an 8mm FWHM isotropic Gaussian kernel. The fMRI time series data were modeled by a series of events convolved with a canonical hemodynamic response function. The modeled events were "Positive Learning" (sequence [CORRECT, correct], "Negative Learning" (sequence [ERROR, correct or error]) and "Application" (sequence [correct, CORRECT]), which were time-locked to the moment of feedback. Other trials including trials where participants responded too late were modeled separately, but were not included in the analysis (events of no interest). The events (trials) were used as covariates in a general linear model, along with a basic set of cosine functions that high-pass filtered the data. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used in pair-wise contrasts. The resulting contrast images, computed on a subject-by-subject basis, were submitted to higher-level analyses, also referred to as group analyses. We calculated the contrast Learning (positive and negative feedback combined) > Application to examine which areas contribute to learning from positive and negative feedback. To investigate valence effects, the contrasts Positive Learning > Negative Learning and Negative Learning > Positive Learning were calculated. We used a 10-voxel spatial extent combined with a stringent FWE-corrected intensity threshold to produce a desirable balance between Types I and II error rates (Bennett, Wolford, & Miller, 2009; Forman et al., 1995).

Region-of-interest Analysis

In order to examine transitions in age effects in more detail, region-of-interest (ROI) analyses were performed with the Marsbar toolbox in SPM8 (Brett et al., 2002). The contrast used to generate functional ROIs was Learning (positive and negative) > Application (FWE corrected, $p < .05$, > 10 contiguous voxels). We chose this contrast because it is not biased towards positive or negative feedback, but at the same time reveals task-relevant areas. The resulting ROIs spanned several brain regions. Therefore, the ROIs were subdivided by masking the functional ROI with the following anatomical Marsbar ROIs (based on Automated Anatomical Labeling): left and right DLPFC (Middle Frontal Gyrus), pre-SMA/ACC (Supplementary Motor Area; left and right combined), left and right SPC (Superior Parietal Gyrus). These ROIs were selected based on prior research indicating that these areas show age differences for feedback learning (Crone et al., 2008; van Duijvenvoorde et al., 2008). The DLPFC ROIs were very large (right: 28488 mm; left: 28240 mm), therefore, we created 6 mm radius spheres based on four local maxima within the DLPFC regions (two per hemisphere). These areas are referred to as 'superior DLPFC' and 'mid-DLPFC'. Centre-of-mass MNI (x y z) coordinates were: pre-SMA/ACC: 0 9 58; right superior DLPFC: 21 9 57; left superior DLPFC: -24 3 57, right mid-DLPFC: 42 18 39; left mid-DLPFC: -42 24 39; right SPC: 27 -62 55; left SPC: -23 -64 50.

Structural MRI analysis

Cortical reconstruction was measured automatically using FreeSurfer version 5.0 (http://surfer.nmr.mgh.harvard.edu/) (Dale, Fischl, & Sereno, 1999; Fischl & Dale, 2000). To extract cortical thickness from the functional ROIs, we performed the following steps: 1) Each functional ROI (bilateral mid- and superior DLPFC, pre-SMA/ACC, bilateral SPC) was registered automatically to the FreeSurfer "fsaverage" template and inspected for accuracy of registration. Of note, as FreeSurfer calculates cortical thickness per hemisphere, the pre-SMA/ACC ROI was split into a left and right structural ROI. 2) Individual cortical thickness data was mapped to the "fsaverage" template. 3) Cortical thickness in mm was extracted for each ROI and individual separately. 4) For functional analyses, the bilateral pre-SMA/ACC cortical thickness ROI was averaged back to one ROI. We used an average weighted procedure by taking into account hemispherical differences in surface size maps. In all areas, cortical thickness decreased with age (pre-SMA/ACC: $r = -.40$, $p < .001$; right superior DLPFC: $r = -.15$, $p = .022$; left superior DLPFC: $r = -.21$, $p = .002$, right mid DLPFC: $r = -.19$, $p = .004$; left mid DLPFC: $r = -.29$, $p < .001$; right SPC: $r = -.27$, $p < .001$; left SPC $r = -.27$, $p < .001$).

Statistical Analyses

Behavioral and Region-Of-Interest fMRI data were analyzed with SPSS 19 (Armonk, NY: IBM Corp.). Age effects were investigated by calculating Pearson's correlations between age and sev-

eral measures of interest. For the behavioral and ROI analyses, age was used as a categorical variable in ANOVAs, divided in 10 age groups (8/9, 10, 11, 12, 13, 14, 15, 16, 17, 18-25). We reasoned that this approach allows for a precise index of age effects and allows for comparison with prior literature. LSD post-hoc tests were performed to further investigate significant results. Hierarchical linear regression analyses were performed with SPSS 19 to investigate contributions of age and puberty to performance and neural measures. Age and puberty were added as continuous variables in regression analyses, because this provides the best method to test the relative contributions of both developmental indices. Finally, the contributions of cortical thickness to ROI activity were investigated with a hierarchical linear regression with age and cortical thickness as predictors.

Results

Behavior

On average, participants needed 138.66 trials (SD = 9.11, range 117-165) to complete the task. In general, participants had a high learning rate (M = 93.39%, SD = 5.11%). A t-test indicated no sex differences for learning rate (t (266) = .53, p = .590, d = .06). Learning rate was positively correlated with age (r = .44, p < .001). We calculated positive and negative learning rate separately to investigate possible differences across development. Both positive (r = .42 p < .001) and negative learning rate (r = .34, p < .001) correlated with age, but we did not find a correlation between age and the ratio between positive and negative learning rate.

We also performed these analyses with categorical age groups to pinpoint the exact ages of change. An ANOVA with age group as between-subjects variable and learning rate as dependent variable showed that learning rate improved with age (see Figure 2, F (9,258) = 11.13, p < .001, η^2 = .28). LSD post-hoc tests indicated that ages 8/9, 10, 11 and 13 performed poorer than the adult group (all ps < .05). An additional ANOVA including only age groups 8/9 to 13 confirmed a steep increase in learning rate in this age range, F (4, 140) = 7.34, p < .001, η^2 = .17), and an analysis including age groups 14 to 18-25 confirmed no additional development in this age range (F (4,118) = 1.83, p = .129, η^2 = .06). A regression analysis was performed with age (continuous) and puberty (PDS, testosterone and estradiol) as predictors for learning rate, for boys and girls separately. We tested whether puberty explained additional variance above age by using the Enter-method, with age as first predictor and puberty as second predictor. The model with puberty as a second predictor was a significantly better predictor in males than the model with age alone, with PDS as the only significant puberty predictor (step 1: R^2 = .29; age: β = .54, p < .001; step 2: ΔR^2 = .08; age: β = .76, p < .001, PDS: β = -.41, p = .002). No such effect was found for testosterone: β = -.06, p = .680 or estradiol: β = .06, p = .570. In girls, puberty (PDS, testosterone or estradiol) did not explain additional variance.

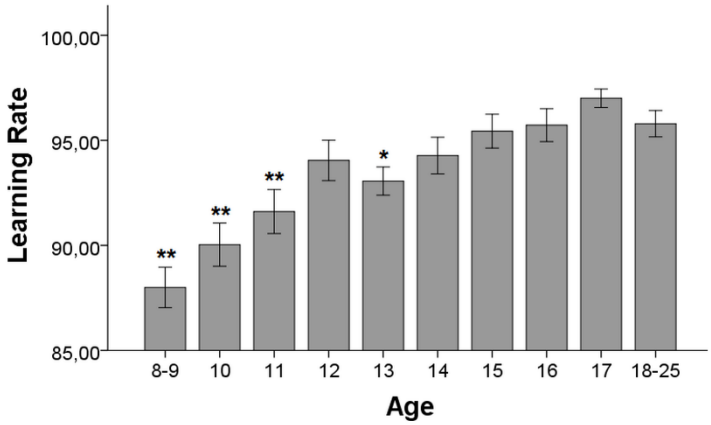


Figure 2: Learning rate per age group (± SEM). Asterisks represent a difference with the adult group, with one asterisk (*) indicating p < .05, and two asterisks (**) indicating p < .01.

Whole Brain analyses

Learning versus Application

To examine which areas were important for feedback learning, we calculated the contrast Learning (positive and negative feedback combined) > Application. This contrast revealed widespread activation in the frontoparietal network including bilateral DLPFC, pre-SMA/ACC, and bilateral SPC (see Figure 3 a, Table 2). Next, to test for developmental effects, we performed an analysis on the contrast Learning > Application with age as a continuous positive linear regressor, which resulted in activation in a similar but less widespread network (see Figure 3 b, Table 2). We also tested for positive effects of age while controlling for learning rate, to identify areas with specific sensitivity to increasing age. This resulted in largely overlapping activated areas compared to the analysis without performance control (see Figure 3 c, Table 2). Finally, we performed an analysis on the contrast Learning > Application with learning rate as a positive regressor (controlled for age), to unravel activity patterns that were associated with better performance. We only found activity in a small cluster in left pre-SMA/ACC (local maximum at -9, 15, 51, T = 5.23, size = 13 voxels). The same analysis but with learning rate as a negative regressor did not result in any activation.

Table 2: MNI coordinates local maxima activated for the contrast Learning > Application. Anatomical labels were acquired with Automated Anatomical Labeling (AAL).

	Area of activation	***x***	***y***	***z***	***voxels***	***T***
	Learning > Application					
Frontal Cortex/	Superior medial frontal gyrus	0	24	42	11313	32.28
Subcortical	R putamen	30	21	0		30.74
	R superior frontal gyrus	24	9	60		29.73
Parietal Cortex	R inferior parietal lobule	45	-42	48	8653	29.86
	L inferior parietal lobule	-48	-45	48		25.92
	R precuneus	6	-66	48		24.35
Temporal Cortex	L middle temporal gyrus	-57	-30	-9	113	12.77
	R middle cingulum	3	-30	27	12	5.94
	Learning > Application, age as positive regressor					
Frontal Cortex	R middle frontal gyrus	30	9	60	521	9.29
	R inferior frontal gyrus	48	12	33		6.57
	R inferior frontal gyrus	45	33	24		6.29
	L precentral gyrus	-36	0	60	732	8.24
	L precentral gyrus	-48	3	51		7.68
	L superior frontal gyrus	-21	12	63		7.37
	L inferior frontal gyrus	-30	30	0	65	5.58
	L inferior frontal gyrus	-45	21	-3		5.42
	L middle frontal gyrus	-36	54	15	30	5.50
Subcortical	R caudate nucleus	6	21	0	260	7.00
	L caudate nucleus	-9	9	0		6.79
	R caudate nucleus	9	9	0		6.63
Parietal Cortex	R inferior parietal lobule	42	-39	48	2127	10.15
	R superior parietal lobule	21	-69	54		9.36
	L superior parietal lobule	-27	-63	54		9.07
Occipital Cortex	L inferior occipital gyrus	-48	-66	-15	407	8.54
	L fusiform gyrus	-33	-51	-18		5.54
	L middle occipital gyrus	-36	-84	6		5.24
Temporal Cortex/	R inferior temporal gyrus	51	-60	-15	292	8.06
Occipital Cortex/	R cerebellum	33	-69	-21		5.75
Cerebellum	R inferior occipital gyrus	27	-93	-12		4.96
	L cerebellum	-3	-81	-18	52	6.13
	L lingual gyrus	-12	-93	-15		4.89

Learning > Application, age as positive regressor, controlled for learning rate						
Frontal Cortex	R middle frontal gyrus	30	12	60	76	7.56
	L precentral gyrus	-36	0	60	164	6.60
	L superior frontal gyrus	-21	12	63		6.19
	L precentral gyrus	-51	12	42		5.54
	R inferior frontal gyrus	51	12	33	25	5.02
	R middle frontal gyrus	51	12	45		4.79
	R inferior frontal gyrus	45	33	24	10	4.98
Subcortical	R caudate nucleus	6	18	0	76	5.88
	L caudate nucleus	-6	9	0		5.70
	L caudate nucleus	-3	18	0		5.19
Parietal Cortex/	R inferior parietal lobule	42	-39	48	626	8.16
Occipital Cortex	R superior parietal lobule	21	-69	54		7.21
	R middle occipital gyrus	33	-72	33		7.08
	L superior parietal lobule	-27	-63	54	480	6.99
	L inferior parietal lobule	-42	-45	57		6.50
	L inferior parietal lobule	-39	-45	48		6.34
	L inferior occipital gyrus	-48	-66	-15	123	6.75
	L fusiform gyrus	-33	-72	-18		5.05
Temporal Cortex/	R inferior temporal gyrus	51	-60	-15	104	6.25
Occipital Cortex	R inferior occipital gyrus	42	-78	-15		5.52

Negative Learning versus Positive Learning

The contrast Negative Learning > Positive Learning and the reverse contrast were calculated to identify areas sensitive to feedback valence during learning. Negative Learning > Positive Learning resulted in activity in bilateral DLPFC, pre-SMA/ACC, anterior and mid cingulum, SPC and several subcortical areas (see Figure 4 a; Table 3). Next, this analysis was performed with age entered as a positive linear regressor, which showed increased activity in bilateral SPC, bilateral DLPFC and pre-SMA/ACC with increasing age (see Figure 4 b, Table 3). This analysis was also performed with age as a positive regressor while controlling for learning rate, to investigate activity that is specifically linked to age and not performance differences. This resulted in increased activity in superior frontal gyrus, SPC and occipital areas (see Figure 4 c, Table 3). Age as a negative regressor did not result in any activity. We also calculated this contrast with learning rate as a positive regressor (controlled for age), but this did not result in any significant activation. The same contrast, but with age as a negative regressor resulted in activity in a small cluster in the right insula (local maximum at 30, -27, 24, $T = 5.53$, size = 19 voxels).

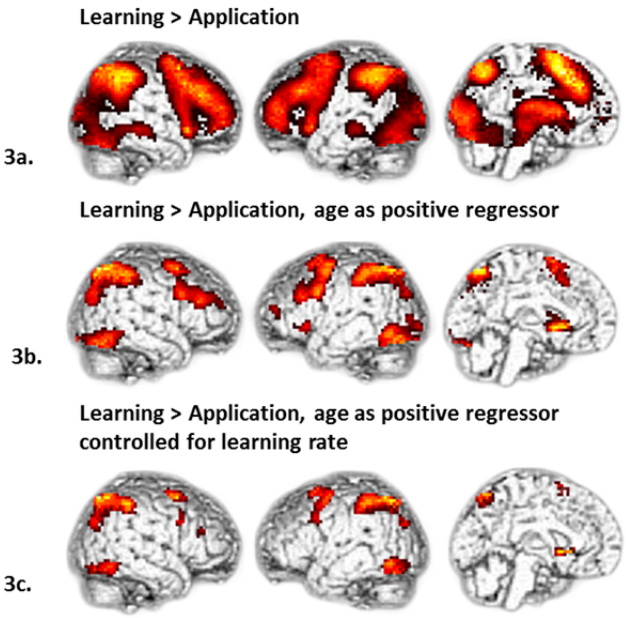


Figure 3: a) Learning (positive and negative) > Application. b) Learning > Application, with age as a positive linear regressor. c) Learning > Application, with age as a positive linear regressor, controlled for learning rate.

Table 3: MNI coordinates for local maxima activated for the contrast Negative Learning > Positive Learning.

	Area of activation	***x***	***y***	***z***	***voxels***	***T***
	Negative Learning > Positive Learning					
Frontal Cortex	R superior medial frontal gyrus	9	27	39	3638	17.88
	R insula	30	24	0		16.18
	R superior frontal gyrus	21	9	57		14.75
	L middle frontal gyrus	-42	24	39	60	6.80
	L middle frontal gyrus	-30	51	9	75	6.30
Frontal Cortex/	L insula	-30	21	-3	552	11.23
Subcortical	R caudate nucleus	15	18	9		9.42
	L caudate nucleus	-12	15	9		9.41
Parietal Cortex	R angular gyrus	54	-51	33	923	15.25

	L inferior parietal lobule	-51	-54	39	303	8.47
	L inferior parietal lobule	-42	-42	42		7.73
	R precuneus	9	-63	48	38	7.15
	L precuneus	-9	-63	48	14	5.84
Temporal Cortex	L calcarine gyrus	-9	-90	6	94	7.12
	R superior temporal gyrus	51	-24	-6	61	6.81
	Negative Learning > Positive Learning, age as positive regressor					
Frontal Cortex	R superior frontal gyrus	24	9	57	150	6.52
	R inferior frontal gyrus	45	9	30	79	5.76
	L supplementary motor area	-12	6	66	69	5.65
	L superior frontal gyrus	-18	12	60		5.37
	L middle frontal gyrus	-30	6	57		5.04
	R middle frontal gyrus	48	39	21	23	5.08
Parietal Cortex/	R superior parietal lobule	27	-63	54	891	7.31
Occipital Cortex	R middle occipital gyrus	30	-69	30		7.07
	R angular gyrus	27	-60	39		6.58
Temporal Cortex/	R inferior temporal gyrus	51	-51	-12	265	6.28
Occipital Cortex	R inferior occipital gyrus	30	-90	-3		5.62
	R inferior occipital gyrus	36	-87	-12		5.30
Occipital Cortex	L middle occipital gyrus	-24	-93	3	290	5.55
	L inferior occipital gyrus	-27	-93	-6		5.42
	L middle occipital gyrus	-30	-84	0		5.20
	L inferior occipital gyrus	-48	-75	-12	25	5.14
Negative Learning > Positive Learning, age as positive regressor, controlled for learning rate						
Frontal Cortex	R superior frontal gyrus	21	9	51	96	6.31
	L superior frontal gyrus	-21	12	57	15	5.14
	L supplementary motor area	-12	6	66		4.94
Parietal Cortex	L superior parietal lobule	-21	-69	45	29	4.87
	L superior parietal lobule	-24	-69	54		4.86
	L superior parietal lobule	-15	-63	54		4.72
Parietal Cortex/	R middle occipital gyrus	30	-66	30	306	6.24
Occipital Cortex	R superior parietal lobule	24	-63	54		5.94
	R superior occipital gyrus	24	-60	39		5.58
Temporal Cortex	R inferior temporal gyrus	45	-57	-9	75	5.62
	R supramarginal gyrus	42	-42	42	13	5.00
Occipital Cortex	L inferior occipital gyrus	-27	-93	-12	64	5.15
	L middle occipital gyrus	-27	-93	9		5.09

L middle occipital gyrus	-30	-78	30	16	5.08
R inferior occipital gyrus	27	-93	-3	50	5.04
R middle occipital gyrus	33	-84	0		4.95
R inferior occipital gyrus	36	-87	-12		4.82

Positive Learning versus Negative Learning

Positive Learning > Negative Learning revealed activation in bilateral precuneus, SPC, DLPFC, pre-SMA/ACC and subcortical areas (see Figure 4d, Table 4). The analyses with age and learning rate as regressors were already covered in the previous paragraph on the contrast Negative Learning > Positive Learning.

Table 4: MNI coordinates local maxima activated for the contrast Positive Learning > Negative Learning.

	Area of activation	***x***	***y***	***z***	***voxels***	***T***
	Positive Learning > Negative Learning					
Frontal Cortex/	L precentral gyrus	-24	-27	66	14051	13.93
Parietal Cortex/	L precuneus	-9	-57	12		13.11
Subcortical	R caudate nucleus	21	-6	30		12.88
Frontal Cortex	L medial orbital frontal gyrus	-3	60	-3	372	12.27
	L superior medial frontal gyrus	-9	66	12		9.96
	L superior medial frontal gyrus	-9	63	21		9.15
	R inferior frontal gyrus	54	33	6	21	7.91
Parietal Cortex	L angular gyrus	-45	-75	27	117	7.87
Temporal Cortex/	L middle occipital gyrus	-27	-93	0	132	8.51
Occipital Cortex	L inferior occipital gyrus	-33	-90	-9		7.16
	L lingual gyrus	-9	-93	-15		5.47

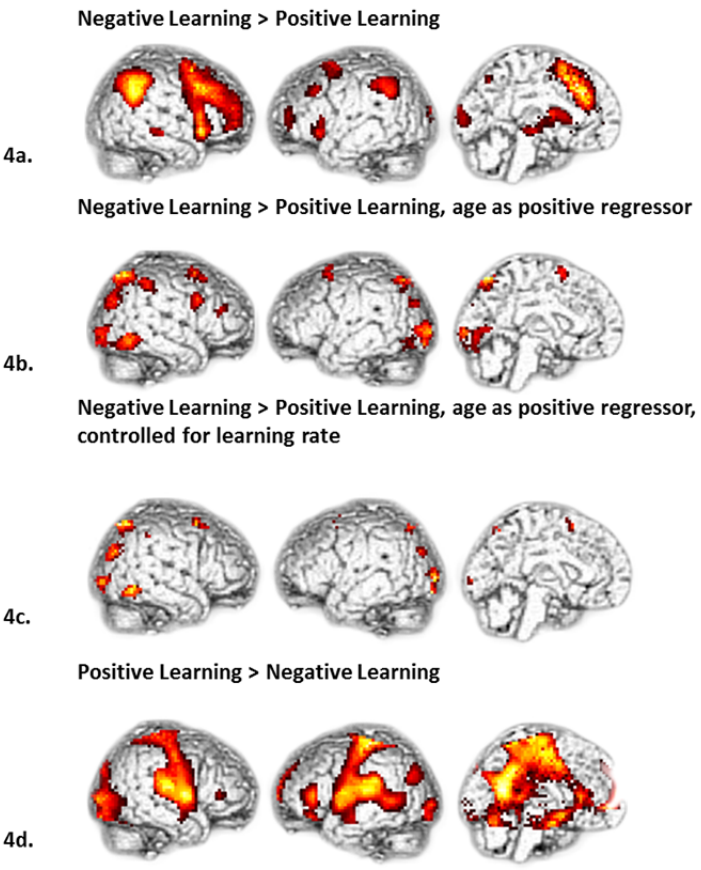


Figure 4: a) Negative Learning > Positive Learning b) Negative Learning > Positive Learning with age as a positive linear regressor c) Negative Learning > Positive Learning, with age as a positive linear regressor, controlled for learning rate. d) Positive Learning > Negative Learning.

Region of Interest analysis: Valence effects

To pinpoint the exact ages at which changes in activation patterns occurred, we used ROI analyses. We focused on activity in pre-SMA/ACC, bilateral superior and mid-DLPFC and bilateral SPC in relation to feedback valence, because prior studies found age differences for feedback learning in these areas (Crone et al., 2008; Van Duijvenvoorde et al., 2008). Age was analyzed as a categorical age group, to identify the precise time of change. The addition of sex did not result in any main or interaction effects. Therefore, effects of sex are not further discussed. In the following analyses, we will more specifically focus on valence effects and discuss them separately per region.

Right superior DLPFC

The age x valence ANOVA for right superior DLPFC resulted in a main effect of valence (F (1,258) = 244.02, $p < .001$, $\eta^2 = .44$) and an age x valence interaction (F (9,258) = 5.38, $p < .001$, $\eta^2 = .09$) (see Figure 5). To follow up on this interaction, we tested for age differences in Positive Learning and Negative Learning separately. For Negative Learning, there was a difference across age groups (F (9,258) = 5.36, $p < .001$, $\eta^2 = .16$), but no age group effect was found for Positive Learning (F (9,258) = 0.48, $p = .886$, $\eta^2 = .02$). A second set of post hoc analyses tested for the valence effects within each age group. Paired-samples t-tests were performed for each age group between Positive Learning and Negative Learning. All ages except 8/9-year olds showed a significant difference between Negative Learningand Positive Learning (all ps < .01; see Figure 5). A third post hoc analysis tested at which the age the neural pattern for valence was adult-like. For this analysis, we calculated the difference between Negative Learning and Positive Learning, which differed across age groups (F (9,258) = 5.375, $p < .001$, $\eta^2 = .16$). An LSD post-hoc comparison with the adult group as baseline indicated that ages 8 to 13 differed significantly from the adult group (8/9y: $p < .001$, 10y: $p = .011$, 11y: $p = .003$, 12y: $p = .002$, 13y $p = .003$), whereas ages 14 to 17 did not differ significantly from adults.

Left superior DLPFC

The age x valence ANOVA for left superior DLPFC showed a main effect of valence (F (1,258) = 63.13, $p < .001$, $\eta^2 = .18$) and an age x valence interaction (F (9,258) = 3.38, $p = .001$, $\eta^2 = .09$) (see Figure 5). To follow up on this interaction, we tested for age differences in Positive Learning and Negative Learning separately. There were no differences across age groups for Negative Learning (F (9,258) = 1.78, $p = .072$, $\eta^2 = .06$) and Positive Learning (F (9,258) = 1.02, $p = .421$ $\eta^2 = .03$). Further post hoc analyses tested for the valence effects within each age group. All ages from 14 to 18-25 showed a significant difference between Negative Learningand Positive Learning (all ps < . 01; see Figure 5). A final set of post hoc analyses tested for the age at which the neural pattern for valence was adult-like. For this analysis, we used the difference score between Negative Learning and Positive Learning, which differed across age groups (F (9,258) = 3.38, $p = .001$ $\eta^2 = .11$). An LSD post-hoc comparison with the adult group as baseline indicated that ages 8/9, 11, 12 and 13 differed significantly from the adult group (8/9y: $p = .006$, 11y: $p = .006$, 12y: $p = .018$, 13y $p = .021$), whereas ages 10 and 14 to 17 did not differ significantly from adults. The interaction between valence x age x hemisphere was not significant (F (9,258) = .50, $p = .872$, $\eta^2 < .001$), indicating there was no significant difference between left and right superior DLPFC.

Right mid-DLPFC

The age x valence ANOVA for right mid-DLPFC resulted in a main effect of valence (F (1,258) = 196.12, $p < .001$, $\eta^2 = .41$) and an age x valence interaction (F (9,258) = 2.53, $p = .009$, $\eta^2 = .05$). Follow-up tests indicated that Negative Learning differed across age groups (F (9,258) = 2.21, $p = .022$,

η^2 = .07), but there was no age effect for Positive Learning (F (9,258) = 1.47, p = .158 η^2 = .05). Next, valence effects were tested for age groups separately. Negative Learning resulted in stronger activations than Positive Learning for all age groups (all ps < .05; see Figure 5). Further post-hoc tests were performed on differential activation for Negative Learning compared to Positive Learning (which differed across age groups: F (9,258) = 2.53, p = .009, η^2 = .08) to test when activation patterns were adult-like. An LSD post-hoc comparison indicated that none of the age groups differ significantly from the adult group.

Left mid-DLPFC

The age x valence ANOVA for left mid-DLPFC resulted in a main effect of valence (F (1,258) = 43.790, p < .001, η^2 = .14), but no age x valence interaction (F (9,258) = 1.78, p = .072, η^2 = .05). Therefore, no further follow-up tests were performed. However, the valence x age x hemisphere interaction was not significant (F (9,258) = .73, p = .684, η^2 = .002), indicating that left and right mid-DLPFC showed a similar pattern of activation.

Right SPC

The age x valence ANOVA for right SPC did not result in a main effect of valence (F (1,258) = 3.32, p = .070, η^2 = .01), but there was a significant age x valence interaction (F (9,258) = 5.14, p < .001, η^2 = .15). Post-hoc comparisons of Positive Learning and Negative Learning separately indicated that there was an effect of age group for Negative Learning (F (9,258) = 7.56, p < .001, η^2 = .21) but not for Positive Learning (F (9,258) = 1.26, p = .262, η^2 = .04).

A further follow-up analysis for each age group separately indicated that ages 8/9, and 16 to 18-25 year olds, differentiated between Positive Learningand Negative Learning (all ps < .05). Notably, 8/9-year olds show more activation after Positive Learning, whereas ages 16 to 18-25 years show more activity after Negative Learning (all ps < .05). Participants aged 10 to14 showed no significant differentiation between Positive Learningand Negative Learning.

A final post-hoc test was performed to investigate when activation patterns reached adult levels. A one-way ANOVA indicated that differential activation for Negative Learning compared to Positive Learning differed across age groups (F (9,258) = 5.143, p < .001 η^2 = .15). LSD post-hoc comparisons indicated that ages 8 to 14 years differed significantly from the adult group (8/9y: p < .001, 10y: p = .007, 11y: p = .001, 12y: p = .005, 13y: p = .002, 14y: p = .045).

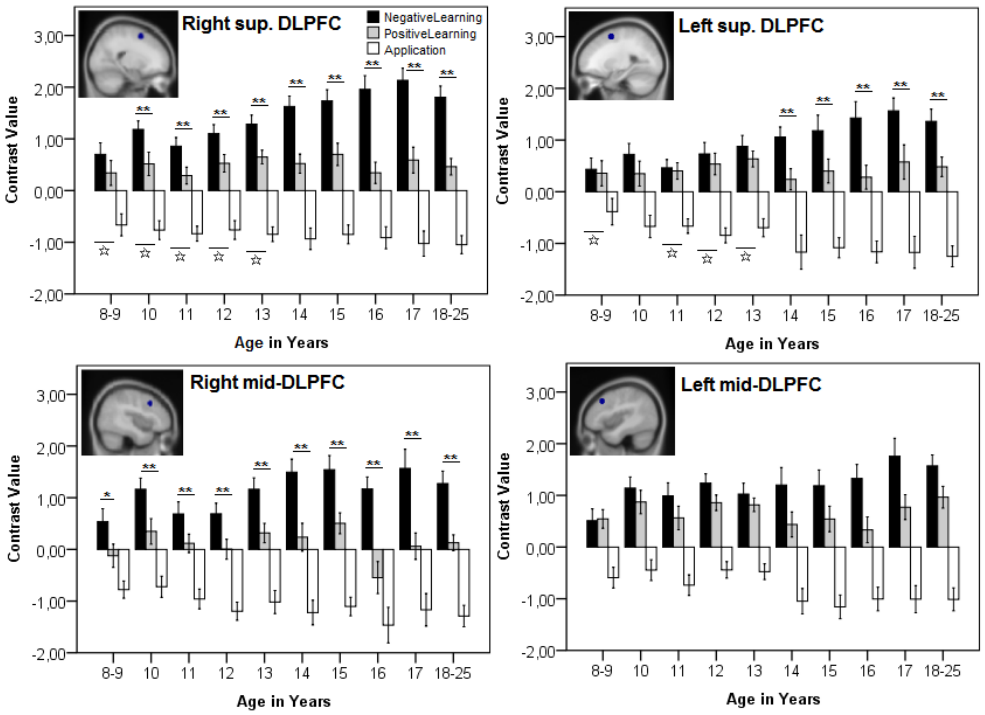


Figure 5a: Activity for Positive Learning, Negative Learning and Application in DLPFC (± SEM). Asterisks represent a significant difference between Positive Learning and Negative Learning, with one asterisk (*) indicating p < .05, and two asterisks (**) indicating p < .01. Stars (☆) indicate a significant difference with the adult group for difference scores between Negative Learning and Positive Learning.

Left SPC

The analyses for left SPC also showed no main effect of valence (F (1,258) < .001, p = .990, η^2 < .001) but a significant age x valence interaction (F (9,258) = 3.70, p < .001, η^2 = .11). Follow-up tests indicated that Negative Learning differed across age groups (F (9,258) = 2.78, p = .004, $\eta 2$ = .09), but there was no age effect for Positive Learning (F (9,258) = .58, p = .811 η^2 = .02). Negative Learning differed from Positive Learning in ages 11, 13 and 18-25 (all ps < .05; see Figure 5), with importantly, 11 and 13-year-olds showing more activity after Positive Learning, but 18-25-year-olds showing more activity after Negative Learning. Further post-hoc tests were performed on differential activation for Negative Learning compared to Positive Learning (which differed across age groups: F (9,258) = 3.70, p < .001, η^2 = .11) to test when activation patterns reached adult levels. LSD post-hoc comparisons indicated that ages 8/9 to 13 (except age 10, p = .060) differed signifi-

cantly from the adult group (8/9y: p = .001, 11y: p < .001, 12y: p = .008, 13y: p = .002). The interaction between valence x age x hemisphere was significant, (F (9,258) = 2.58, p = .007, η^2 = .003), showing that there was a significant difference between left and right SPC.

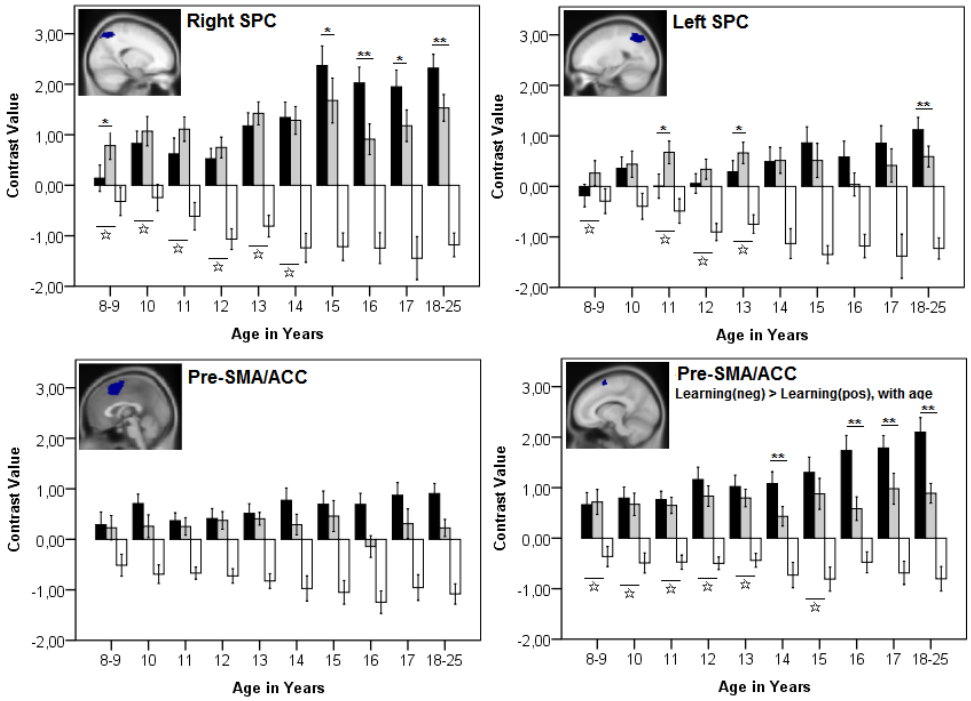


Figure 5b: Activity for Positive Learning, Negative Learning and Application in SPC and pre-SMA/ACC (± SEM). Asterisks represent a significant difference between Positive Learning and Negative Learning, with one asterisk (*) indicating p < .05, and two asterisks (**) indicating p < .01. Stars (☆) indicate a significant difference with the adult group for difference scores between Negative Learning and Positive Learning.

Pre-SMA/ACC

The age x valence ANOVA for pre-SMA/ACC resulted in a main effect of valence (F (1,258) = 28.43, p < .001, η^2 = .09) but no significant age x valence interaction (F (9,258) = 1.83, p = .063, η^2 = .05), indicating that valence effects were present but did not differ per age group (See Figure 5). It was unexpected that no age effects were found in the pre-SMA/ACC, because age effects in this area were found in the whole brain analyses Positive Learning > Negative Learning with age as a

positive regressor. To further investigate age effect s in the pre-SMA/ACC, we selected this wholebrain activated cluster and masked it with an anatomical mask of the pre-SMA/ACC. Subsequently, we performed follow-up analyses on these ROI-extracted values to test which age groups were driving this effect. We first tested for age differences in Positive Learning and Negative Learning separately. For Negative Learning, there was a difference across age groups (F (9,258) = 3.880, $p < .001$ $\eta^2 = .11$), but no age group effect was found for Positive Learning (F (9,258) = .523, $p = .858$, $\eta^2 = .02$). A second set of post hoc analyses tested for valence effects within each age group. Paired-samples t-tests were performed for each age group between Positive Learning and Negative Learning. Ages 14, 16, 17 and 18-25 showed a significant difference between Negative Learning and Positive Learning (all ps < . 01; see Figure 5). A third post hoc analysis tested for the age at which the neural pattern for valence was adult-like. For this analysis, we calculated the difference between Negative Learning and Positive Learning, which differed across age groups (F (9,258) = 3.956, $p < .001$ $\eta^2 = .12$). An LSD post-hoc comparison with the adult group as baseline indicated that ages 8/9 to 13, and 15 differed significantly from the adults (8/9y: $p < .001$, 10y: $p = .001$, 11y: $p < .001$, 12y: $p = .003$, 13y $p = .001$, 15y: $p = .016$).

Effects of puberty on brain activation

To test for effects of puberty on brain activity, a regression analysis was performed with age as first predictor (continuous) and puberty (PDS, testosterone and estradiol) as second predictor for the difference scores for Negative Learning and Positive Learning, for boys and girls separately. PDS, testosterone and estradiol were not significant additional contributors to neural activation.

Effects of structural brain development

To investigate whether structural brain development influenced the relation between age and brain activity, we performed a regression analysis with age as first predictor and cortical thickness as second predictor for the difference scores for Negative Learning and Positive Learning. We found an additional effect of cortical thickness above age in right SPC (step 1: $R^2 = .10$; age: $\beta = .32$, $p < .001$; step 2: $\Delta R^2 = .02$; age: $\beta = .36$, $p < .001$, cortical thickness: $\beta = .13$, $p = .042$), but not in other regions.

Discussion

In this study, we investigated the development of feedback learning in a large sample of children, adolescents and adults. The results indicated that with increasing age, the frontoparietal network becomes increasingly more activated in response to feedback during learning compared to application of rules, but contributed differentially to learning from positive and negative feedback across development. Specifically, we found that a) Activity in DLPFC, SPC and pre-SMA/ACC after negative feedback increased with age, but activity for positive feedback remained constant,

b) Activity in superior DLPFC, SPC and pre-SMA/ACC reached adult levels around age 14/15, c) DLPFC, pre-SMA and SPC contributed differently to learning in children, with SPC showing a shift in activity from positive to negative feedback. In the next paragraphs, these findings will be described in more detail.

Sensitivity to negative feedback increases with development

In adults, we found more activation following negative compared to positive feedback in DLPFC, pre-SMA/ACC and SPC, consistent with prior studies (Holroyd et al., 2004; Zanolie et al., 2008). In childhood and early adolescence (age 8-13/14), children developed into faster learners and showed increasing activation in superior DLPFC, SPC and pre-SMA/ACC after negative feedback with increasing age. When controlling for learning performance, the more anterior portions of the DLPFC were no longer correlated with age, suggesting that these regions might be related more to performance than maturation per se. Intriguingly, in the SPC young children showed increased activation following positive feedback compared to negative feedback, a pattern that has also been observed in a prior study (van Duijvenvoorde, et al., 2008). Also in this area, reactions to positive feedback did not change with development, whereas reactions to negative feedback increased with age. It therefore seems that this effect occurred mainly because of decreased activation after negative feedback in young children, rather than increased positive feedback compared to older participants. Together, these findings suggest that the frontoparietal network functions in a different way in children than in adults.

Transition to adult-like activity patterns

One of the aims of this study was to investigate at which age this network functions in an adult-like way. In terms of behavior and neural activation, adolescents of age 14/15 years and older no longer differed from adults. These findings are the first to provide such a precise index of development, given that prior neuroimaging studies collapsed across wide age ranges (e.g. 13-17 years; (van den Bos, et al., 2009)); or selected specific age groups (e.g. 11-13-year-olds versus adults; (van Duijvenvoorde, et al., 2008)). The transition to adult-like cognitive control functions is consistent with behavioral literature, which has suggested that cognitive development reaches adult levels in early to mid-adolescence (Huizinga et al., 2006; Luna, Garver, Urban, Lazar, & Sweeney).

One possible cause for immature activity in the frontoparietal network is immature structural brain development (Klingberg et al., 2002; Lu et al., 2009). In this study, we found preliminary evidence for a role of cortical thickness on activity patterns, such that right SPC cortical thickness explained additional variance in activity in addition to age. Future longitudinal studies are needed to confirm this finding. Additionally, developmental changes in white matter connectivity have been reported well into adolescence (Lebel & Beaulieu, 2011), which may also explain immature activity patterns in children. Other explanations come from resting state research. The network for cognitive control seems already in place in children, but the strength of connectivity

within this network undergoes developmental changes (Jolles, van Buchem, Crone, & Rombouts, 2011). Also, the strength of short-range connections seems to decrease with age, whereas the strength of long-range connections increases with adolescent development (Dosenbach et al., 2010; Fair et al., 2009).

Another explanation for immature activity is that children differ from adults in strategy use. This is supported by findings that adults generally perform learning tasks in an efficient hypothesis-testing approach, whereas children are more likely to use an inefficient trial-and-error approach (Schmittmann et al., 2006). An intriguing question is whether there is a bias for positive feedback in young children. Possibly, it is adaptive for children to focus attention on positive feedback, because they do not have the capacity to use information derived from negative feedback during hypothesis testing (Schmittmann et al., 2006; van Duijvenvoorde et al., 2008).

Effects of pubertal development

Although this study provides evidence for a transition point in feedback processing at the age of 14/15, future research is needed to unravel the mechanisms behind this transition. One mechanism we explored was pubertal development. Theoretical models have suggested a dual processing network for adolescent development, such that the development of limbic areas is influenced by puberty, whereas the functional development of the frontoparietal network occurs relatively independent of hormones (Nelson, Leibenluft, McClure, & Pine, 2005; Steinberg et al., 2008). Our findings are consistent with this model, and show no additional influence of self-reported puberty stage, testosterone and estradiol on neural activity. Interestingly though, we did find an additional effect of pubertal development in boys on behavioral performance. However, given that no relation between puberty and neural activation in the frontoparietal network was found, the results support the assumption of the dual processing network that cortical development occurs relatively independently from pubertal development.

Limitations

This study has several limitations. First, we only analyzed the trials where learning from feedback was successful, and we have not investigated what happened in the small percentage of trials where learning did not occur. This would be an interesting direction for future research with a task that is more difficult and results in more errors during learning. Second, we only investigated the effects of testosterone, estradiol and PDS in this study. Further research could focus on a more diverse array of puberty measures, such as having a physician assessing pubertal status or collecting multiple hormonal measurements on consecutive days. Finally, the current study was cross-sectional which does not allow us to draw conclusions about change within individuals. Future longitudinal studies will be important to unravel within-person variance over time in performance, strategy use and associated brain activity.

Conclusion

This study provides an overview of the development of feedback learning across adolescent development. Due to the large sample size, it was possible to pinpoint developmental trajectories across adolescence. This study is the first to show that activity in DLPFC, SPC and pre-SMA/ACC after negative feedback increases with age until approximately age 14/15, after which adult levels are reached. We also demonstrated that the SPC shows a qualitative change in recruitment, with more activity in children after positive feedback, but more activity in late adolescents and adults after negative feedback. These findings are interpreted in terms of separable contributions of the frontoparietal network in childhood and more integrated function in adolescence and adulthood. These findings provide important starting points for searching for flexible periods for learning and eventually tailoring educational programs to the needs of children at different stages in development.

Chapter 5

Influence of strategy use and age on neural activity for feedback learning

This chapter is based on:

Peters, S., Koolschijn, P.C.M.P., Crone, E.A., van Duijvenvoorde, A.C.K. & Raijmakers, M.E.J. (2014). Strategies influence neural activity for feedback learning across child and adolescent development. Neuropsychologia, 62, 365-374.

Abstract

Learning from feedback is an important aspect of executive functioning that shows profound improvements during childhood and adolescence. This is accompanied by neural changes in the feedback learning network, which includes pre-supplementary motor area (pre- SMA)/anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC), superior parietal cortex (SPC), and the basal ganglia. However, there can be considerable differences within age ranges in performance that are ascribed to differences in strategy use. This is problematic for traditional approaches of analyzing developmental data, in which age groups are assumed to be homogenous in strategy use. In this study, we used latent variable models to investigate if underlying strategy groups could be detected for a feedback learning task and whether there were differences in neural activation patterns between strategies. In a sample of 268 participants between ages 8 to 25 years, we observed four underlying strategy groups, which cut across age groups and varied in the optimality of executive functioning. These strategy groups also differed in neural activity during learning; especially the most optimal performing group showed more activity in DLPFC, SPC and pre-SMA/ACC compared to the other groups. However, age differences remained an important contributor to neural activation, even when correcting for strategy. These findings contribute to the debate of age versus performance predictors of neural development, and highlight the importance of studying individual differences in strategy use when studying development.

Introduction

An important component of cognitive development is the ability to control and adapt behavior in response to changing environmental demands, also referred to as executive functions (Diamond, 2013; Zelazo, 2006). Executive functions are thought to consist of three core functions: inhibition, working memory and cognitive flexibility (Diamond, 2013). Higher-order executive functions such as reasoning, planning and learning from prior experiences rely upon combinations of these three core functions. The ability to adapt behavior based on prior experiences (i.e. adaptive control) shows a marked improvement during childhood and adolescence (Tamnes, Walhovd, Torstveit, Sells, & Fjell, 2013). For example, in the classic Wisconsin Card Sorting Task (WCST), there is a developmental improvement in flexibly adapting behavior based on positive and negative feedback (Huizinga et al., 2006) and in probabilistic feedback learning tasks there is a developmental improvement in adapting behavior successfully based on informative versus non-informative feedback (Eppinger et al., 2009; Jansen, van Duijvenvoorde, & Huizenga, 2014; van den Bos et al., 2009; Van Duijvenvoorde, Jansen, Griffioen, Van der Molen, & Huizenga, 2013). Despite these convincing developmental patterns, there are large individual differences in adaptive control within age ranges, i.e., not all children and adolescents are equally proficient at learning from positive and negative feedback. Why is it that some children are better at learning compared to their peers? Studying the behavioral and neural mechanisms underlying successful learning is important to advance our understanding of executive control processes and their development.

Most prior studies on the development of feedback learning have focused on performance improvements with age and the accompanying changes in brain activity. Research in adults indicated that during feedback learning, a large brain network is activated, including pre-supplementary motor area (pre-SMA)/anterior cingulate cortex (ACC) (Holroyd et al., 2004; Mars et al., 2005; Monchi et al., 2001; Ullsperger & von Cramon, 2003), (dorso)lateral prefrontal cortex (DLPFC) (Dove et al., 2000; Lie et al., 2006; van Veen et al., 2004; Zanolie et al., 2008), basal ganglia (Monchi et al., 2001; Tricomi et al., 2006), and superior parietal cortex (SPC) (Zanolie et al., 2008). It is thought that a dopamine-initiated alarm signal in pre-SMA/ACC signals that outcomes are worse than expected. Subsequently, the DLPFC is a primary site for implementation of adaptive control (Kerns et al., 2004; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004).

Prior developmental studies have shown that this feedback learning network becomes increasingly activated with age (Crone et al., 2008; Peters, Braams, Raijmakers, Koolschijn, & Crone, 2014; van den Bos et al., 2009; van Duijvenvoorde et al., 2008). However, it is unclear whether these neural changes reflect age differences (i.e., a maturational viewpoint), or whether they are related more to performance differences rather than age (Andersen et al., 2014; Jolles & Crone, 2012; Koolschijn et al., 2011).

Effects of performance versus age are only scarcely investigated in developmental feedback learning studies. Moreover, most studies have assumed that performance differences are continuous, implying that all participants within an age group perform the task using the same strategy. However, performance is not constant within age groups: Some children perform at levels similar to adults, whereas others never seem to reach the highest performing levels. It is possible that these individual differences in performance can be described by differences in strategy use. Such differences in performance and strategy use within age groups pose a considerable problem for traditional ways of analyzing developmental data, because these are based on the assumption of homogenous strategy use within age groups.

A robust approach for analyzing individual differences are categorical latent variable models, which allow for detection of different strategies based on individuals' responses across trials. Such techniques have been applied by a number of studies that distinguished distinct learning strategies within age groups (Andersen et al., 2014; Raijmakers, Dolan, & Molenaar, 2001; Schmittmann et al., 2012, 2006; Speekenbrink, Lagnado, Wilkinson, Jahanshahi, & Shanks, 2010). For instance, Schmittmann et al. (2006) showed that two distinct learning strategies (resulting in relatively fast or slow learning) could be distinguished in a category-learning task. The fast and slow strategy groups both employed a learning strategy based on hypothesis-testing (as opposed to incremental, associative learning), but participants in the slow group were less efficient in their hypothesis testing compared to the fast group. This difference in efficiency was categorical. That is, with age, children were increasingly likely to belong to the faster strategy group; they were not simply less efficient in employing the same strategy. In the current study, we applied these methods to a feedback learning task and investigated whether distinct learning strategies were also observable at the neural level.

In the current paradigm, we built on prior studies on the development of feedback learning, such as a rule switch task used in Crone et al. (2008) and a rule search and application task used by Van Duijvenvoorde et al. (2008), and constructed a paradigm in which correct responses could be inferred through a process of hypothesis-testing. Additionally, different deductive reasoning steps could be applied to use a more efficient hypothesis testing strategy. This made the task suitable for differentiating between categorically different strategies, rather than simply assessing performance differences within one strategy. We asked 268 participants ranging from 8 to 25 years to sort stimuli in one of three locations by using positive and negative feedback. An efficient way of solving this task was to not only focus on feedback for the current stimulus but also to remember the locations for the other two stimuli. We recorded trial-by-trial data on learning efficiency and analyzed this with latent variable modeling approaches (Markov models and finite mixture models), to investigate if latent strategy groups could be detected (van der Maas & Straatemeier, 2008). As a further addition to prior research, we investigated if underlying strategy groups could be distinguished at the neural level (see also Andersen et al., 2014). We hypothesized that age differences in neural activity for feedback learning are largely attributable

to differences in strategy use. Thus, we tested whether age differences in neural activity were influenced by strategy use, or if there was also neural activity related to maturational processes per se, independent of strategy use. The main developmental effects have previously been reported in Peters et al., 2014. This dataset presents a unique opportunity for analyzing strategy-related versus age-related neural changes in feedback learning given the large-sample size across a broad developmental range.

Methods

Participants

The sample included 268 participants (138 female) between 8.01 and 25.95 years old (M = 14.22, SD = 3.63), who were recruited through local schools and advertisements. See Table 1 for the number of participants per age and per sex.

Table 1: Number of participants per age and sex.

Age	N Female	N Male	N Total
8 years	6	4	10
9 years	14	5	19
10 years	11	12	23
11 years	13	14	27
12 years	19	11	30
13 years	16	20	36
14 years	10	17	27
15 years	10	11	21
16 years	11	9	20
17 years	12	11	23
18-25 years	16	16	32
N Total	138	130	268

Adult participants (18-25 years) were grouped together. A chi square test indicated that the proportion of males to females was similar across age groups (χ^2 (10) = 9.20, p = .514). IQ scores were estimated with two subtests of the WAIS-III or WISC-III (Similarities and Block Design). Estimated IQ scores ranged from 80 to 143 (M = 110.25, SD = 10.62) and showed no correlation with age (r = -.09, p = .155). None of the participants reported a history of neurological or psychiatric disorders or current use of psychotropic medication. All anatomical MRI scans were reviewed and cleared by a radiologist. The study was approved by the Institutional Review Board at the Leiden University Medical Center and all participants (or participant's parents for minors) provided

written informed consent. Adults received payment for participation, and children and their parents received presents and a fixed payment for travel reimbursement.

Exclusion criteria

Twenty-five participants were excluded (not included in Table 1) from further analyses after participation for the following reasons: Nineteen participants were excluded because movement in the MRI scanner exceeded 3.0 mm in any direction, three participants were excluded because of technical problems and three participants were excluded because they were outliers (more than three times the interquartile range) on the total percentage of positive feedback, indicating they did not perform the feedback learning task adequately.

Feedback Learning Task

Participants performed a feedback learning task in the MRI scanner (see also Peters et al., 2014). On each trial, they saw three empty squares, under which one of three different stimuli was presented (see Figure 1). We explained to the participants that each stimulus belonged in one of the three squares and they had to find the correct location for all stimuli by using performance feedback. Performance feedback was a plus-sign for positive feedback and a minus-sign for negative feedback. After either 12 trials, or when the participant correctly applied the correct location twice in total for each stimulus, the sequence ended and a new sequence was presented with three new stimuli. There were 15 sequences in total, resulting in a maximum of 180 trials. Stimuli were presented in a pseudorandom order, with a maximum of two identical stimuli in a row. Before the MRI session, all participants practiced three sequences. During the MRI session the task was divided into two runs of eight and seven sequences. Each trial started with a 500 ms fixation cross. Consecutively, stimuli were presented for 2500 ms, during which time window the response had to be given. Participants saw the words "Too Late" if they did not respond within this time window, after which the sequence continued. After the response, performance feedback was presented for 1000 ms. Inter-trial intervals were jittered to optimize the timing for fMRI based on OptSeq (Dale, 1999) with intervals between 0 and 6 seconds.

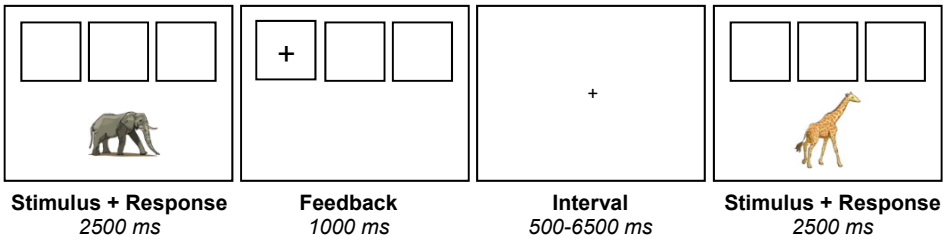


Figure 1: Display of task sequence.

Modeling strategies

To model latent behavioral strategies for the task we first recoded the trial-by-trial accuracy data into new trial-by-trial data, which categorizes different response types that revealed reasoning level. Each trial was scored as one of these four response types, in ascending order of the complexity of reasoning involved:

(a) *Mistake*: repeating a previously made error or making an error after an earlier correct response for the same stimulus. This can be seen as a short-term memory (STM) error (Diamond, 2013), where the subject has forgotten the correct answer it has seen before; no active working memory (WM) calculations were necessary to avoid a mistake. (b) *Inefficient*: when the location of one picture was known, the participant failed to deduce that this location could not be the correct location for another picture. This could be described as a working memory error. In this trial the subject could have avoided an error by means of simple WM calculations. (c) *Suboptimal*: when participants received negative feedback for a stimulus in one location, the optimal decision is to place another stimulus in that location, to ensure a 50 percent chance of being correct (instead of 33 percent). Failing to use this strategy is suboptimal in reasoning, i.e., the subject did not optimize the probabilities of a correct choice. The avoidance of suboptimal trials involves complex executive functions, such as planning and reasoning (Diamond, 2013). (d) *Optimal*: all other cases, i.e. choosing the best possible option given the information that is acquired from previous trials.

The resulting trial-by-trial data on response types was used to distinguish individuals with distinct learning strategies. To detect latent strategies in trial-by-trial data, we used two types of categorical latent variables models. The first type is a static, finite mixture distribution model. Finite mixture distribution models group individuals per learning strategy (based on their response pattern), which is assumed to be constant across the task. The second type is a dynamic Markov model that defines changes in learning strategies across trials (Rabiner, 1989; Visser, 2011). Simulation studies show that such latent variable models are robust statistical techniques, which are necessary to make a reliable decision about the number of strategies and about the nature of the strategies, compared to more ad hoc methods for distinguishing between different learning strategies (van der Maas & Straatemeier, 2008). For fitting models to the data, we calculated maximum likelihood estimates of the parameters in the model by using the statistical R-package depmixS4 (Visser & Speekenbrink, 2010). To determine the most parsimonious, best fitting model to the data, i.e., the optimal model, the Bayesian Information Criterion (BIC) (Schwarz, 1978) was used.

First, we used static, finite mixture distribution models to determine whether performance differences could best be described by either a continuous variation in performance (with age as a covariate) or by a number of categorically different strategies. The former is modeled by including age as a covariate on the response probabilities. According to this model, each individual responds with specific probabilities for the four response types (optimal, inefficient, mistake, suboptimal) and this response pattern depends on age, but without the presence of any latent

strategies. Next, we fitted finite mixture distribution models to the data with a varying number of groups. Comparing these models by means of BIC will show if a model with a certain number of categorically different strategies is better than one with continuous performance differences. Second, we applied a dynamic Markov model to the data. Markov models define changes in learning strategies across trials (Rabiner, 1989; Visser, 2011). We thus extended the mixture of static models (first class of models) to a mixture of (dynamic) Markov models to allow for changes during the learning process. The first Markov models include different states within a learning event. These models could describe, for example, subgroups of participants switching between an inefficient towards an efficient state within a sequence (i.e., different phases during one learning event). The second type of Markov models we tested included a continuous change of strategies across trials within a sequence (i.e., for one or more latent strategies, response probabilities continuously vary with trial number).

Model-comparisons by means of BIC indicated that models with the most parsimonious strategies, corresponding to the firstly described static finite mixture distribution models were optimal for this data compared to the dynamic Markov models. In the results section, we will therefore focus on the static finite mixture distribution models and discuss the number of latent strategies (one to five) that is optimal to describe the data. When the results indicated that the optimal model for the data consisted of multiple strategies, we assigned individuals to the strategy that they were most likely applying based on the posterior probabilities of the data given the model (Visser, 2011). Subsequently, this assignment of individuals was used in the fMRI analysis.

FMRI analyses: Learning and Application

For the fMRI analysis, we used a contrast that reveals brain areas with sensitivity to feedback with learning value. That is, we aimed to find areas that respond more to feedback that provides new information, compared to feedback that provides information that is already known. In our opinion, distinguishing between 'useful' and 'less useful' feedback is one of the key aspects of feedback learning, which is why we expected that this contrast is related to individual differences in performance and strategy use. To identify 'useful' and 'less useful' feedback, we distinguished between a learning phase and an application phase for each stimulus. The learning phase was defined as those trials in which participants had not yet responded with the correct location for the stimulus, and were thus still trying to find the correct solution. We only included trials that actually resulted in learning (M = 96.35 %, SD = 3.03 % of all trials). For learning from positive feedback (Positive Learning), this meant choosing the same location on a next trial for that same stimulus, and for learning from negative feedback (Negative Learning), that meant *not* choosing the same location on a next trial. The application phase was defined as those trials in which a stimulus was sorted correctly on a preceding trial, and which continued to be sorted correctly. All analyses were based on the contrast Learning (Positive Learning & Negative Learning) > Application, in which Positive Learning and Negative Learning were combined and compared to Appli-

cation. In doing this we followed a similar approach to a prior study focusing on feedback differing in informative value for learning (Eliassen et al., 2012). One potential confound is that effects are due to negative feedback rather than feedback which signals learning. However, whole-brain results were highly similar when the Learning (Positive Learning & Negative Learning) > Application and the contrast Positive Learning > Application are compared (see Figure 2). We additionally tested if including negative feedback in the contrast did not lead to additional activation when compared to Positive Learning > Application. Thus we tested with an exclusive mask if Learning (Positive Learning & Negative Learning) > Application showed any significant activation that was not in the contrast Positive Learning > Application. This analysis did not result in significant remaining activations. From now on, the contrast Learning (Positive Learning & Negative Learning) > Application will be abbreviated to Learning > Application.

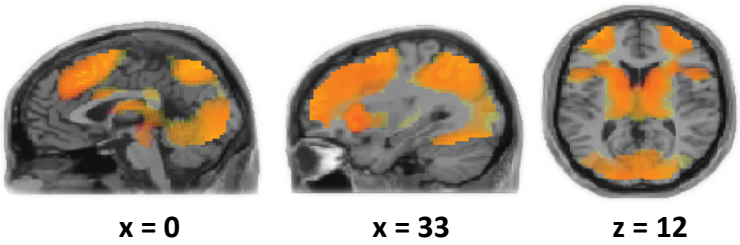


Figure 2: Overlay showing activity for Positive Learning > Application (yellow), Learning (Positive Learning & Negative Learning) > Application (red) and their overlap (orange) (N = 268). Both contrasts were FWE-corrected, p < .05, > 10 contiguous voxels.

Data Acquisition

MRI scans were acquired with a standard whole-head coil on a Philips 3.0 Tesla MRI scanner. Functional scans were acquired during two runs with T2*-weighted echo-planar imaging (EPI). The first two volumes were discarded to allow for equilibration of T1 saturation effects. Volumes covered the whole brain (TR = 2.2 s, TE = 30 ms, sequential acquisition, 38 slices, slice thickness = 2.75 mm, Field of View (FOV) = 220 x 220 x 114.68 mm). A high-resolution 3D T1-FFE scan for anatomical reference was obtained after the experimental tasks (TR = 9.76 ms, TE = 4.59 ms, 140 slices, voxel size = 0.875 mm, FOV = 224 × 177 × 168 mm). The experimental task was projected on a screen that was viewed through a mirror. Before the MRI scan, participants were accustomed to the MRI environment and sounds with a mock scanner.

FMRI Data Analysis

All data were analyzed with SPM8 (Wellcome Department of Cognitive Neurology, London). Images were corrected for slice timing acquisition and rigid body motion. Structural and functional volumes were spatially normalized to T1 templates. The normalization algorithm used a 12-parameter affine transform together with a nonlinear transformation involving cosine basis functions and resampled the volumes to 3 mm cubic voxels. Templates were based on the MNI305 stereotaxic space (Cocosco et al., 1997), an approximation of Talairach space (Talairach & Tourneaux, 1988). Functional volumes were spatially smoothed with an 8mm FWHM isotropic Gaussian kernel. The fMRI time series data were modeled by a series of events convolved with a canonical hemodynamic response function. The modeled events were "Positive Learning", "Negative Learning", and "Application", which were time-locked with 0-duration to the moment of feedback. All other trials (e.g., trials that did not result in learning or too-late trials) were modeled as events of no interest. The trials were used as covariates in a general linear model; along with a basic set of cosine functions that high-pass filtered the data. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used in pair-wise contrasts. The resulting contrast images, computed on a subject-by-subject basis, were submitted to group analyses. All fMRI analyses were initially calculated with a stringent FWE-corrected threshold at $p < .05$, with at least 10 contiguous voxels. In order to examine whole brain effects in more detail, region-of-interest (ROI) analyses were performed with the Marsbar toolbox in SPM8 (Brett et al. 2002).

Results

Modeling results

We fitted finite mixture distribution models to the trial-by-trial data on response types (see Methods section) with a different number of groups (one to five subgroups). In addition, we fitted a 1-state model (i.e., no underlying strategy groups) to the data that assumes continuous individual variation in behavior related to age. By comparing the BIC between models we selected the optimal model to describe the data. Table 2 shows the fit statistics of the different models. First, it appeared that a continuous variation of response probabilities with age was not the optimal way to describe the data. That is, the 1-state age model does not have the lowest BIC. The optimal model assumed the presence of four latent groups (4-states model). We did not test more than five groups because the BIC increased with five groups compared to four groups. Table 3 shows the number of participants per group and how the four groups are defined by percentages of response types.

As can be seen in Table 3, the four strategy groups differed in the number of trials per trial type. We described these four groups as follows: 1) Low strategy group: Relatively many mistakes (no reasoning required, errors in STM or motor response), inefficient responses (not

taking into account feedback for other stimuli: WM errors) and suboptimal responses (not maximizing chance of correctness). 2) Medium-Suboptimal group: Fewer mistakes and inefficient decisions than the Low group, but still relatively many suboptimal decisions. 3) Medium group: Comparable to the Medium-suboptimal group, but fewer suboptimal decisions. 4) High strategy group: Almost all trials are optimal in terms of strategy use.

Table 2: Fit statistics of the static latent variable models with different number of latent strategies.

Model	Loglike	df	BIC
1 state	-15106.51	3	30244.56
1 state age	-14898.26	6	29859.60
2 states	-14605.19	7	29283.97
3 states	-14496.1	11	29107.84
4 states*	-14461.58	15	29080.87
5 states	-14461.32	19	29122.39

Note: The models are explained in the methods section. A 1-state model assumes one strategy; a 2-state model two strategies, etc. The 1-state age model assumes continuous variation between individuals that is related to age. Loglike is the loglikelihood of the fitted model; df = degrees of freedom, which is in this case the number of freely estimated parameters; BIC = Bayesian Information Criterion.

** The model with the lowest BIC is the optimal model.*

Table 3: The number of participants per group and percentage of trials per response type.

Strategy group	N	Mistake	Inefficient	Suboptimal	Optimal
Low	33	7.04 %	8.14 %	7.52 %	77.30 %
Medium-Suboptimal	46	2.12 %	3.58 %	6.77 %	87.53 %
Medium	76	2.91 %	4.51 %	2.22 %	90.36 %
High	113	0.74 %	1.45 %	1.04 %	96.77 %

A repeated-measures analyses with the frequencies of the four response types as the within-subjects variables, and strategy group as between-subjects variable yielded a significant interaction between response-type and strategy group (F (9, 792) = 109.94, p < .001), which is to be expected beforehand from the mixture distribution analysis. Further follow-up one-way ANOVA tests indicated that there were significant between-group differences in the frequency of optimal trials (F (3, 264) = 50.84, p < .001, with Bonferroni post-hoc tests indicating that all groups differed from each other (all ps < .002); except the Medium group and the High group); the frequency of mistakes (F (3,264) = 194.15, p < .001, post-hoc tests showed all groups differed from each other (all ps < .017)); the frequency of inefficient responses (F (3,264) = 189.44, p < .001, post-hoc tests indicated that all groups differed (all ps < .015)); and finally, the frequency of suboptimal responses (F (3,264) = 276.56, p < .001, with post-hoc tests showing that all groups differed from each other (all ps < .005)). These results provide evidence that the High strategy group is most efficient in the

hypothesis-testing process because almost no reasoning mistakes were made. In the next paragraphs, we will describe how the four strategy groups are divided over age groups, and subsequently how the groups differed in brain activity during the feedback learning task.

Division of age and performance over the different strategy groups

In Figure 3, the division of strategy groups over discrete age groups is displayed. A chi-square test indicated that age groups differed across the strategy groups (χ^2 (30) = 82.80, $p < .001$). This indicates that age is an important factor contributing to strategy use. However, even in the two extreme groups (Low and High strategy), children and adults of different ages were distributed over these strategy groups, such that, for instance, some young children (from 9 years onwards) were present in the High strategy group and one adult was present in the Low strategy group. Note that IQ did not differ across strategy group as indicated by a one-way ANOVA (F (3, 262) = 1.82, $p = .145$). The division across strategy groups was similar for males and females (χ^2 (3) = 0.25, $p = .969$).

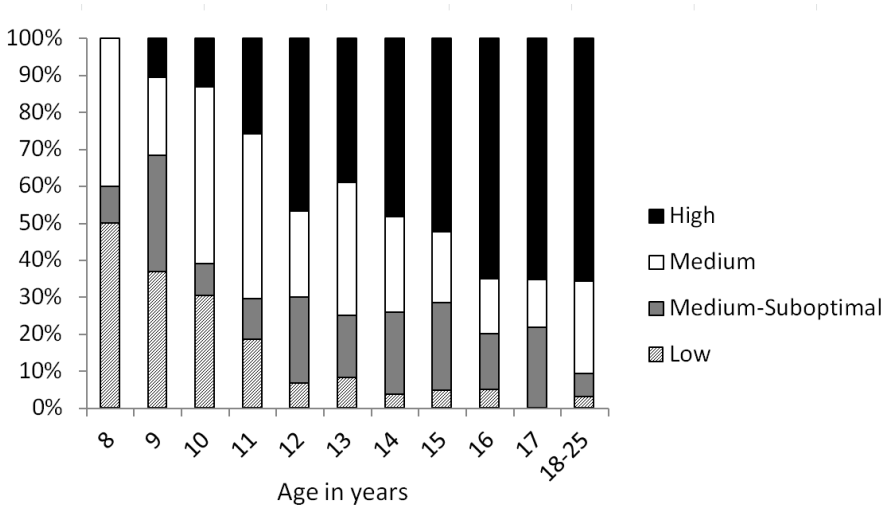


Figure 3: Division of strategy groups (in percentages) over age groups.

FMRI analysis

The percentage of trials for Learning and Application out of the total number of trials per strategy group are displayed in Table 4. Note that the total number of trials could differ per participant due to the nature of the task. 'Other trials' are trials which did not result in learning, incorrect

applications or trials where the participant responded too late. These trials were not used for further fMRI analyses.

Table 4: Percentage of trials for Learning and Application (Mean (SD))

Strategy Group	Positive Learning	Negative Learning	Application	Other trials
Low	24.86 (2.66)	25.36 (3.93)	35.65 (3.77)	14.13 (4.81)
Medium-SO	29.98 (2.18)	21.33 (2.63)	42.82 (2.35)	5.87 (2.37)
Medium	29.23 (2.47)	19.92 (2.06)	42.93 (2.39)	7.91 (3.01)
High	32.87 (2.02)	17.46 (2.02)	46.85 (1.68)	2.82 (1.91)

To assess potential neural differences between strategy groups, we focused on the contrast Learning > Application. General age and main effects of feedback learning were also described in Peters et al. (2014). Here, we focused on the strategy versus age related differences in neural activation. Our hypothesis was that age effects for the contrast Learning > Application can partly be explained by strategy differences. To investigate this, we first calculated the contrast Learning > Application with age as a positive regressor, to see which areas were more active with increasing age (see Figure 4, Table 5).

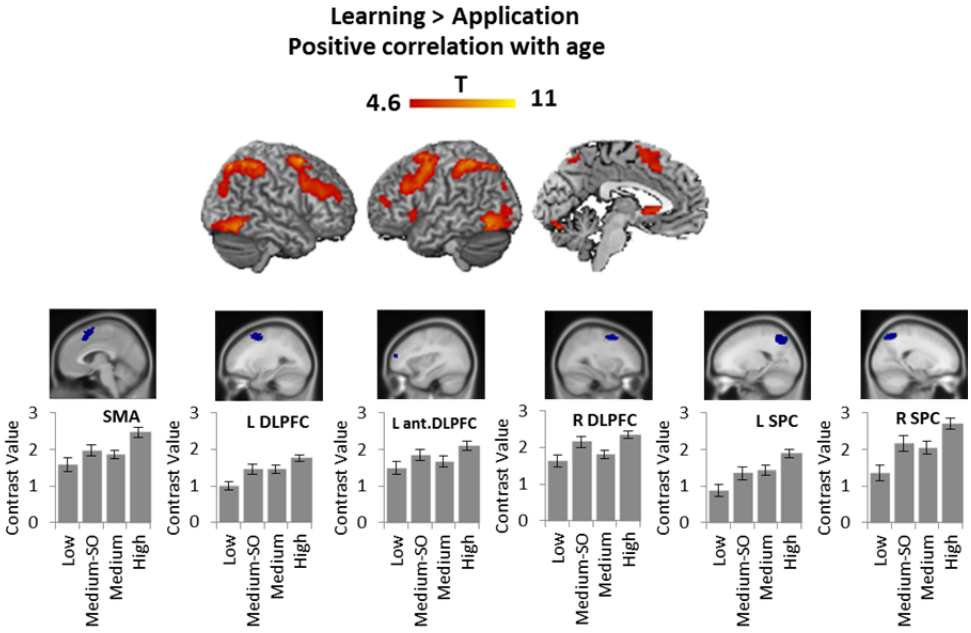


Figure 4: Areas showing activation for the contrast Learning > Application, with age as a positive regressor, FWE corrected at p < .05, > 10 contiguous voxels. Differential activity for Learning > Application in ROIs based on this contrast is shown per strategy group (SO = Suboptimal).

Table 5: Areas showing activation for the contrast Learning > Application, with age as a positive regressor, FWE corrected at p < .05, > 10 contiguous voxels.

Area of activation	x	y	z	voxels	T
R inferior parietal lobule	42	-39	48	2127	10.15
R superior parietal lobule	21	-69	54		9.36
L superior parietal lobule	-27	-63	54		9.07
R middle frontal gyrus	30	9	60	521	9.29
R inferior gyrus	48	12	33		6.57
R inferior gyrus	45	33	24		6.29
L inferior occipital gyrus	-48	-66	-15	407	8.54
L fusiform gyrus	-33	-51	-18		5.54
L middle occipital gyrus	-36	-84	6		5.24
L precentral gyrus	-36	0	60	732	8.24
L precentral gyrus	-48	3	51		7.68
L superior frontal gyrus	-21	12	63		7.37
R inferior temporal gyrus	51	-60	-15	292	8.06
R cerebellum	33	-69	-21		5.75
R inferior occipital gyrus	27	-93	-12		4.96
R caudate nucleus	6	21	0	260	7.00
L caudate nucleus	-9	9	0		6.79
R caudate nucleus	9	9	0		6.53
L cerebellum	-3	-81	-18	52	6.13
L lingual gyrus	-12	-93	-15		4.89
L inferior frontal gyrus	-30	30	0	65	5.58
L inferior frontal gyrus	-45	21	-3		5.42
L middle frontal gyrus	-36	54	15	30	5.50

Note: Abbreviations: L = Left; R = Right

From these functional activations we created ROIs. The resulting ROIs spanned several brain regions, therefore we applied anatomical masks (based on Marsbar Automated Anatomical Labeling) for the key regions implicated in the development of feedback learning (Peters et al., 2014; Crone et al, 2008; Van Duijvenvoorde et al., 2008)): left and right DLPFC (AAL mask: Middle Frontal Gyrus), pre-SMA/ACC (Supplementary Motor Area, left and right combined), and left and right SPC (Superior Parietal Gyrus). Centre-of-mass MNI (x y z) coordinates were: right DLPFC: x = 39, y = 22, z = 41; left DLPFC: x = -35, y = 12, z = 49; left anterior DLPFC: x = -35, y = 52, z = 14 ; pre-SMA/ACC: x = -4, y = 12, z = 58; right SPC: x = 27, y = -62, z = 55; left SPC: x = -24, y = -64, z = 50 (see Figure 4). (Note that the anatomical masking process resulted in two left DLPFC regions, one of which we called: 'left anterior DLPFC').

ROI analyses

To test if strategy groups differed in neural activity within these areas, we created ROIs based on this contrast (see Figure 4). For these ROIs, we first tested whether the four groups differed in neural response; then we tested if strategy group explained variance above age, and finally we tested with mediation analyses whether age effects were mediated by strategy group. Thus we computed difference scores for Learning > Application, which were different for the four strategy groups in all ROIs (all *p*s < .005; see Figure 4). The patterns for each region showed lowest activity for the Low strategy group, higher activity for the Medium and Medium-Suboptimal group, and highest activity for the High strategy group. LSD post-hoc tests for each region separately indicated that for pre-SMA/ACC, all groups showed lower activity than the High group, but there were no other differences (all *p*s < .005). For right SPC, we found that all groups except Medium and Medium-Suboptimal differed from each other (all *p*s < .024), and for left SPC, all groups except Medium and Medium-Suboptimal and Low and Medium-Suboptimal differed from each other (all *p*s < .014). We found that for right DLPFC, the Low and Medium-Suboptimal, Low and High, and Medium and High showed significant differences (all *p*s < .017); for left DLPFC that all groups except Medium and Medium-Suboptimal were different (all *p*s < .025).; and for left anterior DLPFC that the Low and Medium group showed less activity than the High group (all *p*s < .007). Note that the general pattern (see Figure 4) is that the Low group shows the least activity and the High group shows the most activity, which for the High group is consistently different from the other groups except in right DLPFC and left anterior DLPFC.

To investigate whether age differences in neural activity can be attributed to differences in strategy use (over and above age), we used hierarchical linear regression analyses with neural activity for Learning > Application (difference score) as dependent variable, age entered as first predictor and strategy group entered as second predictor. We found significant effects of strategy group above age for four of the seven ROIs: pre-SMA/ACC (step 1: R^2 = .15; age: *B* = .39, *p* < .001; step 2: R^2 = .18; age: *B* = .33, *p* < .001, strategy group: *B* = .16, *p* = .007), left DLPFC (step 1: R^2 = .22; age: *B* = .47, *p* < .001; step 2: R^2 = .24; age: *B* = .43, *p* < .001, strategy group: *B* = .13, *p* = .030), left SPC (step 1: R^2 = .23; age: *B* = .48, *p* < .001; step 2: R^2 = .25; age: *B* = .43, *p* < .001, strategy group: *B* = .14, *p* = .012) and right SPC (step 1: R^2 = .25; age: *B* = .50, *p* < .001; step 2: R^2 = .26; age: *B* = .45, *p* < .001, strategy group: *B* = .12, *p* = .033). To summarize, these results indicate that strategy explained additional variance in neural activity above age in pre-SMA/ACC, bilateral SPC and left DLPFC.

Mediation analyses

For the four ROIs that showed a significant contribution of strategy group above age, we performed mediation analyses with the R package for causal mediation analysis (Imai, Keele, & Tingley, 2010) to investigate the relation between age and strategy use in explaining variance in brain activity. We applied the analysis on ROI activity with age (continuous) as direct predictor and strategy group (nominal variable) as mediator variable (see Figure 5).

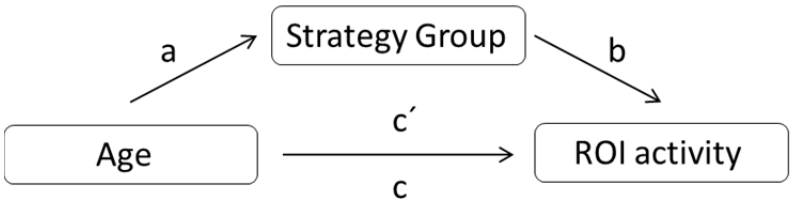


Figure 5: Mediation analysis paths for strategy group as a mediator between age and ROI activity.

With causal mediation analysis, a mediation effect is present if 1) age predicts strategy group (path a), 2) strategy group predicts ROI activity if age is simultaneously entered as a predictor (path b), 3) if age predicts ROI activity (path c), and 4) if strategy group is entered simultaneously as a predictor, the effect of age on ROI activity decreases (path c') (Preacher & Hayes, 2008). We first tested for collinearity problems between age and strategy group, which are indicated by a Variance Inflation Factor greater than 10 (Myers, 1990), a tolerance value less than 0.1 (Menard, 1995) and condition indices greater than 10 (Belsley, David, Kuh & Welsch 1980). For this data, we found no indication for collinearity problems (VIF = 1.14, tolerance = 0.86, condition indices < 10). For the mediation analysis, we report unstandardized regression coefficients. Path a (effect of age on strategy group) was the same for all four mediation analysis and resulted in a significant effect: $B = .10$, $p < .001$).

For pre-SMA/ACC, the effect of age on ROI activity (path c: $B = 0.13$, $p < .001$) was partly mediated by strategy group (path b: $B = 0.19$, $p = .007$; path c': $B = .11$, $p < .001$; mediation effect (ab) = .02, $p = .01$ (95 % confidence interval (CI) = .004 - .036); proportion mediated (ab/c) = .15). In left DLPFC, the effect of age on ROI activity (path c: $B = 0.13$, $p < .001$) was also partly mediated by strategy group (path b: $B = 0.12$, $p = .031$; path c': $B = .11$, $p < .001$; mediation effect (ab) = .01, $p = .04$ (95 % confidence interval (CI) = .0005 - .023); proportion mediated (ab/c) = .09). For right SPC, the effect of age on ROI activity (path c: $B = 0.21$, $p < .001$) was partly mediated by strategy group (path b: $B = 0.18$, $p = .033$; path c': $B = .19$, $p < .001$; mediation effect (ab) = .02, $p = .04$ (95 % confidence interval (CI) = .001 - .037); proportion mediated (ab/c) = .09). A similar effect was found for left SPC, where the effect of age on ROI activity (path c: $B = 0.17$, $p < .001$) was partly mediated by strategy group (path b: $B = 0.18$, $p = .012$; path c': B = .15, $p < .001$; mediation effect = 0.018, $p = .02$ (95 % CI:.004 –.034); proportion mediated = .11). In summary, we found that a significant portion of the variance in ROI activity in pre-SMA/ACC, left DLPFC and bilateral SPC was explained by strategy group, because strategy group partly mediated the relation between age and brain activity.

Whole brain analysis

In addition, we also tested if there were differences on a whole-brain basis between strategy groups, to investigate whether the ROI results were also observable on a whole-brain level, and to test whether there were effects of strategy group (above age) in areas outside of the ROIs. We used an ANOVA model with the four strategy groups as between-subjects variable and age as an additional regressor. We added age as a regressor because the strategy groups were not equally divided over age groups. First, we calculated the contrast Learning > Application for the four strategy groups separately, corrected for age (see Figure 6, Table 6). These analyses showed widespread activity (FWE-corrected at $p < .05$, > 10 contiguous voxels) for all strategy groups in the bilateral frontoparietal network, as well as in the pre-SMA/ACC, basal ganglia and occipital/temporal cortex. An F-test (with age-correction) was performed to see whether there was a main effect of strategy group. There were no significant clusters which survived FWE-correction.

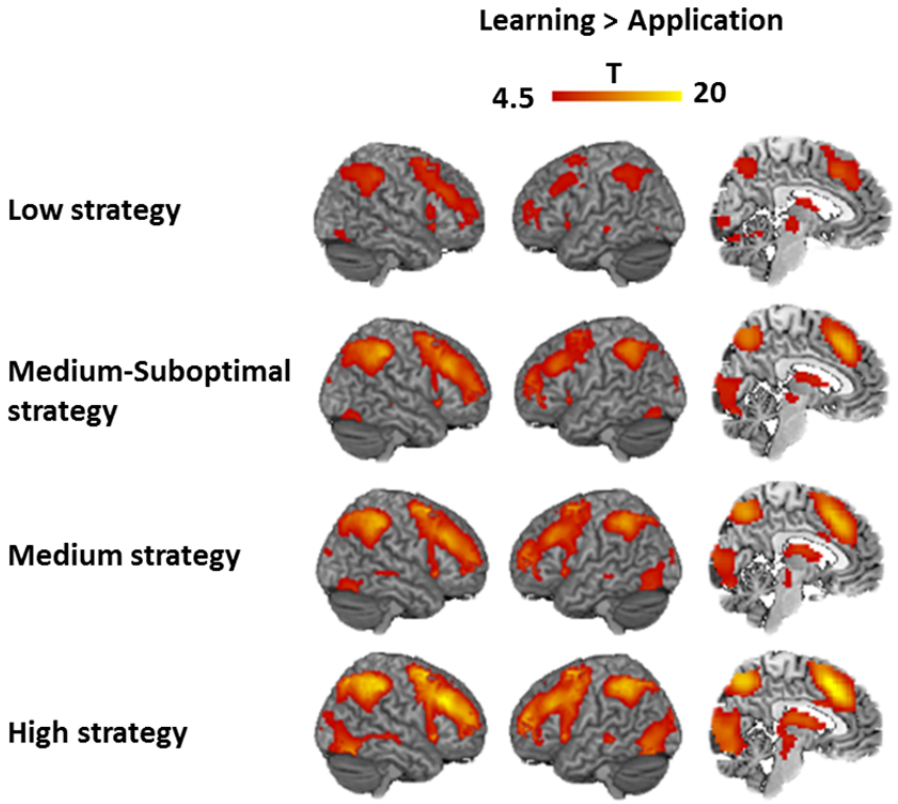


Figure 6: Areas showing activation for the contrast Learning > Application, separately for each strategy group (corrected for age), FWE corrected at p < .05, > contiguous 10 voxels.

Table 6: MNI coordinates local maxima activated for the contrast Learning > Application per strategy group, FWE corrected at p < .05, > contiguous 10 voxels

Strategy Group	Area of activation	x	y	z	voxels	T
Low	R inferior parietal lobule	45	-42	48	1778	10.39
	L inferior parietal lobule	-35	-45	42		7.54
	L inferior parietal lobule	-48	-39	45		7.39
	L superior medial gyrus	-3	24	45	2588	9.99
	R superior frontal gyrus	24	9	57		9.50
	R middle frontal gyrus	42	30	33		9.14
	R putamen	30	21	0	1018	3.06
	L insula lobe	-30	21	-3		8.43
	L caudate nucleus	-15	6	12		7.13
	L middle frontal gyrus	-30	51	18	277	6.93
	L middle frontal gyrus	-36	48	6		6.62
	L superior orbital gyrus	-30	54	-3		6.59
	R middle temporal gyrus	57	-30	-9	53	6.42
	L calcarine gyrus	-3	-90	-9	71	5.59
	R inferior occipital gyrus	42	-78	-12	36	5.16
	R fusiform gyrus	36	-75	-18		5.05
	R cerebellar vermis	6	-75	-18	21	5.11
	L cerebellum	-6	-78	-18		4.91
	R cerebellar vermis	3	-54	-15	27	5.04
Medium-SO	R inferior parietal lobule	45	-42	48	3087	17.34
	R angular gyrus	36	-57	45		16.69
	L inferior parietal lobule	-48	-42	48		12.67
	Superior medial gyrus	0	24	45	6258	16.17
	R putamen	30	21	0		14.77
	R middle frontal gyrus	45	27	36		14.67
	L superior occipital gyrus	-9	-99	6	972	9.09
	R cuneus	12	-96	12		8.21
	L fusiform gyrus	-39	-69	-18		7.78
	R middle temporal gyrus	57	-30	-9	18	6.45
Medium	R superior frontal gyrus	24	9	60	7784	18.16
	Superior medial gyrus	0	24	42		17.18
	R putamen	30	21	0		15.17
	R inferior parietal lobule	45	-42	48	5104	17.92

	L inferior parietal lobule	-48	-45	48		15.11
	R precuneus	6	-66	51		14.13
	L middle temporal gyrus	-57	-30	-9	16	6.28
High	R supplementary motor area	3	21	45	9890	22.07
	L supplementary motor area	-3	15	48		21.40
	L supplementary motor area	-6	9	54		20.57
	R supramarginal gyrus	39	-42	42	7336	21.61
	L inferior parietal lobule	-36	-48	45		18.52
	L inferior parietal lobule	-45	-42	45		18.52
	L middle temporal gyrus	-57	-33	-9	78	9.55

Note: Abbreviations: L = Left; R = Right

Age effects

Additionally, we investigated effects of age within our ANOVA model, to see which brain activation is related to age while controlling for strategy group. As can be seen in Figure 7, a positive relation with age (FWE-corrected, $p < .05$, > 10 contiguous voxels) was found in the frontoparietal network, pre-SMA/ACC, basal ganglia and occipital cortex (see also Supplementary Table 1), when controlling for strategy.

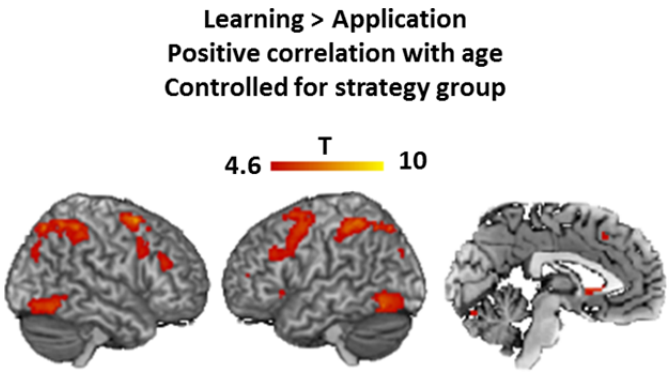


Figure 7: Areas showing activity for the contrast Learning > Application that correlated positively with age, while controlling for strategy group. Displayed at an FWE-corrected threshold at p < .05, > 10 contiguous voxels.

Discussion

An important question in research on executive functioning is whether developmental differences in behavior and neural activation can be explained by strategy differences. This study tested this hypothesis in a large sample of participants between ages 8-25 years in which statistical modeling approaches were combined with neuroimaging. The results showed that 1) learning from feedback in a sorting task resulted in variance in performance, which could be distinguished into four latent strategy groups, 2) even though age was strongly linked to strategy use, such that older participants were more often present in high performing strategy groups and young participants in low strategy groups, there was still considerable variance in strategy groups within age groups, 3) strategy explained additional variance in neural activity above age in pre-SMA/ACC, left DLPFC and bilateral SPC, and 4) there was still unique variance related to age differences in neural activity during feedback learning.

Underlying strategy groups in the feedback learning task

Consistent with prior developmental studies (Andersen et al., 2014; Raijmakers et al., 2001; Schmittmann et al., 2012, 2006), we found latent strategy groups that differed in performance on the feedback learning task. In the current feedback learning paradigm, participants were instructed to sort stimuli in one of three locations by using positive and negative feedback. Several strategies could be used to ensure optimal learning, such as focusing not only on feedback for the current stimulus, but also on feedback for the other two stimuli. In our sample, four different strategy groups could be distinguished. The highest performing strategy group responded almost perfectly in terms of the efficiency of strategy use. This strategy required optimal use of short-term memory, working memory and more complex executive functions (Diamond, 2013). The lowest performing strategy group, however, made more mistakes and did not adequately take into account the information from the other stimuli, thereby missing opportunities for learning. The lowest strategy seemed to involve regular flaws in STM, WM and more complex executive functions. The other two groups represented intermediate variants.

The division over strategy groups was related to age: Results indicated that younger children were more likely to be in a lower performing group, and older children and adults were more likely to be in a higher performing group. Still, there was considerable variation in strategy group within age groups: Some young children (8/9 years) belonged to the same strategy group as some of the adults, but there were also adults who were outperformed by young children. This supports the notion that it is important to study performance and strategy differences as opposed to age differences alone in learning tasks. The presence of different underlying strategies within age groups could make the interpretation of prior developmental studies that compared age groups difficult.

Neural differences between strategy groups

After discovering that different strategy groups could be distinguished in our sample, we investigated if these strategy groups explained developmental differences in neural activity. We used an fMRI contrast that revealed areas that are sensitive to feedback with learning value, i.e., areas that respond more to feedback that provides new information compared to feedback that provides information that is already known. We tested which areas showed a positive relation with age, and created ROIs based on this contrast for pre-SMA/ACC, DLPFC and SPC (see also Peters et al., 2014). We found that strategy group explained additional variance above age in pre-SMA/ACC, left DLPFC and bilateral SPC. In these areas, the highest performing group consistently showed more activation for these regions compared to the other groups. Moreover, with mediation analyses, we found that the proportion of the age effect on brain activity mediated by strategy was between 0.09 and .15. These mediated proportions were relatively low, especially compared to the study by Andersen et al. (2014) which used a similar approach but found stronger effects of strategy despite a much smaller sample size. This could be due to the fact that the task used in Andersen et al. was relatively more difficult, e.g., a majority of even the adult participants did not use the most optimal strategy. The relative contribution of age might be larger compared to strategy in a task such as in the current study where there is a clear developmental 'end point', i.e. most adults demonstrate optimal performance.

The finding that especially the highest performing group differed from the other groups in frontoparietal areas and the pre-SMA/ACC relates to prior studies which indicated that these areas are important for executive functioning. The pre-SMA/ACC is thought to be important to detect conflict between competing representations (Carter & van Veen, 2007) and for top-down control of response selection and preparation (Schulz, Bedard, Czarnecki, & Fan, 2011). The DLPFC and SPC are important for executive functions such as working memory (Klingberg et al., 2002), and the DLPFC is also associated with inhibition (Nyffeler et al., 2007) and cognitive flexibility (Ravizza & Carter, 2008). These executive processes are important to achieve the highest level of performance in the current feedback learning paradigm. Possibly, these fMRI results indicate that high performing learners are better at distinguishing feedback that is important for learning compared to relatively uninformative feedback (i.e., feedback during the application phase, which does not provide new information).

In addition, a whole-brain approach showed that the feedback learning network (pre-SMA/ACC, DLPFC, SPC and basal ganglia) was activated in all strategy groups (after age correction) while participants received feedback during learning compared to applying known rules. These findings are consistent with prior studies that showed that this network is implicated in feedback learning across development (Crone et al., 2008; van den Bos et al., 2009; van Duijvenvoorde et al., 2008). However, when testing for differences between the groups while correcting for age, there were no neural differences which survived correction for multiple comparisons. Together these results showed that age effects in pre-SMA/ACC, SPC and left DLPFC

during a feedback learning task are partly explained by strategy differences, although the effects of strategy may be relatively small compared to the effects of age.

Age versus strategy effects

Because of the relatively weak mediation effects of strategy on brain activity, we also investigated which brain areas show a positive correlation with age, while correcting for strategy group, to find unique activation related to age. We found widespread activation in the frontoparietal network, pre-SMA/ACC and basal ganglia, which survived correction for multiple-comparisons. Together with the ROI findings, these results suggest that age is the most significant contributor to neural activation patterns compared to strategy use in this study. This is consistent with the previously mentioned maturational viewpoint which suggests that age is a vital contributor to the development of patterns of neural activation (Dosenbach et al., 2010).

Therefore, our results only partly correspond to prior studies which found that age-related differences in neural activation can be explained by performance differences (Booth et al., 2004; Bunge et al., 2002; Koolschijn et al., 2011). For instance, in a longitudinal study, when performance and age were used as predictors for neural activation change in regions such as the DLPFC, SPC and pre-SMA/ACC, performance was a better predictor for changes in neural activity than age (Koolschijn et al., 2011). This highlights the need for longitudinal studies to investigate the development of learning and the underlying neural processes. Future research should investigate if participants progress to faster strategy groups with development, and if this is accompanied by more robust changes in neural activity than in this cross-sectional design. We propose that latent variable models provide a valuable method to detect performance-related versus age-related influences on neural activity.

Limitations

There are several limitations to this study. First, as is often the case in developmental studies, children received relatively more negative feedback compared to adults (e.g., Koolschijn et al., 2011). It is possible that our results were influenced by these differences in the amount of trials per feedback type, although these differences between age groups were relatively small. Second, even though the wide age range in this sample has many benefits, it was not possible to investigate effects of strategy in smaller age ranges, due to the unbalanced division of participants across strategy groups. A study with a similarly large sample but with less variation in age would be better suited to investigate effects of strategy within smaller age ranges. In addition, the study was cross-sectional and future studies should test changes in strategy use and neural activation patterns within the same individuals, to examine whether switching to a different strategy is accompanied by neural change within individuals. Finally, since we did not fit confirmative process models to the behavioral data, we cannot exclude the possibility that performance was not only organized in qualitatively different performance groups, but that there was also some con-

tinuous performance variation. However, there are good reasons to believe that qualitatively different strategy groups played an important role in performance differences of individuals. First, model comparisons indicated that the fit of a model that presumes continuous variation was relatively worse. Second, in the selected model the posterior probabilities to belong to a specific strategy group was for most participants either high (around 1.0) or low (around 0.0). Finally, strategy differences (as opposed to continuous performance variation) were found by studies that did fit confirmative models for feedback learning processes in other learning tasks (Raijmakers et al., 2001; Schmittmann et al., 2012, 2006).

Conclusion

In this study, we showed that in a feedback learning task, different underlying strategies could be detected within age groups, which were distinguishable at the neural level. These findings have important implications for traditional ways of analyzing developmental data. In future studies, it will be important to take into account individual differences in performance and strategy use, rather than comparing age groups alone. This research is informative in the context of unraveling the mechanisms underlying learning and learning difficulties and may contribute to interventions teaching children to adapt more efficient strategies to enable faster learning.

Chapter 6

Longitudinal development of the frontoparietal network: Contributions of age, performance, working memory and brain structure

This chapter is based on:

Peters, S., Van Duijvenvoorde, A.C.K., P.C.M.P. Koolschijn & Crone, E.A. Longitudinal development of neural activity in the frontoparietal network: Contributions of age, performance, working memory and brain structure (in revision, 2015).

Abstract

Even though it is well conceptualized that neural activity in the frontoparietal network changes during childhood and adolescent development, there is surprisingly little consensus about the direction of change, and a comprehensive study is lacking. Using a large-scale longitudinal fMRI study, we aimed to test growth patterns across childhood and adolescence in frontoparietal activity during a feedback learning task capturing multiple aspects of cognitive control. Our first aim was to test for linear and non-linear developmental trajectories of activity in the dorsolateral prefrontal cortex (DLPFC), superior parietal cortex (SPC) and the pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC). Our second aim was to test which factors drive developmental change in the frontoparietal network besides age. Contributions of task performance, working memory and cortical thickness were investigated. To these ends, a developmental sample (N = 208, 8-27 years old) was tested twice across a period of two years. The results showed that developmental patterns for neural activity in DLPFC and SPC were best characterized by a quadratic age function leveling off/peaking in late adolescence, and by a linear increase across age in pre-SMA/ACC. In addition to age, task performance explained variance in DLPFC and SPC activity, but not in pre-SMA/ACC activity. In contrast, cortical thickness explained additional variance in pre-SMA/ACC activity, but not in DLPFC and SPC. Together, these findings provide a novel perspective of developmental changes in the frontoparietal network, arguing against a simple imbalance model with linear development of cognitive control regions.

Introduction

Although adolescent development of neural activity during cognitive control is studied extensively, results from cross-sectional studies remain inconclusive. A meta-analysis (Crone & Dahl, 2012) showed that, although many fMRI studies reported age-related increases in prefrontal and parietal recruitment during cognitive tasks, a substantial number of studies reported age-related decreases in these brain regions. Until recently most studies only used cross-sectional comparisons to characterize development, which has several disadvantages compared to longitudinal designs. Here, we used a large longitudinal dataset to test growth patterns across adolescent development. Specifically, we tested whether frontoparietal activity follows a linear pattern (i.e. monotonic development over time, no adolescent-specific changes), a quadratic pattern (i.e., adolescent-specific effects) or a cubic pattern (adolescent-emergent; e.g. stable levels during childhood, steep changes in adolescence and stabilization in adulthood) (Braams et al., 2015; Somerville et al., 2013). A second goal was to test which other factors influence changes over time in frontoparietal activity in addition to age. Here we tested the contributions of increased task performance, working memory capacity and cortical thickness.

Cognitive development has been hypothesized to follow a linear developmental trajectory according to 'dual-systems models' of adolescent brain development, with steadily increasing frontoparietal recruitment from childhood to adulthood (Ernst et al., 2006; Somerville & Casey, 2010; Steinberg, 2008). Several authors have argued for adaptations to dual-systems models, by suggesting that the frontoparietal network is not necessarily immature or inaccessible in adolescents, but sensitive to different situations than adults depending on e.g. motivational salience, or increased specialization of brain areas for more specific tasks (Casey, 2015; Crone & Dahl, 2012; Johnson, 2011; Pfeifer & Allen, 2012). This is for instance supported by results from feedback learning paradigms which demonstrated that adults showed increased recruitment of prefrontal and parietal cortex following negative compared to positive feedback, whereas children showed more activity in these same regions for positive compared to negative feedback (Peters, Braams, et al., 2014; van den Bos et al., 2009; van Duijvenvoorde et al., 2008). These findings indicate that young adolescents are capable of recruiting frontoparietal regions but in different situations than adults, arguing against a simple frontoparietal immaturity model with linear development in cognitive control regions.

The current study builds on prior cross-sectional findings by testing for linear and non-linear development in neural activity during feedback learning using an experimental paradigm inspired by the Wisconsin Card Sorting Test (Milner, 1963). We designed this task to capture multiple aspects of cognitive control functioning and we focused on three regions within the frontoparietal network, dorsolateral prefrontal cortex (DLPFC), pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC) and superior parietal cortex (SPC). Prior studies revealed that these are key regions for cognitive control, as demonstrated by meta-analyses across

multiple executive functions (Kim et al., 2012; Niendam et al., 2012). In addition, our rationale for focusing on these three regions is based on findings showing that DLPFC, pre-SMA/ACC and SPC show pronounced age-related changes in activity during cognitive control tasks such as working memory (Klingberg et al., 2002; Thomason et al., 2009) and feedback learning paradigms (Crone et al., 2008; Peters, Braams et al., 2014; van Duijvenvoorde et al., 2008). However, to date little is known about how activity in the frontoparietal network changes longitudinally.

Longitudinal designs have critical advantages over cross-sectional designs. For instance, previous studies demonstrated important individual differences in developmental trajectories that can be overlooked in cross-sectional designs (Koolschijn et al., 2011; Ordaz, Foran, Velanova, & Luna, 2013; Shaw et al., 2013). Furthermore, longitudinal designs have increased power to detect developmental change, because testing within-individual changes reduces error related to cohort differences (Fjell et al., 2010; Koolschijn et al., 2011).

Besides investigating age-related patterns of neural activity, a second goal of this study was to investigate other factors influencing time-related changes in frontoparietal activity in addition to age. There are multiple processes closely related to advancing age that may drive changes in neural activity. That is, an increase in age could be the only factor explaining time-related increases or decreases in activity, but other factors might also play a role. The factors investigated in this study were task performance, working memory capacity and structural brain development. Task performance has been shown to influence neural activity, and there is evidence that a portion of developmental changes attributed to advancing age are related more to increases in performance (Church et al., 2010; Koolschijn et al., 2011). Here we tested whether performance on the feedback-learning task partly explained changes in neural activation over time. Working memory has previously been argued to be a core prerequisite for cognitive development (Case, 1992) and cognitive control functions (Huizinga et al., 2006), and as such was investigated as an important contributor to changes over time in neural activity during feedback-learning. That is, we aimed to study whether a portion of changes in neural activity during feedback learning was explained by individual differences in working memory capacity. A final factor that was investigated is cortical thickness. Several cross-sectional studies have suggested a link between functional activity and structural gray matter in adults (Harms, Wang, Csernansky, & Barch, 2013; Hegarty et al., 2012) and children (Lu et al., 2009; Wendelken, O'Hare, Whitaker, Ferrer, & Bunge, 2011). It is likely that developmental changes in neural activity are at least partly influenced by structural development of these brain regions, although the longitudinal relation between structural maturation and development of brain function is not well understood.

Taken together, in this study, we tested developmental trajectories of activation in the frontoparietal network in a large longitudinal fMRI sample across a wide age range (N = 208, 8-27 years) with a two year interval between the first and second time point. Our aims were 1) to examine growth trajectories of core areas in the frontoparietal network (DLPFC, pre-SMA/ACC and SPC) and to define the shape of age-related changes, 2) to test the additional contributions of task

performance, working memory capacity and structural development to changes over time in neural activity.

Methods

Participants

At time point 1 (TP1), a total of 299 participants between ages 8-27 years underwent an MRI scan, of which 293 participants completed the feedback learning task in the MRI scanner. Of these, 25 participants were excluded from further analyses because of excessive movement (movement > 3.0 mm: *n* = 19), artifacts (*n* = 3) or because they were extreme outliers in task performance (> 3x the interquartile range: *n* = 3). In total, 268 participants were included at TP1 (*Mean* Age = 14.52 years, *SD* = 3.55; published in Peters, Braams et al., 2014). At time point 2 (TP2), a total of 254 of the initial 299 participants were scanned again approximately two years later (mean time = 1.99 years, SD = 0.10 years, range = 1.66-2.47 years). Reasons for not collecting a scan at TP2 (*n* = 45) were braces (*n* = 32) or no interest in participating again (*n* = 13). Further exclusions at TP2 were because of excessive movement at TP2 (*n* = 9), scanner artifacts (*n* = 5), loss of signal (*n* = 3) or extreme outliers (> 3x the interquartile range) on task performance (*n* = 2).

Only those participants who were included at both TP1 and at TP2 were included in the analyses (*N* = 208). All analyses were performed on these 208 participants, except for the analyses including working memory capacity and cortical thickness. For working memory, data were incomplete for five participants at TP1 and for two participants at TP2. For the analyses involving structural MRI data, visual quality control led to exclusion of 28 out of 208 participants: Three exclusions for insufficient quality data at both TP1 and TP2, 16 for TP1 and nine for TP2. These participants were only excluded from the analyses where cortical thickness was a factor. Taken together, the analyses with fMRI in the model contained a total of 208 participants (105 females), the analyses with working memory a total of 201 participants and the analyses with structural MRI in the model contained a total of 177 participants.

IQ was estimated with two subtests of the WAIS-III or WISC-III (Similarities and Block Design at TP1, Vocabulary and Picture Completion at TP2). The estimated IQ-scores of the 208 included participants were within the normal range at TP1 (85-143, *Mean* = 110.91, *SD* = 9.74) and TP2 (80-147, *Mean* = 108.92, *SD* = 10.18). The study was approved by the Institutional Review Board at the Leiden University Medical Center and all participants (or participant's parents in case of minors) provided written informed consent. Adults received payment for participation and children and their parents received small presents and payment for participation. Participants did not report psychiatric or neurological diagnoses, and no current use of psychotropic medication. All anatomical MRI scans were reviewed and cleared by a radiologist.

Feedback Learning Task

Participants performed a child-friendly feedback learning task in the MRI scanner described in detail earlier (Peters, Braams, et al., 2014; Peters, Koolschijn, Crone, Van Duijvenvoorde, & Raijmakers, 2014). In short, on each trial, participants viewed a screen with three boxes at the top part of the screen (Figure 1a). At the bottom part of the screen, a stimulus picture was presented, which was one of three possible stimuli. Participants were informed that all pictures belonged in one of the three boxes and that they had to find the correct box for each picture. Performance feedback was provided in the form of a plus-sign ('+') for correct choices (positive feedback) and a minus-sign ('-') for incorrect choices (negative feedback). Stimuli were presented in a pseudorandom order (maximum two identical pictures in a row). The sequence ended after 12 trials, or when participants chose the correct location twice for all three stimuli. Subsequently, a new sequence with three new pictures was presented. In total, participants completed 15 sequences, which resulted in a maximum of 180 trials. The task was divided into two blocks of eight and seven sequences, respectively. Before performing the task in the MRI scanner, participants practiced three sequences in a separate practice session. All trials started with a 500 ms fixation cross, followed by a 2500 ms time window during which the stimulus was presented and a response needed to be given. Feedback was presented for 1000ms. Inter-trial intervals were jittered with OptSeq (Dale, 1999), with intervals between 0-6 seconds in addition to the 500 ms fixation cross.

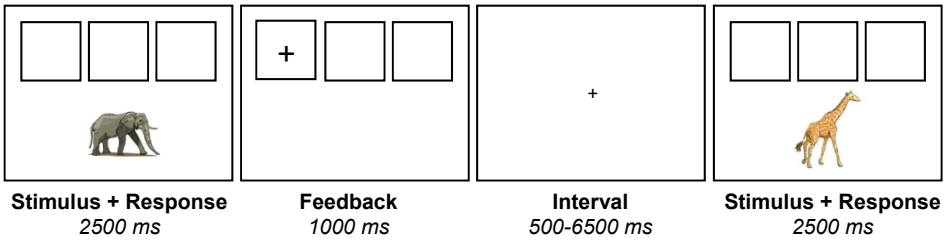


Figure 1a: Display of task sequence for the feedback learning task. During the last 500 ms of the Interval screen, a fixation cross was presented to prepare the participant for the next upcoming stimulus.

Feedback types

We distinguished between a learning phase and an application phase for all stimuli. The learning phase was defined as the trials where participants had not yet found the correct location for each stimulus and were still using feedback to find the correct locations. The application phase was defined as the trials where each stimulus was sorted correctly previously and was continued to be sorted correctly on subsequent trials. We excluded trials in the learning phase that did not result in learning, i.e., the trials where the feedback was not successfully used on the subsequent trial

(5.71 % of the trials). Based on the learning-application distinction, we defined the following three feedback types: Learning phase. a) Positive Learning: A first correct feedback for a stimulus followed by a correct sort on the next trial for this stimulus. b) Negative Learning: A first encountered incorrect feedback for a stimulus followed by a choice for another location on the next trial of this stimulus. Application phase. c) Application: Correct (i.e., positive) feedback for a stimulus that was sorted correctly before.

As a task performance measure, we calculated the 'learning rate' for each participant. This was defined as the percentage of trials in the learning phase for which feedback was successfully used on the next trial, compared to the total number of trials during the learning phase (including trials which did not result in learning according to the participants' behavior on the next trial).

Working Memory task

Working memory was assessed with the Mental Counters task (Larson, Merritt, & Williams, 1988), which has been shown in a prior developmental study to be a well-suited task to measure the latent factor of working memory in children (Huizinga et al., 2006). Similar to the feedback learning paradigm, the mental counters task has both a spatial aspect (because two counters at two different locations have to be remembered) and a verbal rehearsal component, allowing us to assess the working memory aspects of the feedback learning task contributing to neural activity. Participants were presented with a screen with two horizontal lines (the 'counters') placed next to each other (Figure 1b).

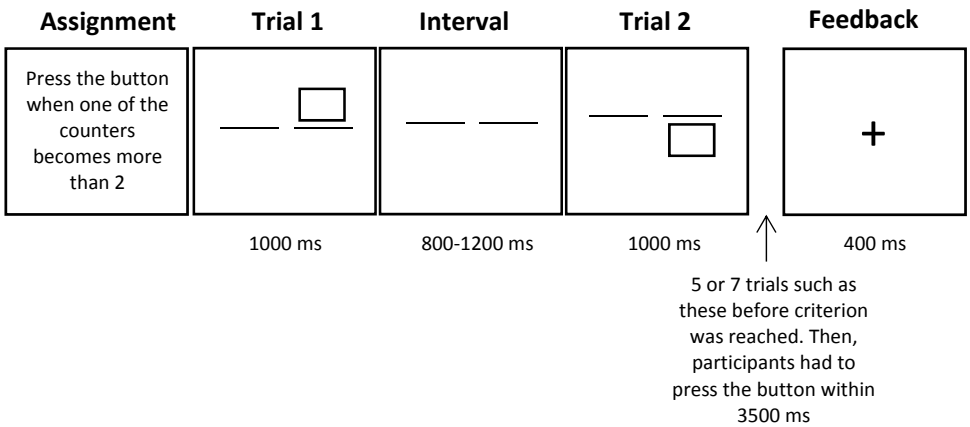


Figure 1b: Display of task sequence for the Mental Counters task as a measure of working memory capacity.

At each trial, a square randomly appeared on top of or below one of the two horizontal lines. The participant was instructed to keep track of the 'score' of the two counters. The value of the counters changed on each trial: e.g., a square appearing above the left counter changed the score to 1-0. If on a next trial a square appeared above the right counter, the score changed to 1-1. A square appearing below one of the counters meant a point had to be subtracted for that counter. Participants were instructed to press a button when one of the counters reached a certain criterion value, e.g. 'press when the score for one of the counters reaches more than two points', and the criterion changed for each series. The amount of trials before criterion was reached was set to either five or seven trials. In total, 16 series were presented. Trials were separated by 800-1200 ms intervals and participants had a time window of 3500 ms to respond once criterion was reached. Feedback was provided as a '+' for a correct button press, a '-' for an incorrect button press and an 'x' for an omission. Performance on the working memory task was defined as the proportion of correct responses (at TP1: $M = 0.83$, $SD = 0.15$; at TP2: $M = 0.87$, $SD = .12$).

FMRI Data Acquisition

We used the same Philips 3.0 Tesla MRI scanner and settings for both time points (Peters, Braams, et al., 2014; Peters, Koolschijn, et al., 2014). Functional scans were acquired with T2*-weighted echo-planar imaging, for which the first two volumes were discarded to allow for equilibration of T1 saturation effects. The following settings were used: TR = 2.2 s, TE = 30 ms, sequential acquisition, 38 slices, slice thickness = 2.75 mm, Field of View (FOV) = 220 x 220 x 114.68 mm. We acquired a high-resolution 3D T1-FFE anatomical scan after the experimental task (TR = 9.76 ms, TE = 4.59 ms, 140 slices, voxel size = 0.875 mm, FOV = 224 × 177 × 168 mm). The experimental task was projected on a screen that was viewed through a mirror attached to the head coil. Participants were accustomed to the MRI environment and sounds with a mock scanner prior to the actual MRI scan.

FMRI Data Analysis

We performed two types of analyses: a whole-brain analysis for an illustrative overview of brain activity at TP1 and TP2, and regions-of interest (ROI) analyses for growth curve modeling. For all analyses we used SPM8 (Wellcome Department of Cognitive Neurology, London) to analyze fMRI data. All scans were corrected for slice timing acquisition and rigid body motion. All volumes were spatially normalized to T1 templates, using a 12-parameter affine transform with a nonlinear transformation involving cosine basis functions with resampling of the volumes to 3 mm voxels. T1 templates were based on the MNI305 stereotaxic space (Cocosco et al., 1997), an approximation of Talairach space (Talairach & Tourneaux, 1988). Functional volumes were spatially smoothed with an 8 mm FWHM isotropic Gaussian kernel. The fMRI time series data were modeled by convolving a series of events with a hemodynamic response function. The modeled feedback events were categorized as: "Positive Learning", "Negative Learning", and "Applica-

tion", which were time-locked with 0-duration to the moment of feedback presentation. All other trials (e.g., trials that did not result in learning or too-late trials) were modeled as events of no interest. These events were used as covariates in a general linear model together with a set of cosine functions that high-pass filtered the data. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used in pair-wise contrasts.

The main fMRI contrast was the learning contrast: Learning (Positive and Negative combined) > Application. With this contrast, we investigated neural activity for feedback that is informative for learning, averaged across negative and positive valence, relative to application. The contrast images were submitted to higher-level group analyses. Whole-brain fMRI analyses were performed with an FWE-corrected threshold at $p < .05$.

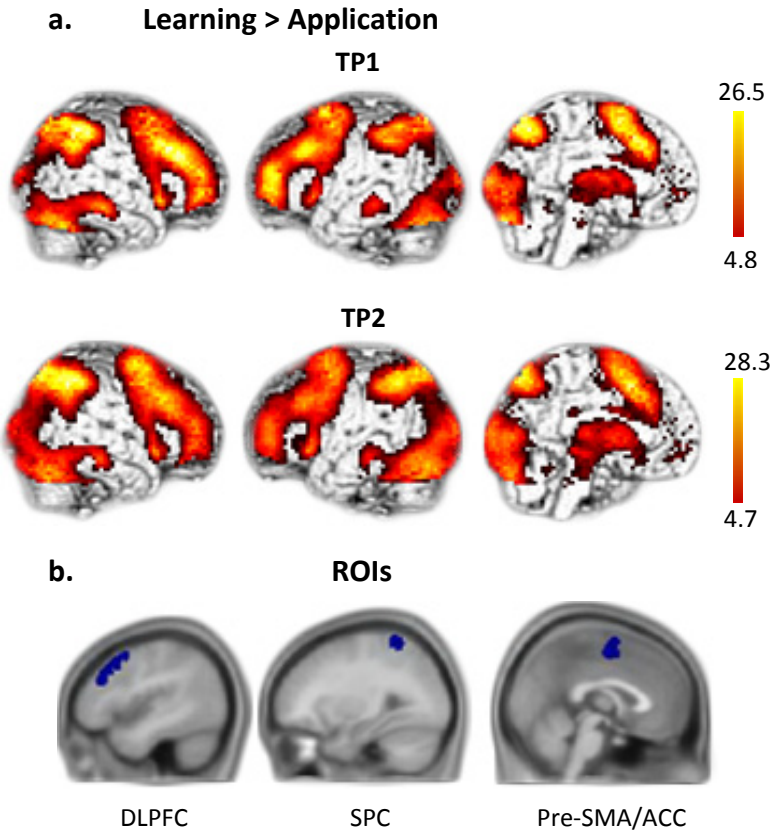


Figure 2a: Whole-brain analyses showing comparable neural activation patterns at TP1 and TP2 (FWE-corrected at p < .05). 2b: Bilateral Regions of Interest in the DLPFC, SPC and pre-SMA/ACC, extracted from the Harvard-Oxford Cortical Atlas. The anatomical ROIs were situated within the activation maps of the Learning > Application contrast.

Region-of-Interest analyses

Region-of-interest (ROI) analyses were performed with the MarsBaR toolbox (v. 0.42) in SPM8 (Brett et al., 2002). We used bilateral anatomical ROIs which were obtained from the probabilistic Harvard-Oxford Cortical Structural atlas. The ROIs were Middle Frontal Gyrus for DLPFC, Superior Parietal Lobule for SPC, and Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex) for pre-SMA/ACC. Because the anatomical ROIs based on probability maps were large and to ensure the selected voxels had a high probability of belonging to the targeted ROI, we created more focal ROIs by thresholding at 50 %, indicating that for each voxel, the probability that the voxel was actually part of the ROI (e.g. DLPFC) was > 50 % (see Figure 2b). These anatomical ROIs fell within activity clusters for the contrast Learning > Application (FWE-corrected at $p < .05$) (see Figure 2a). Beta values reflecting activity for all voxels within each ROI were averaged to produce a mean signal for each ROI per time point.

Structural brain analysis

Cortical reconstruction was measured automatically using FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu). Structural brain maturation of TP1 was reported earlier using a prior version of FreeSurfer (Koolschijn, Peper, & Crone, 2014), therefore we reconstructed and reanalyzed all anatomical scans from the first wave using version 5.3. Details of the surface-based cortical reconstruction and subcortical volumetric segmentation procedures have been extensively documented previously (Dale et al., 1999; Fischl & Dale, 2000; Ségonne et al., 2004).

To extract reliable volume and thickness estimates, images where automatically processed with the longitudinal stream in FreeSurfer (Reuter, Schmansky, Rosas, & Fischl, 2012). Specifically, an unbiased within-subject template space and image (Reuter & Fischl, 2011) is created using robust, inverse consistent registration (Reuter, Rosas, & Fischl, 2010). Several processing steps, such as skull stripping, Talairach transforms, atlas registration as well as spherical surface maps and parcellations were then initialized with common information from the within-subject template, significantly increasing reliability and statistical power (Reuter et al., 2012).

To extract average cortical thickness from the FSL anatomical ROIs, we performed the following steps: 1) Each anatomical ROI (DLPFC, pre-SMA/ACC and SPC) was registered automatically to the FreeSurfer "fsaverage" template with normalized mutual information and inspected for accuracy of registration. Of note, as FreeSurfer calculates cortical thickness per hemisphere, the ROIs were split into a left and right structural ROI. 2) Individual cortical thickness data was mapped to the "fsaverage" template. 3) Average cortical thickness in mm was extracted for each ROI and individual separately. 4) For subsequent analyses, all bilateral ROIs, including pre-SMA/ACC, were averaged to a single ROI as results were comparable between hemispheres. We used an average weighted procedure by taking into account hemispherical differences in surface size maps.

Statistical analyses

To model the shape of individual growth curves, we used mixed model analyses (also termed "random effects", "multilevel modeling", or hierarchical linear model-analyses) on the ROI values for contrasts of interest (Ordaz et al., 2013). This method expands on multiple regression analyses and is suited for longitudinal data because it takes into account the repeated-nature of the data, and controls for the dependency in measures within individuals (i.e., nested data). It was not necessary to calculate change scores, because mixed models take into account all data including individual differences in intercepts. We performed mixed analyses with the NLME package in R (Pinheiro, Bates, DebRoy, & Sarkar, 2007) version 3.1.0. With this package it is possible to test for fixed effects (effects that are similar for all participants) and random effects (effects that vary across participants) of age on brain activity. Models were compared using the Akaike Information Criterion (AIC), a standard measure for model comparison which indicates how well the model describes the data. Lower AIC values indicate a better fit of the model to the data. For nested models (i.e., comparing models with only 1 different term), we additionally tested with log-likelihood tests (χ^2) whether changes in model fit were significant. Our goals were twofold: 1) to test which shape of age (linear, quadratic cubic) best described the developmental pattern for the variables neural activity, task performance, working memory capacity and cortical thickness, and 2) to investigate which factors explain variance in brain activity above age; task performance, working memory capacity and/or cortical thickness. The model-building steps are described in the next paragraphs.

Developmental patterns: linear, quadratic and cubic trajectories

We first tested for each dependent variable (task performance, working memory capacity and activity and cortical thickness for each ROI) which age shape best described the developmental pattern. The base-model included a fixed intercept and a random intercept, with the latter capturing the variation in the intercept to account for the repeated nature of the data. Next, the base-model was tested against three models that tested the shape of the grand mean trajectory for age. We tested for a linear effect of age (i.e., monotonic development), a quadratic effect of age (i.e., an adolescent-specific effect) and a cubic effect (i.e., adolescent-emergent pattern) by adding three polynomial functions for age to the base-model (Braams et al., 2015; Somerville et al., 2013). We only selected a linear, quadratic or cubic model if the age term resulted in a better fit compared to the base-model without age as indicated by the AIC and a log-likelihood test. For the best model, we tested whether specifying age as an effect with a random slope resulted in a better fit (judged by the AIC and a log-likelihood test) compared to age as a fixed effect. A significant random slope would indicate that the effect of age on the dependent measure differs for each individual. An example of a formal notation of such a mixed model to predict, for instance, task performance, would be as follows:

$$\text{Performance} = \pi_{0i} + \pi_{1i}(\text{Age})_{ti} + e_{ti}$$

$$\text{With} \quad \pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

In this example model, substitution of the second level model into the first level model gives the integrated model that was fitted to the data. β_{00} reflects the grand mean intercept of neural activity at the average age of the sample. β_{10} reflects the grand mean slope of age effects on neural activity, and eti represents the residual error term. This example model displays a random intercept (r_{0i}) indicating different starting points of development for each participant, and a random slope of age (r_{1i}), indicating individual variability in the change over time. All models were fit with full information maximum likelihood estimates. A random slope did not improve model fit in any of the analyses we performed, except for the quadratic age effect on working memory capacity ($p < .001$). Therefore, we did not describe this further in the results section.

Explaining development of neural activity with different predictors

The second aim was to investigate which factors explain development of neural activity in the frontoparietal network in addition to age. The predictors we tested for significance above age were task performance, working memory capacity and cortical thickness. Therefore, we tested a combined model including all predictors to account for neural activity in the three ROIs. The model-building procedure consisted of multiple steps. We started with the best fitting age-model (linear, quadratic or cubic) for the ROI determined in the previous analysis. To test wether other measurements explained additional variance above age, we first added task performance as a second predictor to investigate whether performance explained additional variance above age. If this was the case, performance was included in the next step and if this was not the case, we continued to the next step with only the linear, quadratic or cubic age term. The next step was adding working memory capacity the model, followed by cortical thickness of the ROI as a final step. AIC and log-likelihood values were used to test whether model fit improved by adding the predictor. Changing the order of adding performance, working memory and cortical thickness did not change the overall results.

Reliability

For visualization purposes, the whole-brain results for the Learning > Application contrast for TP1 and TP2 across all participants are displayed in Figure 2a and Table 1. The results showed that a comparable network was recruited on TP1 and TP2, including bilateral DLPFC, pre-SMA/ACC, and bilateral SPC.

Table 1: MNI coordinates of local maxima activated for the contrast Learning > Application averaged across the two measurements, across all participants, FWE-corrected at p < .05.

Area of activation	x	y	z	voxels	T
R inferior parietal lobule	45	-43	49	22000	33.64
R supplementary motor area	3	17	49	s.c.	31.87
R middle frontal gyrus	45	29	34	s.c.	31.79
R superior frontal gyrus	27	2	58	s.c.	30.33
R angular gyrus	33	-58	49	s.c.	28.34
Supplementary motor area	0	11	55	s.c.	27.65
L supplementary motor area	-3	8	58	s.c.	27.17
R insula lobe	33	23	1	s.c.	26.91
R precuneus	6	-70	55	s.c.	26.60
R precuneus	9	-73	58	s.c.	26.13
L middle frontal gyrus	-45	23	37	s.c.	25.93
L insula lobe	-30	20	4	s.c.	25.74
L middle frontal gyrus	-48	26	34	s.c.	25.68
L inferior parietal lobule	-48	-43	46	s.c.	25.46
L inferior parietal lobule	-42	-52	55	s.c.	25.04
R middle frontal gyrus	33	53	13	s.c.	25.01
R posterior cingulate cortex	3	-31	25	26	7.21
Cerebellar vermis	0	-52	-8	13	5.68

Abbreviations: L = Left; R = Right; s.c. = same cluster

For all predictors (ROI activity and cortical thickness, task performance and working memory capacity), we also assessed reliability in ROIs from TP1 to TP2.

We calculated intra-class correlation coefficients (ICCs) on the mean signal of each ROI. We used a two-way mixed model with absolute agreement and we reported the average measure. A value of 0 indicates no relation between the first and second time point and a value of 1 indicates perfect agreement. Interpretation of ICC values for reliability was guided by Cichetti (2001): values < .4 were interpreted as poor; values .41-.59 were interpreted as fair, values .60-.74 were interpreted as good, and values > .75 were interpreted as excellent.

As can be seen in Figure 3, most ICC values were in the 'fair to good' range. For a visual comparison of ICCs across different age groups, Figure 3 also shows ICC values for 3 age groups (8-12 years, 13-16 years, 17-25 years). For DLPFC, reliability was lowest (in the poor range) in the oldest age groups (suggesting change over time), and was fair for the children and adolescents. In contrast, for SPC and pre-SMA/ACC, reliability was lowest in the child group (suggesting change over time), and was fair to good for adolescents and adults.

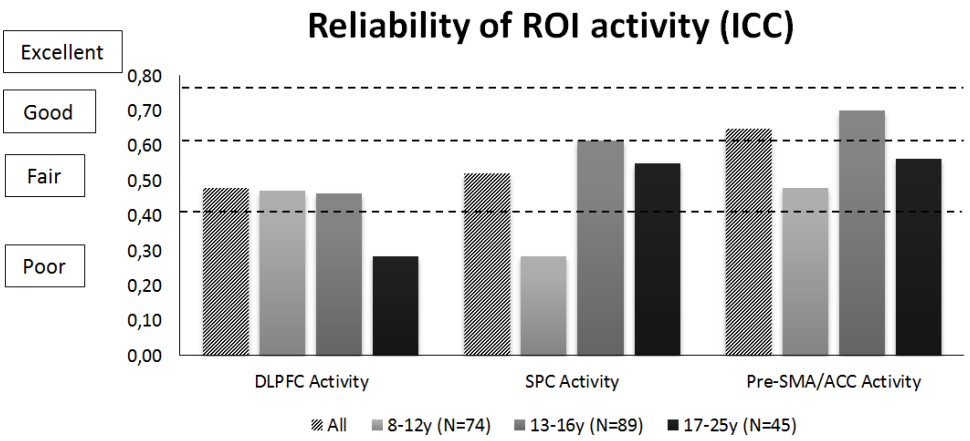


Figure 3: Intra-Class Correlations (ICC) values for each ROI per age group. The labels 'poor', 'fair', 'good' and 'excellent' are based on Cicchetti (2001).

Results

Developmental trajectories

For descriptive purposes, correlations between measures are presented in Table 2. Next, we started the model building procedure. As a first step, we tested for all measures (neural activity, task performance, working memory capacity and cortical thickness) whether a linear, quadratic or cubic age pattern best described developmental change. AIC values and a log-likelihood test were used to test which model best fit the data. AIC values for the base model (without age), linear, quadratic and cubic age model for each measure are listed in Table 3.

For neural activity (i.e., the Learning > Application contrast), we observed distinct developmental changes across the different ROIs. For DLPFC, the relationship between age and neural activity was best described by both a linear and a quadratic term for age, i.e. activity increased with age, and leveled off in late adolescence and young adulthood (see Figure 4). For SPC, there was a quadratic but not a linear effect of age, which indicated that activity increased until adolescence and then decreased into adulthood. In the pre-SMA/ACC, neural activity was best described by a linear effect of age that showed increasing neural activity with increasing age. In all ROIs, there was significant individual variability in mean neural activation as indicated by a random effect of the intercept. See Table 3 for the model comparison values and Figure 4 for the predicted and actual data.

Table 2: Correlations between all variables at both time points.

Time point 1	DLPFC	SPC	SMA	DLPFC CT	SPC CT	SMA CT	WM	Perf	Age TP1
DLPFC	1								
SPC	.463**	1							
SMA	.489**	.574**	1						
DLPFC CT	-.192**	n.s.	n.s.	1					
SPC CT	n.s.	n.s.	n.s.	.390**	1				
SMA CT	n.s.	n.s.	n.s.	.423**	.332**	1			
WM	.203**	n.s.	n.s.	n.s.	n.s.	-.144*	1		
Perf	.343**	.222**	.155*	-.140*	n.s.	n.s.	.437**	1	
Age TP1	.410**	.205**	.210**	-.382**	-.162*	-.307**	.340**	.444**	1
Time point 2	**DLPFC**	**SPC**	**SMA**	**DLPFC CT**	**SPC CT**	**SMA CT**	**WM**	**Perf**	**Age TP2**
DLPFC	1								
SPC	.326**	1							
SMA	.400**	.389**	1						
DLPFC CT	-.146*	n.s.	-.170*	1					
SPC CT	n.s.	n.s.	n.s.	.383**	1				
SMA CT	n.s.	n.s.	n.s.	.491**	.341**	1			
WM	.165*	n.s.	n.s.	n.s.	n.s.	n.s.	1		
Perf	.227**	n.s.	.159*	n.s.	n.s.	n.s.	.285**	1	
Age TP2	.209**	n.s.	.296**	-.380**	n.s.	-.307**	.311**	.167*	1

***: Significant at p < .01, *: Significant at p < .05. Abbreviations: CT = cortical thickness, WM = working memory, Perf = performance, n.s. = not significant.*

Table 3: AIC and loglikelihood p-values for the base model, and linear, quadratic and cubic age models for brain activity, cortical thickness, task performance and working memory capacity.

	Base	**Linear**		**Quadratic**		**Cubic**	
	AIC	**AIC**	***p***	**AIC**	***p***	**AIC**	***p***
DLPFC activity	1562.10	1537.33	<.001	**1533.26**	.014	1535.09	.675
Pre-SMA/ACC activity	1872.18	**1851.03**	<.001	1851.70	.249	1853.45	.618
SPC activity	1688.66	1690.18	.486	**1683.57**	.003	1685.30	.600
DLPFC CT	-425.54	**-553.83**	<.001	-553.17	.248	-553.715	.110
Pre-SMA/ACC CT	-336.29	-421.659	<.001	-421.60	.164	**-430.81**	<.001
SPC CT	-447.86	-478.99	<.001	**-488.05**	<.001	-486.25	.656
Performance	2442.43	2406.12	<.001	**2387.45**	<.001	2386.47	.084
Working Memory	-485.96	-529.56	<.001	**-546.92**	<.001	-548.52	.057

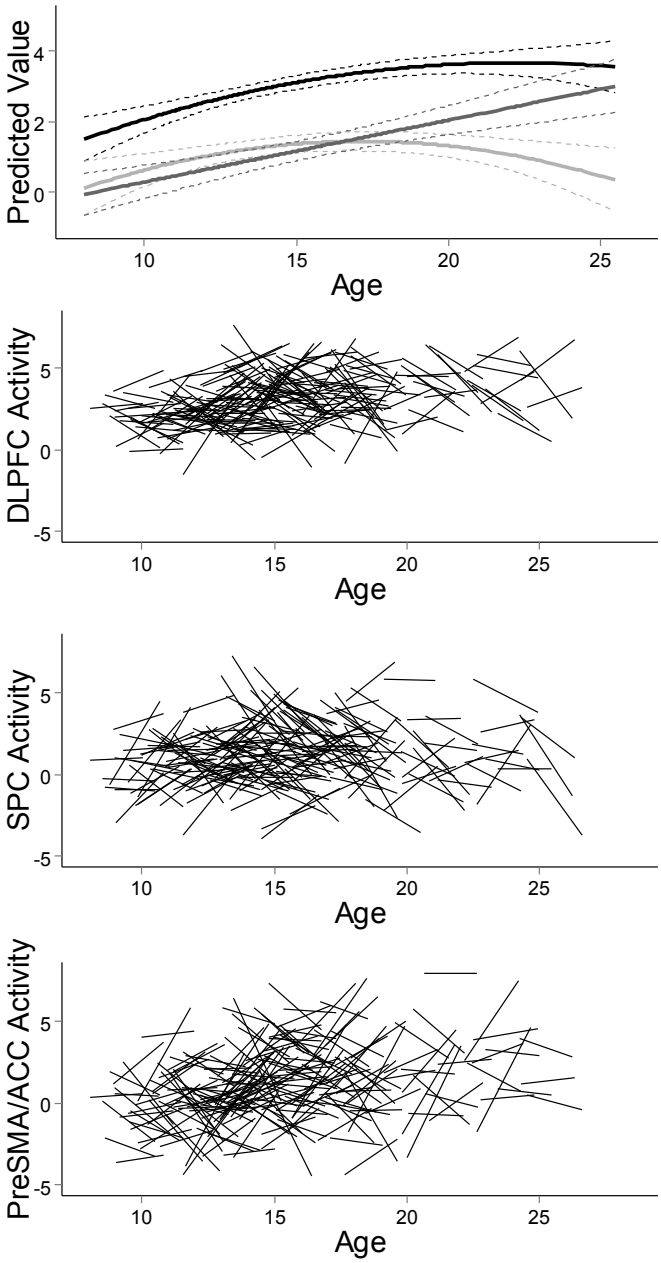


Figure 4a: Predicted data (presented in the upper graphs) and raw data for neural activity in DLPFC, SPC and pre-SMA/ACC. Each line represents one individual at two time points.

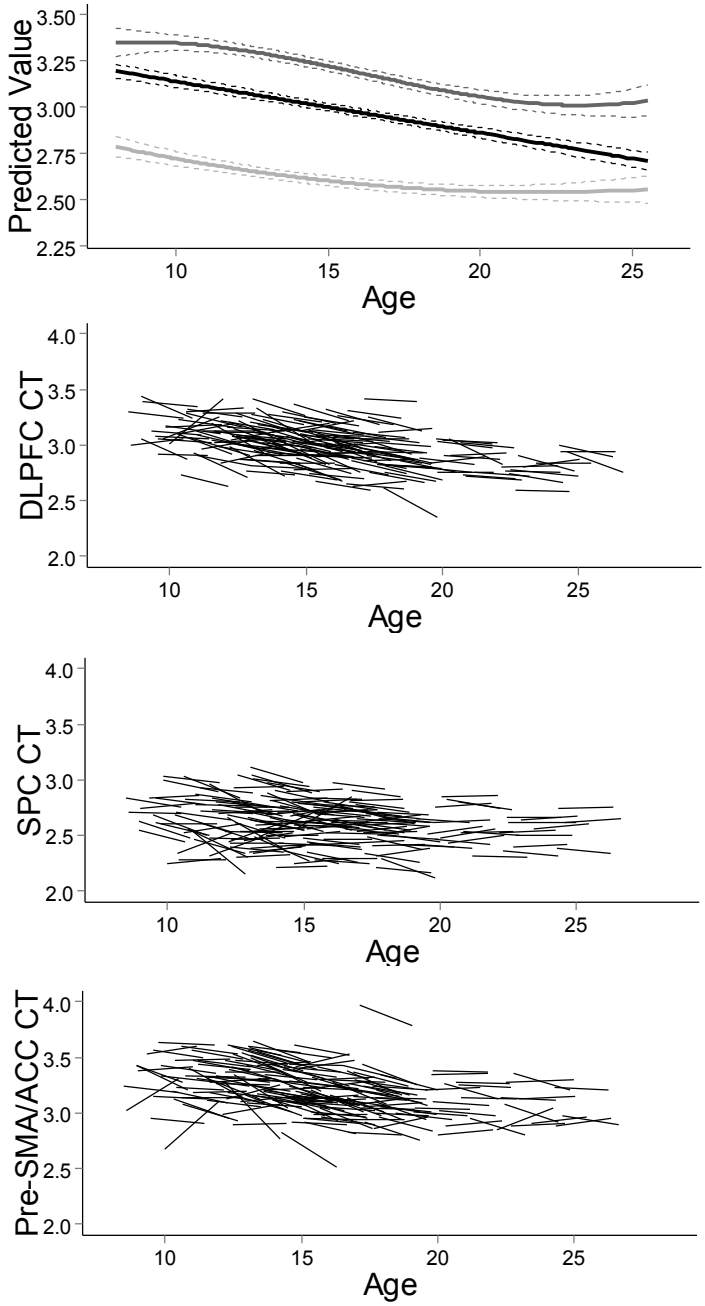


Figure 4b: Predicted data (presented in the upper graphs) and raw data for cortical thickness (CT) in DLPFC, SPC and pre-SMA/ACC. Each line represents one individual at two time points.

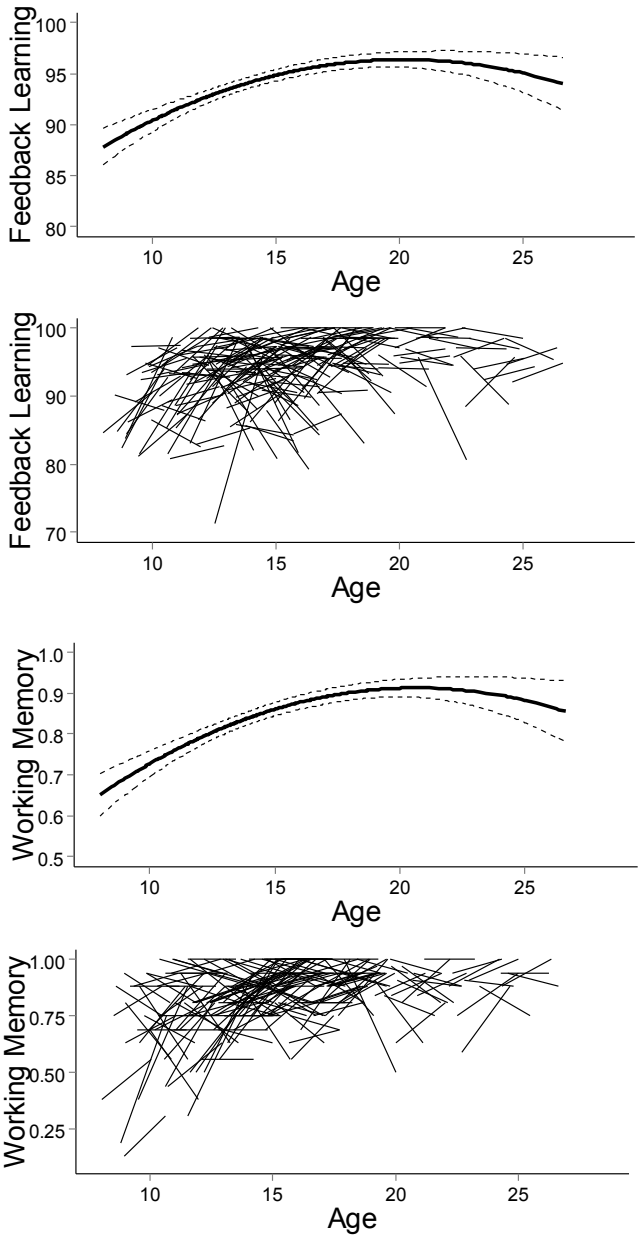


Figure 5: Predicted data and raw data for feedback learning performance and working memory capacity. Feedback learning performance was defined as the percentage of trials during the learning phase where (positive or negative) feedback was successfully used on a subsequent trial.

For task performance, there was a linear and quadratic effect of age on task performance. Note that although the AIC value was slightly lower for a cubic model than for a quadratic model, this difference was not significant according to a log-likelihood test (p = .084). This indicated that task performance improved with age, and then leveled off for older participants (see Table 3 and Figure 5 for the raw data and predicted data).

The age models for working memory indicated that a combined linear and quadratic model described the best fit. Note that the AIC value was slightly lower for the cubic age model, but the accompanying log-likelihood test showed no significant improvement (p = 0.58) (see Table 3; Figure 5). This indicated that working memory performance improved with age, and then leveled off for older participants.

For cortical thickness of the ROIs, we found different developmental patterns depending on the region. For DLPFC, a linear model with decreasing cortical thickness with age best described the data. In contrast, in SPC the relationship between age and cortical thickness was best described by a quadratic model with a significant linear and quadratic term for age. That is, cortical thickness decreased with age and stabilized in late adolescence/young adulthood (see Figure 4). For pre-SMA/ACC, a model with both a linear and cubic (but not quadratic) effect of age best described the data, i.e., cortical thickness showed a relatively stable pattern in young adolescents, decreased steeply in adolescence and stabilized in late adolescence/early adulthood (see Figure 4). All results are described in Table 3 and visualized in Figure 4.

Explaining developmental change with age, performance, working memory and cortical thickness

To investigate which factors additionally drive these developmental changes in neural activity, we tested the contributions of task performance, working memory capacity, and cortical thickness in addition to age in a hierarchical mixed regression model. Neural activity (i.e., the Learning > Application contrast) was the dependent variable (for each ROI separately) and the predictors task performance, working memory and cortical thickness were added in a consecutive order above age. The starting model was the model with the best fitting age shape (linear, quadratic or cubic). Next, task performance, working memory and cortical thickness were added in hierarchical steps (see Methods section). For DLPFC, the model that best explained neural activity was a model including a linear (but no longer quadratic) age effect and a positive effect of task performance, i.e. better performance predicted increased activity. Cortical thickness and working memory did not explain additional variance over and above age and task performance. For SPC, both age and task performance explained a significant amount of variance in neural activity, such that in addition to a quadratic age effect, better task performance was associated with increased activity. Similar to the DLPFC, working memory capacity and cortical thickness did not contribute to the model over and above age and task performance. A different pattern was found for pre-SMA/ACC. Here, we observed a significant positive linear effect for age and cortical thickness,

but not for task performance and working memory. This model indicated a positive relation between cortical thickness and neural activity, such that increased activity was associated with increased cortical thickness. The final model parameters are summarized in Table 4.

Table 4: Model parameters for the best fitting model for DLPFC, SPC and pre-SMA/ACC.

				95% CI	
Area	**Variance**	***β***	***p***	**Lower**	**Upper**
DLPFC					
Random effect					
Intercept	0.69			0.49	0.98
Fixed effects					
Intercept		3.09	<.001	2.92	3.25
Age		5.89	<.001	2.57	9.22
Age^2		-2.17	.18	-5.35	1.00
Performance		0.07	<.001	0.035	0.11
SPC					
Random effect					
Intercept	1.09			0.087	1.36
Fixed effects					
Intercept		1.18	<.001	0.96	1.39
Age		-0.46	.833	-4.74	3.92
Age^2		-4.36	.036	-8.4	-0.31
Performance		0.05	.044	0.0015	0.09
Pre-SMA/ACC					
Random effect					
Intercept	1.51			1.26	1.82
Fixed effects					
Intercept		1.29	<.001	1.01	1.56
Age		0.20	<.001	0.12	0.28
Cortical Thickness		1.48	.042	0.062	2.85

Discussion

The main aim of this study was to examine the developmental trajectory of the frontoparietal network during a feedback-learning task that captures multiple aspects of cognitive control. We tested for different developmental trajectories (linear, quadratic and cubic) in neural activation and for factors contributing to time-related changes in brain activity for cognitive control above age, particularly task performance, working memory capacity and cortical thickness as an index of structural brain development. The results showed that 1) neural activity during a feedback learning task was best characterized by a quadratic age function in DLPFC and SPC, increasing in early adolescence and leveling off/decreasing into adulthood, but a linear pattern for pre-SMA/ACC, i.e. increasing monotonically across adolescence, 2) Task performance explained additional variance above age in DLPFC and SPC, but not pre-SMA/ACC activity, with higher performance predicting increased neural activation, 3) Cortical thickness explained additional variance in pre-SMA/ACC but not in DLPFC and SPC activity, with increased cortical thickness associated with increased neural activation during feedback learning, and 4) There was fair to good reliability for activity in DLPFC, pre-SMA/ACC and SPC across a two-year period for participants aged 8-27 years. These findings are discussed in more detail in the next paragraphs.

Growth trajectories for neural activity, cortical thickness, performance and working memory

Mixed model analyses of the longitudinal data showed that neural activity was described by different patterns in the neural regions of interest. The linearly increasing pattern in pre-SMA/ACC fits with prevailing developmental theories of monotonously increasing cognitive control. However, the quadratic pattern in DLPFC and SPC is more novel and contradicts initial theories of prefrontal cortex maturation, which suggested a linearly protracted developmental pattern until the early twenties. A closer inspection of the studies reported to date shows that no prior large-scale study included our age range of 8-27 years. Instead, prior studies selected age groups, such as adolescents aged 13-17 versus adults aged 25-30 years or other age selections that did not span the whole range of adolescence and early adulthood (Geier, Garver, Terwilliger, & Luna, 2009; Thomason et al., 2009; van den Bos et al., 2009; van Duijvenvoorde et al., 2008; Velanova, Wheeler, & Luna, 2008). Thus, these prior studies may have been underpowered to detect a quadratic pattern, especially because most prior studies used cross-sectional comparisons (Fjell et al., 2010).

Previous studies found contradicting patterns for frontoparietal recruitment, with some showing increases in activation with increasing age and others showing decreases, and both patterns were interpreted as reflecting immaturity in adolescence (Pfeifer & Allen, 2012). Intriguingly, the current findings show peak activity in SPC and to a lesser extent in DLPFC during late adolescence when learning from feedback. For SPC, there was quadratic pattern with a peak in mid-adolescence, whereas for DLPFC, there was a combined linear and quadratic pattern with

peaking and leveling off in late adolescence/early adulthood. Recent research also showed that complex paradigms such as divergent thinking (Kleibeuker, De Dreu, & Crone, 2013) resulted in stronger DLPFC activity in adolescents than in adults. Studies using more basic cognitive control tasks, however, have not reported these peak activations in frontoparietal regions in adolescence (Klingberg et al., 2002; Rubia et al., 2006). Therefore an important direction for future research is to unravel whether, when, and how DLPFC and SPC show peak activity in late adolescence. Recently, it was found that adolescents recruit DLPFC more strongly than adults when financial incentives were offered for performing well (Teslovich et al., 2014), suggesting that adolescents may engage DLPFC more in a context of high motivation. Possibly, the finding of enhanced activity in SPC and DLPFC in late adolescence indicates that this is a time window when the frontoparietal network can be optimally recruited (Crone & Dahl, 2012). This hypothesis should be tested in more detail in future studies using a larger variety of paradigms, and by also including broader age ranges including a larger adult sample than our sample. Nevertheless, we suggest these results seem consistent with a new perspective on nonlinear development of the frontoparietal cortex (see also Casey, 2015).

We also examined the effects of individual differences on time-related changes in neural activity besides age, such as cortical thickness, working memory capacity and performance on the feedback learning task. We first tested the general age patterns for these individual differences, and subsequently related these to changes over time in activation patterns. First, for cortical thickness, all areas showed a decrease with age, which fits with prior studies showing an initial increase in childhood followed by cortical thinning in adolescence (Koolschijn & Crone, 2013; Shaw et al., 2013; Tamnes et al., 2010). In this study the decrease in cortical thickness in DLPFC was best described by a linear pattern. For SPC, a combined linear and quadratic decrease best described the data, i.e. a steep decrease in early adolescence which leveled off towards late adolescence. Cortical thickness decreases in pre-SMA/ACC were best described by a combined linear and cubic function, showing relatively stable levels in early adolescence, decreasing steeply in adolescence, and leveling off in young adulthood. The overall pattern of a developmental decrease is consistent with earlier reports.

Second, both performance on the feedback learning task and working memory capacity showed a quadratic pattern for age, with steadily increasing performance in adolescence which leveled off in adulthood. A close examination of the literature shows that few studies examined cognitive control development across the whole age range of adolescence until adulthood. For example, a large and comprehensive study on the development of executive functions compared children of 7, 11, 15 and 21 years (Huizinga et al., 2006), but not the intermediate ages. Likewise, a study on performance monitoring previously compared performance of 8-9 year-olds, 11-13-year-olds and 18-25-year-old adults, but did not separate between ages within the adult group (van Duijvenvoorde et al., 2008). The strength of the current study is that quadratic changes in task performance and working memory capacity were observed using a longitudinal design.

After testing these general age patterns, we tested whether task performance, working memory capacity and cortical thickness explained additional variance in time-related changes of frontoparietal activity above age. We found that performance explained additional variance above age in DLPFC and SPC but not pre-SMA/ACC activity, with better performance linked to increased activity. This suggests that task performance provides a unique contribution to activity changes that is not captured by age alone. Cortical thickness explained additional variance above age in activity in pre-SMA/ACC but not in DLPFC or SPC. The latter finding is consistent with earlier studies that showed that DLPFC and SPC structure could not explain age differences in neural activity (Haier, Karama, Leyba, & Jung, 2009; Squeglia et al., 2013). The relation between pre-SMA/ACC cortical thickness and neural activity was positive, such that increased cortical thickness (i.e., less mature) is associated with increased activity. This is surprising given that cortical thickness decreased with age. Intuitively, one might expect that pruning of connections during development increases selectivity and effectiveness of synaptic activity and, therefore, leads to stronger activation. Evidence for this line of reasoning, that is, stronger activations are generally found in thinner (i.e. more mature) regions, has been provided by such relationships in frontoparietal regions in children performing linguistic tasks (Lu et al., 2009; Nunez et al., 2011). On the other hand, increases of activation have for instance been reported after training studies (Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005), in areas that were thicker in musicians than non-musicians (Bermudez, Lerch, Evans, & Zatorre, 2009). Interestingly, our findings are consistent with a study in adult participants by Hegarty et al. (2012) who also found a positive relation between cortical thickness and activity in pre-SMA/ACC, but not in other prefrontal areas.

A limitation of this study is that the participants who were excluded for the cortical thickness analyses due to lower quality data (n = 28), were younger on average compared to the included participants, which is unfortunately a common problem in developmental fMRI studies. However, the remaining sample for cortical thickness was still large (n = 180; ages 8-27) and this was one of the first large-scale developmental longitudinal studies assessing the contribution of structural maturation to development of brain function. Taken together, our findings build upon prior research showing relationships between brain structure and function, but the underlying mechanisms and models for these relationships are still open for discussion.

Reliability of neural activity across child and adolescent development

Finally, in this study we measured consistency over time of activity in the frontoparietal network within individuals. There is a growing number of longitudinal studies in adults that examined test-retest reliability in the cognitive control network. The results indicate that over periods of two weeks, reliability is good in DLPFC and SPC during an n-back working memory task (Plichta et al., 2012), there is modest to good reliability of DLPFC activity in visual working memory over a period of three months (Zanto, Pa, & Gazzaley, 2014) and there is modest to good reliability of DLPFC and SPC in a feedback monitoring task over a period of three years (Koolschijn et al.,

2011). The question of reliability in developmental populations has not been consistently examined, but a recent study suggested fair reliability in DLPFC, SPC and ACC in 123 participants between ages 9 to 29 in an oculomotor inhibition task (Ordaz et al., 2013). The current findings demonstrate that in a large sample of 208 participants, reliability across two years is fair to good across ages 8-27 years in bilateral DLPFC, SPC and pre-SMA/ACC. The relatively high reliability of cortical regions is consistent with prediction studies that have shown that future academic achievement can be predicted by activity in SPC one year earlier (Emerson & Cantlon, 2014) and two years earlier (Dumontheil & Klingberg, 2012) during working memory tasks, suggesting that activity in DLPFC and SPC is related to future cognitive outcome.

Conclusions and future directions

This study moved beyond prior cross-sectional comparisons by fitting growth curves based on longitudinal data and thereby improved power for detecting developmental change. An interesting direction for future research is to test how neural activity predicts future changes in behavior. A prior study highlighted the important role of the parietal cortex for predicting future behavior such as academic performance (Dumontheil & Klingberg, 2012). However, other studies showed that activity in the frontoparietal network correlated with working memory capacity across sessions, but activity in subcortical areas (basal ganglia and thalamus) predicted future working memory capacity (Darki & Klingberg, 2014; Ullman, Almeida, & Klingberg, 2014).

In future studies it will be of interest to follow individuals in this study for a third time point. With two time points, it was only possibly to investigate nonlinear patterns on the group level. In addition, an interesting future direction would be to examine whether and how baseline activity in the frontoparietal network predicts future behavioral outcomes. An important question for future research is to examine not only the spatial, but also the temporal dynamics of cognitive control, such as with event-related potentials. Prior studies have highlighted the feasibility of this approach in young children (Eppinger et al., 2009), and this will allow for the investigation of fast evaluative processes. A specific limitation of this study is that the feedback learning task was relatively easy and performance was near-perfect for older participants. Future studies should investigate whether the results from this feedback-learning paradigm can be replicated using other and more cognitively taxing tasks measuring cognitive control. Finally, we only investigated whether time-related changes in neural activity covary with the factors age, task performance, working memory capacity and cortical thickness. Many other factors may contribute to changes over time in activity, such as increased response inhibition or motivation, which should be investigated in future research.

Taken together, this study accentuates the important role of the emerging frontoparietal network in child and adolescent development. With regard to dual-process models of adolescent development, this study provides evidence against a simple linear development of the frontopa-

rietal network and highlights the need for further large-scale longitudinal studies to test adolescent development more reliably.

Chapter 7

Predicting reading and mathematics from neural activity for feedback learning

This chapter is based on:

Peters, S., Van der Meulen, M., Zanolie, C.K.K. & Crone, E.A. Predicting reading and mathematics from neural activity for feedback learning: A longitudinal study (in revision, 2015).

Abstract

Although many studies use feedback learning paradigms to study the process of learning in laboratory settings, little is known about their relevance for real-world learning settings such as school. In a large developmental sample (*N* = 228, 8-27 years), we investigated whether performance and neural activity during a feedback learning task predicted reading and mathematics performance two years later. The results indicated that feedback learning performance predicted both reading and mathematics performance. Activity during feedback learning in left superior dorsolateral prefrontal cortex (DLPFC) and left superior parietal cortex (SPC) predicted reading performance, whereas activity in pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC) predicted mathematical performance. Moreover, left superior DLPFC and pre-SMA/ACC activity predicted unique variance in reading and mathematics ability over behavioral testing of feedback learning performance alone. These results provide valuable insights into the relationship between laboratory-based learning tasks and learning in school settings, and the value of neural assessments for prediction of school performance over behavioral testing alone.

Introduction

Learning from performance feedback is an important skill allowing us to rapidly adjust behavior based on changes in environmental demands (Holroyd & Coles, 2002). Thus, it is an adaptive form of learning which allows individuals to flexibly and creatively adapt to a changing environment in a successful way. Feedback learning is often investigated in controlled laboratory settings to study the process of learning. However, it is unclear how feedback learning in these controlled experimental paradigms relates to real-world learning in settings such as school. In this study, we investigated this question in a large developmental sample of participants between 8-27 years, focusing on both neural and behavioral indices of feedback learning as predictors for school performance two years later.

School performance can be measured in different ways. The most important school performance skills taught in schools across the world are reading and mathematics, given that many courses in school rely on children's ability to read proficiently and perform mathematical calculations. Many children who are poor readers in school keep having difficulties with reading later in life (O'Shaughnessy, Lane, Gresham, & Beebe-Frankenberger, 2003) and research has demonstrated that performance on mathematical tests predicts employability, productivity and salaries in adulthood (Geary, 2000; Rivera-Batiz, 1992).

Although the link between laboratory-based feedback learning tasks and school performance (e.g., mathematics and reading performance) is not yet clear, several studies have provided evidence that both feedback learning and reading and mathematics are linked to executive functions. Executive functions are defined as the ability to behave in goal-directed actions in new situations and to overcome automatic thoughts and behaviors (Garon et al., 2008). Executive functions are thought to consist of three subprocesses, or basic executive functions: (1) working memory, (2) inhibition and (3) switching (Huizinga et al., 2006; Miyake et al., 2000). Prior research using structural equation modeling showed that complex executive function tasks, such as performance on the classic Wisconsin Card Sorting Task (WCST), requires several basic executive functions, such as working memory and task switching (Huizinga et al., 2006; Miyake et al., 2000). It has been argued that complex cognitive tasks which rely on multiple subprocesses of executive functions are the most reliable correlates of cognitive performance in daily life (Barcelo & Knight, 2002), possibly because these tasks are more similar to everyday challenges. Feedback learning can also be interpreted as a complex executive control process, which most likely relies on multiple subprocesses of executive functions (Peters & Crone, 2014), and may rely partly on working memory capacity, given that feedback learning shares commonalities with the classic WCST (Huizinga et al., 2006).

Evidence for the relationship between school performance and executive functioning comes from numerous studies that demonstrated a link between working memory, inhibition and switching on the one hand, and reading and mathematics performance on the other (Blair &

Razza, 2007; Bull & Scerif, 2001; Raghubar, Barnes, & Hecht, 2010; Van der Sluis, De Jong, & Van der Leij, 2004). The link between executive functioning and school performance is not surprising, given that to develop reading and mathematics understanding, children probably need additional cognitive skills. For example, children have to be able to understand grammatical and numerical structure, keep track of the sentences read or mathematical steps taken before, and integrate information from long-term memory with current information to form a coherent view (Cain, Oakhill, & Bryant, 2004; Landi, Frost, Mencl, Sandak, & Pugh, 2013), which are all processes intimately related to executive functioning. This led us to hypothesize that feedback learning in controlled laboratory settings is a valid predictor of real-world learning performance in schools.

A second reason why feedback learning and reading and mathematical ability are expected to be related, is because they rely on similar brain mechanisms. The main neural areas involved during feedback processing are the dorsolateral prefrontal cortex (DLPFC), superior parietal cortex (SPC) and pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC) (Peters, Braams, et al., 2014; Zanolie et al., 2008). Meta-analyses of fMRI-activity during reading also show recruitment of the DLPFC (Ferstl, Neumann, Bogler, & Von Cramon, 2008) and pre-SMA/ACC (Ferstl et al., 2008; Houdé, Rossi, Lubin, & Joliot, 2010) amongst other areas (mostly lateralized to the left hemisphere). Meta-analyses on mathematics-related neural activity also showed involvement of the DLPFC (Arsalidou & Taylor, 2011; Houdé et al., 2010), parietal cortex and pre-SMA/ACC (Arsalidou & Taylor, 2011). Thus, it is to be expected that activity patterns in DLPFC and pre-SMA/ACC are linked to reading and mathematics.

Recently, an increasing body of research has directed attention to predicting school performance from brain measures. A possible advantage of collecting neural measures in addition to behavioral measures is the hypothesis that brain measures can provide unique predictive information over behavioral measures alone (Dumontheil & Klingberg, 2012; Hoeft et al., 2007). In the current study, we investigated the link between learning in a controlled laboratory setting, and reading and mathematical ability as indices for real-world learning. We focused on fluency at reading single words, because this is one of the most crucial aspects of reading determining reading ability at a later stage (Jenkins, Fuchs, Van Den Broek, Espin, & Deno, 2003; Juel, 1988). To assess mathematics proficiency, we used a standardized arithmetic test that is part of the Wechsler Adult Intelligence Scale and the Wechsler Intelligence Scale for Children, which measures numerical reasoning and mathematical problem solving. In addition, we investigated whether individual differences in working memory capacity could explain a possible link between feedback learning and reading and mathematics performance. For instance, Huizinga et al. (2006) found that from the factors working memory, inhibition and switching, only working memory predicted WCST performance, a task that also relies on learning from feedback. We hypothesized that feedback learning performance would predict reading and mathematics performance two years later, and that neural measures would provide additional information over behavioral testing (feedback learning and working memory performance) alone.

Methods

Participants

The initial sample consisted of 299 participants (see also Peters, Braams, et al., 2014; Peters, Koolschijn, Crone, Van Duijvenvoorde, & Raijmakers, 2014), for whom data was collected on two time points (T1 and T2) which were approximately 2 years apart (M = 1.99, SD = 0.10, range: 1.66-2.47 years). The included sample with complete data at T1 for feedback learning and fMRI data consisted of 268 participants. At T1 participants were excluded from analyses for a variety of reasons, such as reported history of neurological or psychiatric disorders or use of psychotropic medication, movement in the MRI scanner exceeding 3.0 mm (N = 19), technical issues (N = 3) or because they were outliers at the lower end (more than three times the interquartile range) on feedback learning performance (N = 3).

At T2, there was complete data on reading and math performance for 228 participants (119 females) who were also included at T1 (aged 8.01 – 24.55 years at T1 (M = 14.35, SD = 3.57) and aged 9.92 – 26.62 at T2 (M = 16.34, SD = 3.58). All analyses were performed on these 228 participants. IQ scores at T1 were estimated using two subtests (Similarities and Block Design) of the WISC-III (participants 8-15 years old) or WAIS-III (participants 16-25 years old). Estimated IQ scores ranged from 85 to 143 (M = 110.78, SD = 9.80). The study was approved by the Institutional Review Board at the University Medical Center and all participants older than 12 (and participants' parents for children under 18) signed an informed consent form. Adults received payment (€60) for participation and children and their parents received brain-related presents and a payment for travel reimbursement (€30 for children 12-17 years, €25 for children 8-11 years).

Materials

Reading Fluency

Technical reading skills were measured with a reading fluency task at T2. We used one of the tests in the Dutch "Three-Minute-Test" (Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). In this task, participants received a list of words and were instructed to read aloud as many words as possible in one minute. The total score is defined as the number of correct words minus the number of incorrect words. The Three-Minute-Test has good validity and reliability (Cronbach's alpha, dependent on age group > 0.92) (Krom et al., 2010).

Mathematics

Mathematical ability was measured with the subscale "Arithmetic" of the Wechsler Intelligence Scales (WISC-III for participants under 16, WAIS-III for participants of 16 years and older). A set of arithmetical problems of increasing difficulty was administered verbally. All arithmetic problems had a time limit of 30 to 75 seconds, depending on the difficulty of the problem. If the partic-

ipants failed to correctly answer three consecutive problems the test was aborted. Both the WISC and the WAIS resulted in raw scores that were converted to norm scores relative to same-aged peers. We used norm scores in further analyses (see also Barnea-Goraly, Eliez, Menon, Bammer, & Reiss, 2005; Li, Hu, Wang, Weng, & Chen, 2013) to ensure comparability between the different ages (reflected in WISC and WAIS scores). In addition, we performed our main analyses with the mathematics subtest with raw scores for the WISC and WAIS group separately.

Working memory

We measured working memory performance at T1 to assess whether feedback learning and reading and mathematics performance were explained by individual differences in working memory. Working memory capacity was measured with the Mental Counters task (Huizinga et al., 2006), in which participants need to keep numerical information active. For this task, two independent counters were presented on a computer screen. The counters were horizontal bars for which the values changed depending on the position of a square. If a square was presented above a counter the participant was instructed to add 1 to the current value, if a square was presented below the counter the participant was instructed to subtract 1 from the current value of the counter. The squares appeared randomly above or below one of the two counters. Participant were instructed to keep track of both counters and to press a button as soon as one of the counters reached a given criterion value (e.g., when one of the counters reached the value 3). The squares were randomly presented in series (the number of trials before criterion was reached) of 5 or 7 trials with inter-trial intervals of 1000 to 1300 ms, with a total of 16 trials. The proportion of correct trials was used as a measure of performance.

Feedback Learning Task

Participants performed a feedback learning task in the MRI scanner (Peters, Braams, et al., 2014; Peters, Koolschijn, et al., 2014). On every trial, three empty boxes were presented in the top half of the screen in the stimulus and feedback display. During presentation of the stimulus display one of three different stimuli was presented in the centre of the bottom half of the screen (see Figure 1). Participants were instructed that each stimulus belonged in one of three boxes for an entire sequence and they had to find the correct location for all three stimuli by using performance feedback. Each trial started with a 500 ms fixation cross, presented in the center of the screen. After fixation the stimulus display was presented for 2500 ms, during which participants were required to sort the stimulus in one of three squares. Participants responded by pressing one of three buttons strapped to their right leg. If participants failed to respond within 2500 ms "Too Late" was presented in the centre of the screen, after which the sequence continued. After the response, performance feedback was presented for 1000 ms. When a participant sorted a stimulus in the correct square a plus-sign (positive feedback) was shown, when a participant sorted a stimulus in the incorrect square a minus-sign (negative feedback) was shown. Inter-trial interval

(blank screen) was jittered to optimize the timing for fMRI based on OptSeq (Dale, 1999) with intervals between 0 and 6 seconds. A sequence was aborted when the participant sorted each stimulus twice in the correct location, or after 12 trials in total. When a sequence ended a new sequence with three new unique stimuli was presented. There were 15 sequences in total, resulting in a maximum of 180 trials. Stimuli were presented in a pseudorandom order, with a maximum of two identical stimuli in a row. Before the MRI session, all participants practiced three sequences. During the MRI session the task was divided into two runs of eight and seven sequences, respectively.

To calculate a performance measure for feedback learning we calculated the percentage of trials in the learning phase where feedback was successfully used on the next trial. For this purpose we divided the number of trials during the learning phase which were succesfully applied in the next trial, by the total number of trials during the learning phase.

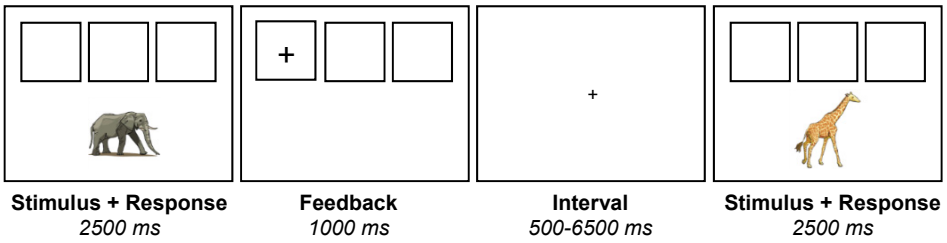


Figure 1: Display of task sequence for the feedback learning task. A trial started with a 2500 ms stimulus display during which the participant responded by sorting the stimulus in one of the three boxes. In this example, the participant (correctly) chose the left box. Next, feedback was presented for 1000 ms by either a '+' for correct feedback or a '–' for incorrect feedback. After an inter-trial interval (varying from 0-6 s) and a 500 ms fixation cross, the next stimulus was presented.

FMRI data acquisition

MRI scans were obtained with a Philips 3.0 Tesla MRI scanner. Functional scans for the feedback learning tasks were acquired during two runs with T2*-weighted echo-planar imaging (EPI). The first two volumes were discarded to allow for equilibration of T1 saturation effects. The following settings were used: TR = 2.20 s, TE = 30 ms, sequential acquisition, 38 slices, slice thickness = 2.75 mm, Field of View (FOV) = 220 x 220 x 114.68 mm. For the structural scan, a high-resolution 3D T1-FFE was obtained after the experimental tasks (TR = 9.76 ms, TE = 4.59 ms, 140 slices, voxel size = 0.875 mm, FOV = 224 × 177 × 168 mm). The experimental task was projected on a screen,

which was visible to participants through a mirror. Participants were accustomed to the MRI environment and sounds with a mock scanner before the actual MRI scan.

FMRI data Analysis

We used SPM8 (Wellcome Department of Cognitive Neurology, London) to analyze fMRI. The following pre-processing steps were used: correction for slice timing acquisition and rigid body motion, spatial normalization to T1 templates (MNI305 stereotaxic space (Cocosco et al., 1997)) using a 12-parameter affine transform together with a nonlinear transformation involving cosine basis functions and resampling of the volumes to 3 mm voxels. Functional scans were smoothed with an 8mm FWHM isotropic Gaussian kernel. For further fMRI analyses, we used a contrast that reveals brain areas with sensitivity to informative feedback for learning (Eliassen et al., 2012; van den Bos et al., 2009), that is, areas responding more to feedback providing new information (i.e., more informative) compared to feedback providing known information. To compare neural activity for 'informative' and 'uninformative' feedback, we distinguished between a learning phase and an application phase for each stimulus. For the learning phase, we included trials where participants had not correctly sorted this particular stimulus yet, and were thus still using feedback to determine the correct location. Only trials for which feedback was used appropriately on the next trial for that stimulus were included. Thus, feedback was categorized as learning, when positive feedback resulted in choosing the same location on a next trial and when negative feedback resulted in sorting in a different location. These trials during the learning phase were compared to the application phase: trials in which a stimulus was sorted correctly on a preceding trial, and continued to be sorted correctly. All further analyses were based on a comparison between the learning phase and the application phase, i.e. the contrast Learning > Application. In order to calculate this contrast for all participants, we first modeled the fMRI time series with events corresponding to the events "Positive Learning", "Negative Learning", and "Application", time-locked with 0-duration to the moment of feedback, which were convolved with a canonical hemodynamic response function. Other trials (e.g., trials during the learning phase that did not result in learning or trials where participants responded too late) were modeled as events of no interest. The events were used in a general linear model; along with a set of cosine functions which high-pass filtered the data. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used for the calculation of the contrast Learning (Positive Learning + Negative Learning) > Application for each subject. The resulting contrast images were submitted to higher-level analyses.

FMRI Region-of-interest analysis

In order to examine neural effects of feedback learning and its relation to reading and mathematics performance, region-of-interest (ROI) analyses were performed with the Marsbar toolbox in SPM8 (Brett et al., 2002). The contrast used to generate functional ROIs was Learning > Applica-

tion (FWE corrected, $p < .05$, > 10 contiguous voxels). The resulting ROIs spanned several brain regions. Therefore, the ROIs were subdivided by masking the functional ROI with the following anatomical Marsbar ROIs (based on Automated Anatomical Labeling (AAL)): left and right DLPFC (Middle Frontal Gyrus in AAL), pre-SMA/ACC (Supplementary Motor Area in AAL; left and right combined), left and right SPC (Superior Parietal Lobule in AAL). These ROIs were selected based on earlier studies demonstrating that these areas show developmental changes for feedback learning (Crone et al., 2008; Peters, Braams, et al., 2014; van Duijvenvoorde et al., 2008) and were also used in a prior study with the same experimental task (Peters, Braams, et al., 2014).

The DLPFC ROIs, even after masking, were still very large (right: 28488 mm; left: 28240 mm), therefore, we created 6 mm radius spheres based on four local maxima within the DLPFC regions (two per hemisphere). These areas are referred to as 'superior DLPFC (sup-DLPFC)' and 'mid-DLPFC'. Centre-of-mass MNI (x y z) coordinates for the ROIs were: pre-SMA/ACC: 0 9 58; right sup-DLPFC: 21 9 57; left sup-DLPFC: -24 3 57, right mid-DLPFC: 42 18 39; left mid-DLPFC: -42 24 39; right SPC: 27 -62 55; left SPC: -23 -64 50 (See Figure 2).

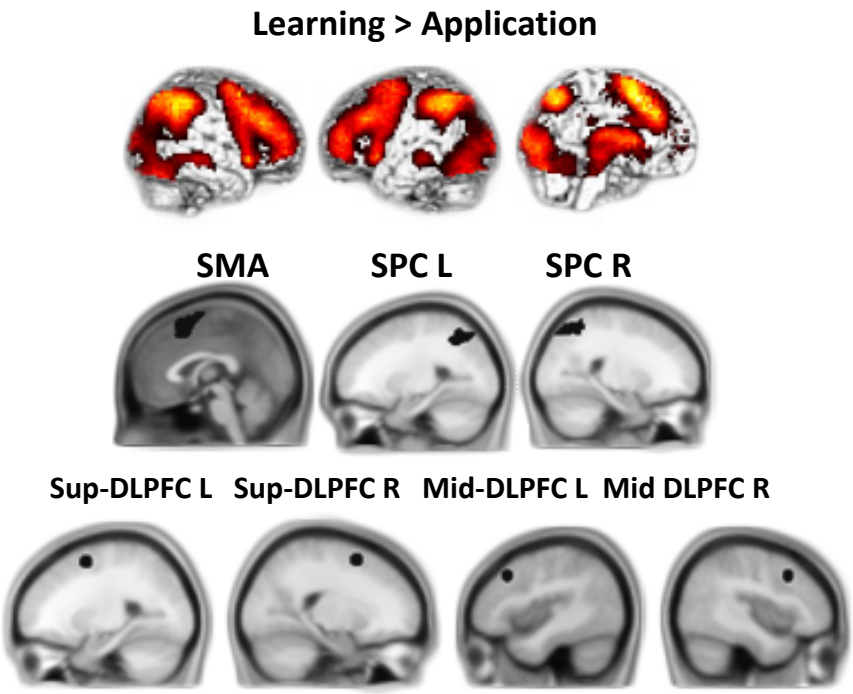


Figure 2: Wholebrain results for the contrast Learning > Application (FWE-corrected at $p < .05$, > 10 contiguous voxels) and the regions-of-interest based on this contrast.

Results

Data checks

We performed several data quality checks by investigating relationships between the main variables of interest (neural activity and behavioral performance for feedback learning, and reading and mathematics) and age, IQ, working memory and sex (See Table 1 for an overview of the values for age, IQ, working memory, feedback learning, reading and mathematics).

Table 1: Descriptive values for age, IQ, working memory, feedback learning, reading and mathematics scores for male and female participants separately. In the right-most column, we indicated the p-value for sex differences.

	Female				**Male**				
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	***p* sex**
Age T1	14.10	3.39	8.01	22.79	14.63	3.75	8.01	24.55	.27
Age T2	16.10	3.40	10.02	24.83	16.60	3.77	9.92	26.62	.30
IQ T1	109.83	10.09	85.00	143.00	111.81	9.40	93.00	138.00	.13
Working Memory T1	0.79	0.17	0.13	1.00	0.86	0.12	0.38	1.00	p<.001
Feedback Learning T1	93.62	5.36	71.29	100.00	93.78	4.40	81.11	100.00	.81
Reading Fluency T2	98.02	14.51	64.00	120.00	97.72	15.46	58.00	120.00	.88
Mathematics T2	11.75	2.88	6.00	19.00	12.44	2.69	4.00	18.00	.06

Age at T1 correlated positively with reading fluency (r = .31, p < .001), working memory (r = .34, p < .001), and feedback learning performance (r = .47, p < .001). Age was also positively related to neural activity for the difference score Learning > Application in all 7 ROIs. Therefore, we corrected for age in further analyses. Even though mathematics scores were norm scores, i.e., scores relative to same-aged peers, there was still a small but significant correlation with age (r = .16, p = .018). We therefore also corrected for age in all further analyses with mathematics scores. Figure 3 shows the relations with age separated in categories for illustrative purposes.

Working memory at T1 correlated positively (corrected for age) with feedback learning performance (r = .33, p < .001), reading fluency (r = .15, p = .026) and mathematics (r = .25, p < .001) but not with neural activity at T1. IQ estimates correlated with mathematics norm scores (r = .32, p < .001, age-corrected) but not with the other measures. Note that the mathematics test (measured at T2) was part of the WISC/WAIS IQ test, although the estimated IQ scores (measured at T1) were measured two years earlier and based on only the subtests Similarities and Picture Comple-

tion. Finally, there was an age-corrected correlation between reading and mathematics scores (r = .20, p = .003).

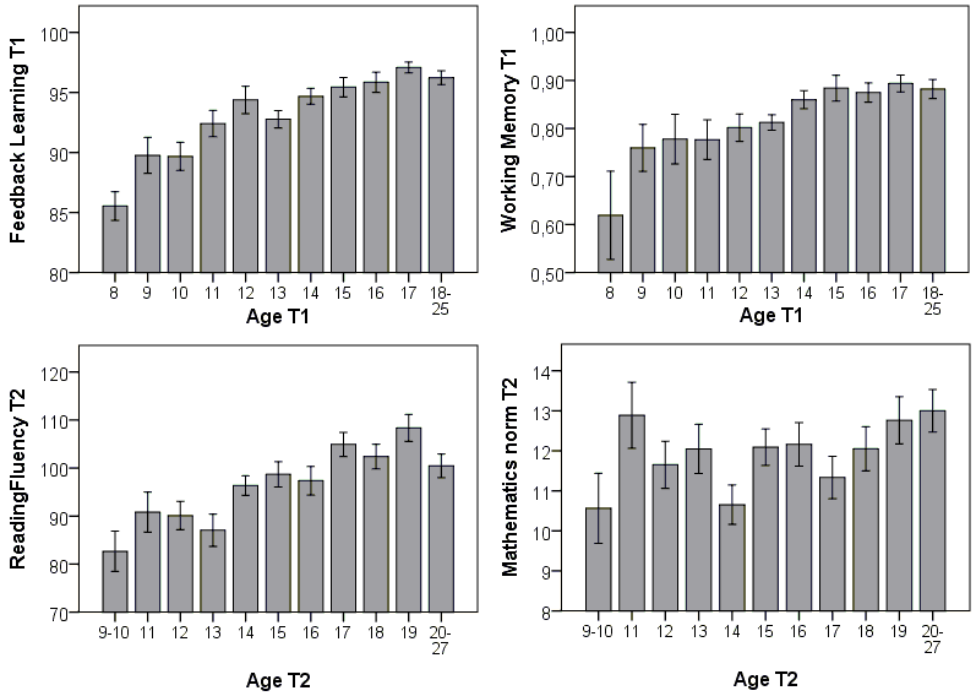


Figure 3: Display of age effects for feedback learning, working memory, reading and mathematics. Note that for T2 one participant was 9.92 years old, therefore the youngest age group at T2 was 9 and 10 years combined.

Predicting reading and mathematics performance at T2 from T1 feedback learning

We first investigated whether reading and mathematics performance at T2 could be predicted from behavioral performance on the feedback learning task at T1. A hierarchical regression with age at T1 entered as a first step and feedback learning performance at T1 as a second step, showed that in addition to age, feedback learning performance significantly predicted reading fluency and mathematics performance two years later (positive relation), see Table 2.

Table 2: Hierarchical linear regression models with age and feedback learning performance as significant predictors for reading and mathematics performance.

Steps	Predictor	B	SE B	β	*p*	*F*	R^2
Dependent: Reading Fluency T2							
1	Overall model					23.10***	.09
	Age T1	1.28	.27	.30	<.001***		
2	Overall model					16.96***	.13
	Age T1	.84	.30	.20	.005**		
	Feedback Learning T1	.67	.21	.22	.002**		
Dependent: Mathematics T2							
1	Overall model					5.47*	.02
	Age T1	.12	.05	.15	.020*		
2	Overall model					10.53***	.09
	Age T1	.02	.06	.02	.760		
	Feedback Learning T1	.16	.04	.28	<.001***		

* $p < .05$ ** $p < .01$ *** $p < .001$

Predicting reading and mathematics performance at T2 from T1 neural activity during feedback learning

Next, we assessed whether brain activity during feedback learning in 7 ROIs at T1 predicted reading and mathematics performance at T2. We performed hierarchical regressions with age at T1 as first step and neural activity in one of the 7 ROIs as second step. These analyses showed that in addition to age, reading fluency was predicted by left SPC and left sup-DLPFC (see Table 3). For mathematics performance at T2, activity in pre-SMA/ACC and right sup-DLPFC were significant predictors above age (see Table 4). For a visual representation of the relationship between right sup-DLPFC activity and mathematics performance, and left sup-DLPFC and reading fluency, see Figure 4.

We also tested whether neural activity for feedback learning explained additional variance in reading and mathematics above age and behavioral performance for feedback learning. We analyzed this with hierarchical regressions with age at T1 as first step, feedback learning performance at T1 as second step, and neural activity (per ROI) as third step. Neural activity explained additional variance for reading fluency (left sup-DLPFC remained significant ($\beta = .20$, $p = .004$), left SPC did not remain significant ($\beta = .12$, $p = .096$) and mathematics (pre-SMA/ACC remained significant ($\beta = .15$, $p = .029$), right sup-DLPFC did not remain significant ($\beta = .11$, $p = .113$)). This indicates that neural activity in left sup-DLPFC and pre-SMA/ACC explained unique variance in reading and mathematics over and beyond age and behavioral feedback learning performance.

Table 3: Hierarchical linear regression models for neural activity in left SPC and left sup-DLPFC as significant predictors above age for reading fluency.

Steps	Predictor	B	SE B	β	*p*	*F*	R^2
Dependent: Reading Fluency T2							
1	Overall model					23.10***	.09
	Age T1	1.28	.27	.31	<.001***		
2	Overall model					13.96***	.11
	Age T1	.97	.30	.23	.002**		
	SPC L	1.897	.90	.15	.036*		
1	Overall model					23.10***	.09
	Age T1	1.28	.27	.31	<.001***		
2	Overall model					17.56***	.14
	Age T1	.90	.28	.21	.002**		
	Sup-DLPFC L	2.78	.84	.23	.001**		

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 4: Hierarchical linear regression models for neural activity in pre-SMA/ACC and right sup-DLPFC as significant predictors above age for mathematics performance.

Steps	Predictor	B	SE B	β	*p*	*F*	R^2
Dependent: Mathematics							
1	Overall model					5.47*	.02
	Age T1	.12	.05	.15	.020*		
2	Overall model					6.31**	.05
	Age T1	.08	.05	.10	.159		
	Pre-SMA/ACC	.54	.21	.18	.009**		
1	Overall model					5.47*	.02
	Age T1	.12	.05	.15	.020*		
2	Overall model					4.73*	.03
	Age T1	.09	.05	.11	.120		
	Sup-DLPFC R	.38	.19	.14	.049*		

* $p < .05$ ** $p < .01$ *** $p < .001$

Adding working memory and IQ as control variables

To assess whether the relationship between feedback learning and reading and mathematics performance could be explained by individual differences in working memory, we tested whether the above effects remained significant when analyzing a hierarchical regression with age as a first step, working memory and IQ at T1 as a second step, and feedback learning performance or neural activity as a third step. Most analyses remained significant: For reading fluency, feedback learning performance was still a significant predictor ($\beta = .20$, $p = .011$) over age ($\beta = .31$, $p < .001$), working memory ($\beta = .16$, $p = .025$) and IQ ($\beta = -.19$, $p = .763$). Reading fluency was also still predicted by left sup-DLPFC ($\beta = .21$, $p = .002$), over age, IQ and working memory. Left SPC, howev-

er, was not a significant predictor anymore (β = .13, p = .065) after adding working memory and IQ. In this model, working memory was a significant predictor (β = .16, p = .025) but IQ was not (β = -.19, p = .763), indicating the lack of significance for left SPC is due to the addition of working memory to the model. For mathematics, feedback learning performance remained a significant predictor (β = .18, p = .015) over age (β = .16, p = .018), working memory (β = .22, p = .001) and IQ (β = .30, p < .001). Pre-SMA/ACC (β = .15, p = .023) was also still significant over age, IQ and working memory, but right sup-DLPFC (β = .12, p = .065) was only marginally significant after adding age, IQ and working memory. Both working memory (β = .22, p = .001) and IQ (β = .30, p < .001) were significant predictors in this model. Together, these results indicate that some of the effects of feedback learning activity are explained by working memory and IQ, but for others feedback learning performance and neural activity explained unique variance that was not explained by working memory or IQ.

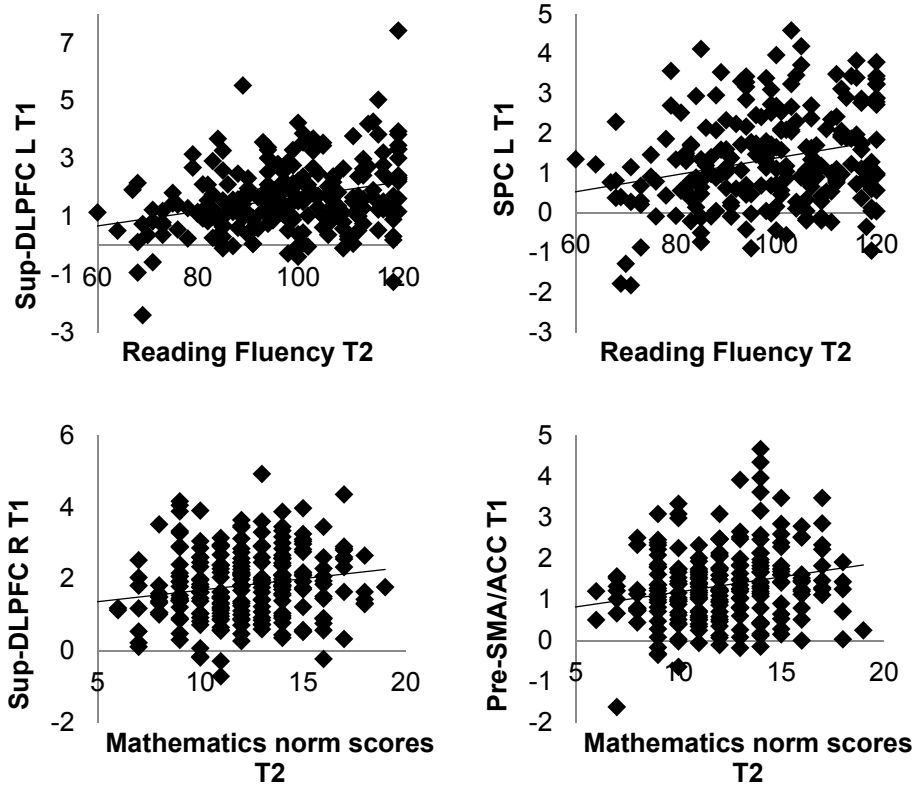


Figure 4: Scatterplot of the significant relationships between reading and mathematics performance at T2 and neural activity at T1 for the contrast Learning > Application.

Mathematics raw scores

All prior analyses used mathematics norm scores. To investigate whether results were also present when using raw scores, we also performed the analyses with feedback learning performance and neural activity as predictors for raw mathematics scores. Because the younger age group (10-15, n = 116) performed the mathematics test from the WISC-III and the older group (16-27, n = 112) the WAIS-III, these age groups were analyzed separately. The results showed that effects were only present in the younger adolescents but not the in the older adolescent/adult group. That is, for the youngest group, mathematics performance was predicted above age by feedback learning performance (β = .14, p = .027) and by pre-SMA/ACC activity (β = .20, p = .033) and there was only a trend for right sup-DLPFC activity (β = .15, p = .094). None of the effects were significant for the participants who were 16 years and older.

Discussion

In this study we investigated whether performance and neural activity during a feedback learning paradigm, used to study learning processes in a controlled laboratory setting, could predict indices of real-world learning performance in school two years later (reading and mathematics performance). The results of this study showed that 1) Feedback learning performance predicted both reading and mathematics performance two years later, 2) Neural activity during feedback learning in left sup-DLPFC and left SPC predicted reading fluency, and neural activity in right sup-DLPFC and pre-SMA/ACC predicted mathematics performance two years later, 3) Left sup-DLPFC and pre-SMA/ACC predicted unique variation in school performance over behavioral testing alone, and 4) Relations between feedback learning performance and neural activity and school performance remained significant when controlling for individual differences in working memory capacity and IQ.

Relation between feedback learning performance and school performance

For both reading and mathematical ability, we found that performance could be predicted by feedback learning performance two years earlier. Possibly, this relation can be explained by underlying individual differences in executive functions. It is well conceptualized that both feedback learning and school performance are related to executive functions (Diamond, 2013). Especially working memory was expected to be an important underlying factor, given that WCST performance (a complex feedback learning task) in a previous study was predicted by working memory in children (Huizinga et al., 2006) and adults (Miyake et al., 2000). Miyake et al. (2000) additionally found that switching was predictive for WCST performance, but this was not replicated in the child-aged sample of Huizinga et al. (2006). Consistent with these prior findings, we found a positive correlation with working memory performance and feedback learning, as well as with reading and mathematics performance, also when controlling for age differences. However, even

when adding working memory as a predictor to the model, feedback learning performance predicted unique variance for both reading and mathematics, suggesting that working memory may explain a part of, but not all variance. We investigated whether differences in general intelligence might explain the relation between feedback learning and school performance, but there was still a significant prediction of reading and mathematics scores by feedback learning when controlling for IQ.

Relation between neural activity for feedback learning and school performance

An important question tested in this study was whether neural activity could predict reading and mathematics performance two years later, and whether neural activity could provide additional information over behavioral testing alone. This was based on prior studies showing that neural measures can predict reading (Hoeft et al., 2007; Maurer et al., 2009) and mathematics performance (Dumontheil & Klingberg, 2012). Consistent with these studies, we found evidence for a relation between neural activity for feedback learning and reading and mathematics ability. First, we found that left sup-DLPFC and left SPC activity predicted reading ability. These findings fit with earlier research showing that a mostly left-lateralized network including DLPFC is involved during reading tasks (Ferstl et al., 2008). Second, right sup-DLPFC and pre-SMA/ACC predicted mathematics ability two years later. This fits with meta-analyses showing involvement of pre-SMA/ACC and DLPFC during arithmetical tasks (Arsalidou & Taylor, 2011; Houdé et al., 2010). Notably, for all areas we found a positive relation, indicating that increased activity predicts better performance on reading or mathematics tests. With the current design, it is not possible to determine whether higher activity might indicate better functioning or perhaps earlier maturation of these regions. Future research could build on this study by analyzing longitudinal fMRI measures and data on structural brain development.

In addition, we performed analyses to assess whether neural measures provided unique information that cannot be captured by behavioral testing alone. The regions that remained significant predictors when controlling for behavioral feedback learning were left sup-DLPFC for reading and right sup-DLPFC and pre-SMA/ACC for mathematics. This indicated that assessing feedback learning ability is useful for predicting reading and mathematics, but adding neural measures in addition to behavioral assessment further enhanced predictive ability. The finding that neural activity measures have added value over behavioral testing alone fits with earlier studies for the prediction of reading (Hoeft et al., 2007) and mathematics (Dumontheil & Klingberg, 2012; Hoeft et al., 2007).

Prior research suggested that working memory is an important component of both feedback learning (Miyake et al., 2000) and reading and mathematics (Alloway & Alloway, 2010), therefore it was possible that working memory is the underlying factor explaining these relations. Indeed, the prediction of reading performance from left SPC activity was no longer significant when controlling for working memory, indicating that working memory might underlie this

relation. However, even when we controlled for working memory and IQ, there was still a significant prediction of reading fluency from feedback learning performance and activity in left sup-DLPFC, and for prediction of mathematics from feedback learning performance and activity in pre-SMA/ACC. This indicates that although working memory plays a role in the relation between feedback learning and reading and mathematics, there is still unique variation in reading and mathematics that is explained by neural activity during feedback learning. Other aspects of feedback learning performance that might be relevant for learning in school settings, are for instance the capacity to monitor one's actions and keep track of performance feedback, ignoring irrelevant aspects of the task, perceived competence and motivation (Fortier, Vallerand, & Guay, 1995; St Clair-Thompson & Gathercole, 2006). Future research is needed to examine this in more detail.

An interesting laterality difference was observed for predicting reading and mathematics in superior DLPFC. That is to say, we found that activity in left superior DLPFC during feedback learning predicted reading ability, whereas activity in right superior DLPFC predicted mathematical ability. The left-right distinction fits nicely with the well-established finding that the neural network for learning is left-lateralized (Frost et al., 1999). There is no conclusive evidence for a possible right-lateralized network for mathematics. The current findings suggest that left-right hemispheric differences may be an important factor explaining differences between reading and mathematics related school processes.

Limitations and future directions

There are several limitations to this study. First, school performance can be measured in many ways. In this study, we measured only two short, well-validated measures for reading and mathematics. Future research could build on this study by relying on a more extensive assessment of school performance involving multiple measures. Second, we only collected reading fluency and mathematics data at the second time point but not at the first time point. An interesting question would be to investigate whether feedback learning and brain measures can predict reading and mathematics even better than tests for reading and mathematics themselves. On the other hand, an advantage of measuring feedback learning or other executive functioning tasks is that it captures abilities that are essential to both reading and mathematics. Third, IQ was assessed with only two subtests of the WISC/WAIS. A more comprehensive assessment of IQ might give a more definite answer to the question whether the relation between feedback learning and school performance is driven by underlying differences in general intelligence. Fourth, mathematics was assessed with the WISC for younger participants (10-15 years at T2) and with the WAIS for older participants (16-27 years at T2). When we performed the analyses with mathematics raw scores rather than norm scores (scores relative to same-aged peers), we needed to perform the analyses in separate age groups. These analyses showed that the prediction of mathematics scores from behavioral performance and neural activity for feedback learning was only present in the youngest age group (10-15 years). One tentative interpretation is that prediction is stronger in the

younger age groups, when brain maturation is still undergoing major changes (Giedd & Rapoport, 2010). Alternatively, it is possible that the WISC scores are more sensitive for picking up change than the WAIS scores. Future studies should use a wider battery of tests to test these competing hypotheses in more detail.

Conclusion

In conclusion, this study found contributions of feedback learning performance and neural activity in predicting school outcomes two years later. This provides evidence that studying learning processes through simplified laboratory tasks provides at least some relevance for real-world learning. In addition, we showed that neural measures explain unique variance in school outcomes two years later that is not captured by behavioral testing of executive functions alone.

Chapter 8

The link between testosterone and amygdala connectivity in adolescent alcohol use

This chapter is based on:

Peters, S., Jolles, D.J., Van Duijvenvoorde, A.C.K., Crone, E.A., & Peper, J.S. (2015). The link between testosterone and amygdala–orbitofrontal cortex connectivity in adolescent alcohol use. Psychoneuroendocrinology, 53, 117-126.

Abstract

Alcohol consumption is one of the most problematic and widespread forms of risk taking in adolescence. It has been hypothesized that sex hormones such as testosterone play an important role in risk taking by influencing the development of brain networks involved in emotion and motivation, particularly the amygdala and its functional connections. Connectivity between the amygdala and the orbitofrontal cortex (OFC) may be specifically related to alcohol use, given the association of this tract with top-down control over behavioral approach tendencies.

In line with this, prior studies in adults indicate a link between alcohol use and functional connectivity between the amygdala and the OFC, as well as between testosterone and amygdala-OFC connectivity. We consolidated these research lines by investigating the association between alcohol use, testosterone and resting state functional brain connectivity within one large-scale adolescent sample (N = 173, aged 12-25 years). Mediation analyses demonstrated an indirect effect of testosterone levels on alcohol use through amygdala-OFC intrinsic functional connectivity, but only in boys. That is, increased testosterone in boys was associated with reduced amygdala-OFC connectivity, which in turn was associated with increased alcohol intake. This study is the first to demonstrate the interplay between adolescent alcohol use, sex hormones and brain mechanisms, thus taking an important step to increase our understanding of the mechanisms behind this form of adolescent risk taking.

Introduction

Adolescents are prone to increased risk taking and impulsive behavior (Steinberg, 2008). Although risk taking behavior in adolescence can be adaptive (Crone & Dahl, 2012), it is also associated with negative consequences for health and safety. Alcohol use is one of the most widespread forms of risk taking in adolescence (Hibell et al., 2012). Adolescent alcohol use is associated with impaired cognitive functioning and school performance (Zeigler et al., 2005) and alcohol-related problems in adulthood (Grant et al., 2006). Understanding the mechanisms behind alcohol use in adolescence is an important step towards preventing alcohol-related problems.

Risk taking behavior, including alcohol use, has been related to the dramatic rise in sex hormones during puberty (Forbes & Dahl, 2010). Indirect evidence for a link between increased sex hormone production during puberty and increased alcohol use comes from studies showing that adolescents with advanced pubertal maturation show relatively higher levels of alcohol intake (Biehl, Natsuaki, & Ge, 2007; Bratberg, Nilsen, Holmen, & Vatten, 2005; Westling, Andrews, Hampson, & Peterson, 2008). Moreover, a direct association was found between higher production of the sex hormone testosterone and an earlier onset of alcohol consumption in adolescent boys (de Water et al., 2013). Consequently, a prominent hypothesis predicts that sex hormones affect risk taking by influencing the development of limbic brain areas involved in emotion and motivation (Peper & Dahl, 2013).

The amygdala is one such limbic brain area that plays a key role in adolescent functional brain organization (Scherf, Smyth, & Delgado, 2013) and is associated with both testosterone and alcohol use. For instance, the amygdala is among the brain areas with the highest density of androgen receptors as shown in animal studies (Simerly, Chang, Muramatsu, & Swanson, 1990). Moreover, the amygdala response to emotional faces can be modulated by testosterone (Derntl et al., 2009; Hermans, Ramsey, & van Honk, 2008; Manuck et al., 2010; Stanton, Wirth, Waugh, & Schultheiss, 2009). With regard to alcohol, the amygdala is one of the key regions of interest in animal studies on alcohol use (McBride, 2002) and human research shows that alcohol ingestion leads to reduced amygdala activity for fearful/angry faces (Gilman, Ramchandani, Crouss, & Hommer, 2012; Gilman, Ramchandani, Davis, Bjork, & Hommer, 2008; Sripada, Angstadt, McNamara, King, & Phan, 2011) and a lower amygdala response to fearful faces is linked to increased risk for future alcohol abuse (Glahn, Lovallo, & Fox, 2007).

Since the amygdala is highly interconnected with other brain regions (Cole, Pathak, & Schneider, 2010), it is important to also take into account the functional connections of the amygdala. The connection with the orbitofrontal cortex (OFC) is of particular interest, as the OFC is directly connected to the amygdala through the uncinate fasciculus (Von Der Heide, Skipper, Klobusicky, & Olson, 2013), an association tract that develops well into adolescence (Lebel & Beaulieu, 2011). In adults, it has been demonstrated that functional connectivity between the amygdala and the OFC during emotional face processing was reduced after alcohol ingestion

(Gorka, Fitzgerald, King, & Phan, 2013). Interestingly, testosterone administration (Bos, Hermans, Ramsey, & van Honk, 2012; van Wingen, Mattern, Verkes, Buitelaar, & Fernandez, 2010) and high endogenous testosterone (Spielberg et al., 2014) showed similar reducing effects on amygdala-OFC functional connectivity.

We therefore argue that it is vital to study the interplay between testosterone and amygdala-OFC connectivity in adolescents to explain individual differences in alcohol consumption. We tested this in a large cross-sectional adolescent sample using a resting state paradigm, which is a valuable tool to investigate functional networks in the developing brain (Uddin, Supekar, Ryali, & Menon, 2011) and has several advantages over task-based brain activity, including high test-retest reliability (Zuo & Xing, 2014) and broader generalizability. We hypothesized that higher levels of testosterone would be associated with increased alcohol consumption, mediated through lower amygdala-OFC connectivity.

Methods

Participants

The included sample consisted of 173 healthy participants (86 girls, 87 boys), between 12.05 and 25.95 years old (M = 15.85, SD = 3.10), for whom data on alcohol consumption, brain imaging and hormonal samples were available. Because this study was part of a larger project also involving younger participants, we collected complete resting state scans, high-resolution functional scans and T1 scans for 295 participants between 8 and 25 years old. Only participants who were twelve years and older (N = 209) were asked to fill out the alcohol questionnaire and younger participants were therefore excluded from further analyses. Other reasons for exclusion were: not completing the alcohol questionnaire (N = 11), missing testosterone levels (N = 16), excessive movement in the MRI scanner (> 3 mm; N = 3) and excessive micromovements (> 20% of volumes with > .05 mm movement; N = 5). Note that several participants were excluded for multiple reasons, e.g. both excessive movement and missing testosterone data.

Participants were recruited through local schools and advertisements. IQ was estimated with two subtests of the WAIS-III or WISC-III (Similarities and Block Design). All estimated IQ scores were within the normal range (M = 109.39, SD = 9.67, range: 80-135) and there was no correlation with age (r = .05, p = .53). Adults (18 years and older) received payment (60 euros) for participation, children received presents and their parents received 30 euros for travel reimbursement. The study was approved by the Institutional Review Board at the Leiden University Medical Center. The participants (or in case of minors, participant's parents) signed a written informed consent. All anatomical MRI scans were reviewed and cleared by a radiologist. None of the participants reported neurological or psychiatric disorders or current use of psychotropic medication. See Table 1 for the demographics for male and female participants in the sample.

Table 1: Demographic variables for boys and girls separately.

	Males (N = 87)		Females (N = 86)	
	Mean	**SD**	**Mean**	**SD**
Age	16.10	3.39	15.60	2.77
IQ	109.63	9.85	109.07	9.66
Lifetime alcohol use	30.82	40.30	28.69	36.23
Recent alcohol use	7.97	14.70	5.68	10.71
Testosterone (pmol/l)	258.37	164.79	26.66	38.71
Right amy-OFC connectivity	0.50	0.79	0.69	0.71
Left amy-OFC connectivity	0.66	0.78	0.77	0.77

Testosterone levels

Testosterone levels were extracted from saliva samples (de Water et al., 2013; Peper, Mandl, et al., 2013). Participants collected saliva by passive drool at home on the day of the MRI scan. In order to minimize effects of diurnal fluctuations of hormonal levels, saliva samples were collected immediately after waking up in all participants. Girls using contraceptives (N = 17) collected saliva on the last day of the stop period (day 7) and post-menarchal girls not using contraceptives (N = 46) collected saliva on the 7th day of the menstrual cycle. Testosterone levels from saliva were determined by isotope dilution - online solid phase extraction liquid chromatography – tandem mass spectrometry (ID-XLC-MS/MS; Intra-essay coefficient of variation (CV) was 11% and 4%, at 10 and 140 pmol/L, respectively and inter-assay CV was 8% and 5%, at 31 and 195 pmol/L, respectively. Testosterone levels were not normally distributed, so a log-transformation of the scores was used for further calculations.

Alcohol Questionnaire

Participants filled in an on-line questionnaire at home on recent and lifetime alcohol use (Ames et al., 2007; de Water et al., 2013; Thush et al., 2008). Self-report measures of alcohol use have been shown to be reliable if confidentiality is ensured (Brener et al., 2002; Sobell & Sobell, 1990). The instructions explicitly stated that participant's answers were confidential and would not be disclosed to anyone. Participants were instructed to fill out the questionnaire at a time point as close as possible to the MRI scan.

Lifetime alcohol use was reported as the lifetime amount of glasses consumed on an 11-point scale (0, 1–10, 11–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80, 81–90, and > 90). Bottles and cans of alcohol had to be counted as 1.5 glasses, because these contain more of the alcoholic beverage than a standard glass in the Netherlands (Thush et al., 2008). Recent alcohol use was reported as the number of glasses of alcohol participants had consumed over the past 30 days on a 10-point scale (0, 1–2, 3–4, 5–6, 7–10, 11–15, 16–20, 21–30, 31–50, and > 50). To create a scale variable,

the ordinal data on quantity of alcohol use were converted by calculating the mean of the answer (for > 50 and > 90, 51 and 91 were used, respectively). On average, participants had consumed 29.76 glasses of alcohol in their lives (*SD* = 38.24) and 6.83 glasses in the last month (*SD* = 12.89). There were no gender differences or gender x age interactions in alcohol consumption. See Figure 1 for the amount of glasses consumed per age group and per gender. Adult participants between 18-25 years were grouped together due to a lower sample size in that age group (sample sizes: 12y: *N* = 27, 13y: *N* = 33, 14y: *N* = 25, 15y: *N* = 20, 16y: *N* = 16, 17y: *N* = 22, 18-25y: *N* = 30).

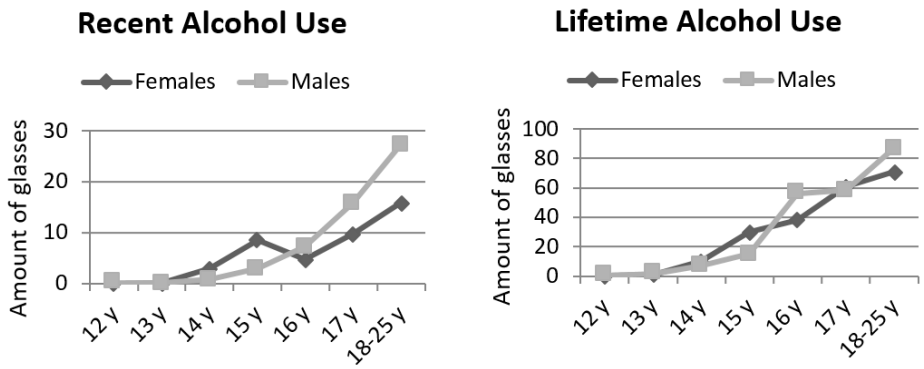


Figure 1: Glasses of alcohol consumed over the past month (recent) and over lifetime, per age and per gender.

MRI data Acquisition

MRI scans were acquired with a standard whole-head coil on a Philips 3.0 Tesla MRI scanner. Functional resting state scans were acquired with T2*-weighted echo-planar imaging (EPI). The first two volumes were discarded to allow for equilibration of T1 saturation effects. The following scan parameters were used: 140 volumes; 38 slices; sequential acquisition; TR = 2200 ms, TE = 30 ms; flip angle = 80°; FOV = 220 x 220 x 114.67 mm; slice thickness = 2.75 mm. A high-resolution anatomical scan (T1-weighted; 140 slices; TR = 9.76 ms; TE = 4.59 ms; flip angle = 8°; FOV = 224 × 177.33 × 168 mm; in-plane resolution = 0.875 x 0.875 mm; slice thickness = 2 mm) and a high-resolution T2*-weighted gradient echo EPI scan (84 slices; TR = 2200 ms; TE = 30 ms; flip angle = 80°; FOV = 220 x 220 x 168 mm; in-plane resolution = 1.96 x 1.96; slice thickness = 2 mm) were acquired after the resting state scan. Participants were instructed to close their eyes during the resting state, and a video was presented during the structural and high-resolution T2*-weighted scan. Before the MRI scan, participants were accustomed to the MRI environment and sounds with a mock scanner.

FMRI data preprocessing

All resting state scans were submitted to visual quality control to check for artifacts. Next, FMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.98, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The following preprocessing steps were used: motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002); non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 5mm; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering of 100 s (Gaussian-weighted least-squares straight line fitting, with sigma = 50 s). The resting state scan was registered with FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001) to the high resolution T2*-weighted scan, which in turn was registered to the T1-weighted scan, and the T1-weighted scan was registered to the 2 mm MNI-152 standard space image.

FMRI data analysis

A seed-based correlation approach (Fox & Raichle, 2007) was used to find brain regions with functional connectivity to the amygdala. Amygdala masks were obtained by using atlas-based masks of left and right amygdala (Automatic Anatomical Labeling atlas; see Figure 2).

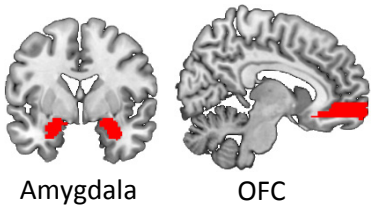


Figure 2: Amygdala and OFC masks based on anatomical masks (Automatic Anatomical Labeling atlas)

Amygdala masks in MNI space were transformed to native space (each individual's resting state scan) with a binary threshold of 0.5. Next, the mean time courses were extracted from each individual's left and right amygdala, i.e. all voxels located within the amygdala mask. These mean time courses were entered as a regressors in a general linear model (separately for left and right amygdala), with nuisance regressors for the white matter signal and CSF signal (obtained from a bilateral 4 mm sphere in white matter (left: x = 54, y = 44, z = 44; right x = 35, y = 44, z = 44) and CS*F* (left: x = 59, y = 55, z = 50; right: x = 30, y = 55, z = 50), global signal, and six motion parameters (rigid body: three translations and three rotations). For participants with excessive micromovements (> .05 mm) between volumes, we included additional regressors (binary for all volumes with movement > .05) to remove the specific volumes where micromovements occurred

from the analysis. Participants for whom more than 20% of volumes were affected by micromovements (> .05 mm) were excluded from further analyses. Next, we performed a first-level analysis with FEAT for each participant. The individual parameter estimate maps and within-subject variance maps were resliced into MNI space and used for higher-level analyses. Higher-level analysis was carried out using FLAME (FMRIB's Local Analysis of Mixed Effects) stage 1 (Beckmann, Jenkinson, & Smith, 2003; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004; Woolrich, 2008). Z statistic images were thresholded with an initial cluster-forming threshold of Z > 2.3 and a (corrected) cluster threshold of $p < .05$ (Worsley, 2001).

ROI analyses

Region-of-interest analyses on the a-priori selected OFC were performed on an OFC anatomical mask (based on Autonomic Anatomical Labeling: Medial Orbital Frontal Gyrus; see Figure 2) in which left and right OFC were combined. OFC masks in MNI space were, similar to the amygdala masks, transformed to native space with a binary threshold of 0.5. Next, we extracted Z-scores for amygdala connectivity from the OFC. The results were further analyzed with SPSS 19 for the indirect effect of testosterone on alcohol use via amygdala-OFC connectivity, with age as a covariate of no interest. All analyses were performed for girls and boys separately. This is preferred in puberty related research given the possible differential effects of testosterone in boys and girls (Bramen et al., 2011) and the difference in timing at which puberty emerges, approximately 1.5 years earlier for girls than boys (Shirtcliff, Dahl, & Pollak, 2009). Moreover, there was a substantial gender difference in testosterone levels in our sample (t (128.67) = 14.59, $p < .001$).

To investigate the interplay between amygdala-OFC connectivity, alcohol use, and testosterone, we performed mediation analyses (Preacher & Hayes, 2008). We investigated the association between testosterone and alcohol use and examined whether amygdala-OFC connectivity mediated this association. This approach is similar to previous work demonstrating that the association between testosterone and adolescent risk taking is mediated by OFC morphology (Peper, Koolschijn, & Crone, 2013). Throughout all analyses, we corrected for age. We used the bootstrapping method of Preacher and Hayes (Preacher & Hayes, 2004, 2008) to test the indirect effect of testosterone on alcohol use (through amygdala-OFC connectivity) for significance. With this method it is possible to test for indirect effects even in the absence of direct effects (Hayes, 2009). A bootstrapped mediation analysis uses re-sampling of raw data to estimate the confidence intervals (CI) to formally test the indirect effects of which the mediation model consists.

Whole-brain analyses

In addition to the ROI analyses, we performed whole-brain analyses (using higher-level FEAT in FSL) to examine whether effects were specific for the coupling between the amygdala and the OFC or whether other regions were implicated as well. We examined the relation between amygdala connectivity and testosterone, as well as the relation between amygdala connectivity and

alcohol use (lifetime and recent use). Age was included as covariate of no-interest. Z statistic images were thresholded with an initial cluster-forming threshold of Z > 2.3 and a (corrected) cluster threshold of $p < .05$ (Worsley, 2001). Left and right amygdala were analyzed separately.

Results

Mediation effect in boys

To test the hypothesis that testosterone influences alcohol use through an effect on amygdala-OFC connectivity, we used mediation analyses with alcohol use as outcome variable and amygdala-OFC connectivity as mediator variable (corrected for age). For boys, the results showed that testosterone influenced alcohol use through amygdala-OFC connectivity, for recent alcohol use (right and left amygdala-OFC connectivity: 95% CI = 0.16 – 5.50 and 0.12 - 4.69, respectively) and lifetime alcohol use (right amygdala-OFC connectivity only (95% CI = 0.34 – 10.61) (Figure 3). That is, higher testosterone levels were associated with less functional connectivity between the amygdala and OFC (path a), and less amygdala-OFC connectivity was in turn associated with more alcohol use (path b). There was no direct relation between testosterone and alcohol consumption (path c and c') (see Figure 3 for the statistical values), but a significant direct effect is not a prerequisite for a significant mediation effect (Hayes, 2009; MacKinnon, Krull, & Lockwood, 2000; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Rucker, Preacher, Tormala, & Petty, 2011; Shrout & Bolger, 2002; Zhao, Lynch, & Chen, 2010).

No mediation effects in girls

No significant effects for any of the mediation paths were found for girls, even when girls using contraceptives ($N = 17$) were excluded from the analyses. Note that there was no correlation between age and amygdala-OFC connectivity for either girls or boys. Therefore, the follow-up analyses described in the next paragraphs were performed for boys only.

Follow-up analyses in smaller age ranges

Because of the relatively large age range in our sample, we also tested for mediation effects in separate age groups in boys. We created three age groups of equal size ($N = 29$, young adolescents 12.3-14.0 y, mid-adolescents 14.0-16.7 y, late adolescents/young adults 16.7-25.9 y). The mediation effect was only significant for the mid-adolescent group for right amygdala-OFC connectivity for recent use (95% CI = 0.19 – 5.69) and lifetime use of alcohol (95% CI = 0.23 – 10.56). The mediation in this mid-adolescent group showed comparable effects to the results in the group as a whole. For these analyses in smaller age ranges, similar results were found for mediation analyses performed with and without age correction. These analyses might indicate that the effects of testos-

terone on alcohol use via amygdala-OFC connectivity are mostly driven by the mid-adolescent age group.

A possible explanation for the lack of a direct effect of testosterone on alcohol use is that alcohol use is strongly related to age (recent: $r = .67$, $p < .001$; lifetime: $r = .68$, $p < .001$) and age in turn is strongly related to testosterone ($r = .50$, $p < .001$). Possibly, effects of testosterone are less pronounced due to the correction for age. When analyzing the relationship between testosterone and alcohol use in the smaller age groups (without age correction), this relation was significant for lifetime use in the young adolescent group ($r = .37$, $p = .05$, but this disappears with age correction: $r = .31$, $p = .11$) and for lifetime use in the late adolescent/young adult group ($r = .46$, $p = .01$, with age correction: $r = .44$, $p = .02$).

Age vs. testosterone effects on amygdala-OFC connectivity

We found no correlation between amygdala-OFC connectivity and age in either boys or girls. This correlation with age was not significant independent of whether we did, or did not control for testosterone levels. There was a significant correlation between amygdala-OFC connectivity and testosterone in boys, which remained significant when controlling for age (right: $r = -.26$, $p = .02$, left: $r = -.23$, $p = .03$). These results suggest that connectivity between the amygdala and the OFC is associated more with testosterone rather than age per se.

Whole-brain analyses

To examine the specificity of the mediation effects in boys to the OFC, we investigated effects of alcohol use and testosterone on amygdala connectivity with the rest of the brain. We performed an analysis with testosterone as a regressor, including age as a regressor of no interest. The results indicated reduced functional connectivity between the right amygdala and the orbitofrontal cortex and other medial frontal areas with relatively higher levels of testosterone (see Figure 4, Table 2). The pattern in the left amygdala was in the same direction, albeit less pronounced. Besides the medial frontal cortex, we did not find any other regions that showed an effect of testosterone on amygdala connectivity.

For the analyses with alcohol use as a regressor, we found reduced functional connectivity in boys between the right amygdala and the orbitofrontal cortex and other medial frontal areas for both recent and lifetime alcohol use (see Figure 4, Table 3). Similar but less pronounced effects were found for the left amygdala. There were no other regions that showed an effect of alcohol use on amygdala connectivity. These results provide further evidence for the hypothesis that testosterone and alcohol use are related to similar brain mechanisms and indicate that the effects are specific to amygdala-OFC coupling.

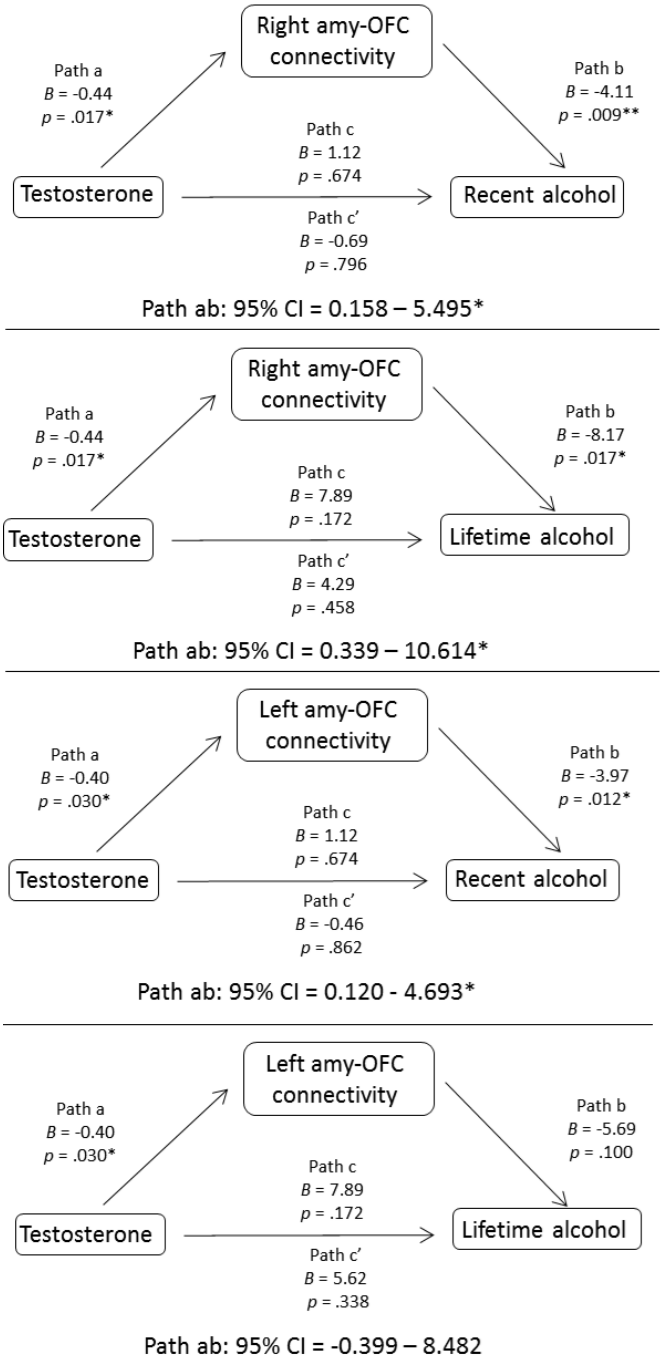


Figure 3: Mediation analyses for the relation between testosterone, amygdala-OFC connectivity and alcohol consumption for left and right amygdala-OFC connectivity and lifetime and recent alcohol use in boys.

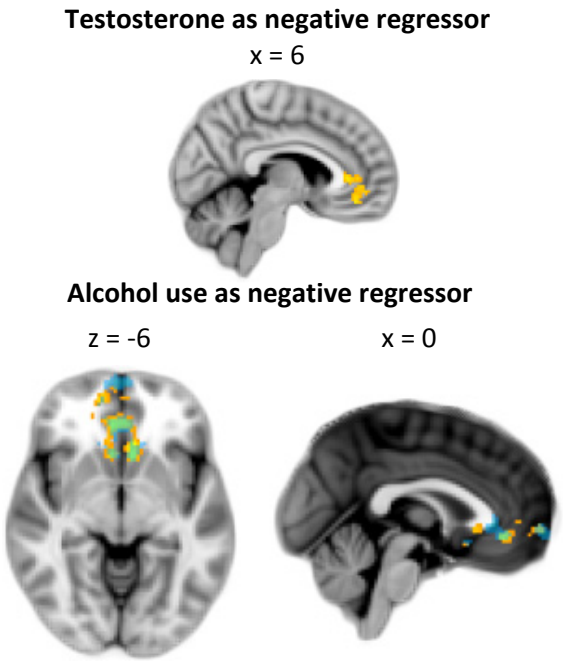


Figure 4: a) Right amygdala connectivity in boys with testosterone as a negative regressor and age as regressor of no interest. b) Right amygdala connectivity in boys with alcohol use as a negative regressor and age as regressor of no interest (yellow=recent alcohol use, blue=lifetime alcohol use, green = overlap).

Table 2: MNI-coordinates for local maxima for right amygdala connectivity in boys with testosterone as a negative regressor and age as regressor of no interest.

Brain Region	Z	x	y	z	voxels
R cingulate gyrus	3.73	10	38	4	397
R frontal medial cortex	3.43	12	34	-10	s.c.
R frontal pole	3.26	20	42	-16	s.c.
R frontal medial cortex	3.2	6	40	-14	s.c.
R frontal pole	3.16	28	52	-12	s.c.
R frontal pole	3.12	22	32	-6	s.c.

Note: s.c. = same cluster

Table 3: Local maxima for connectivity with the right amygdala with alcohol use as a negative regressor and age as regressor of no interest.

Brain Region	Z	x	y	z	voxels
Recent alcohol use					
L subcallosal cortex	4.16	-6	24	-6	786
R subcallosal cortex	3.98	8	20	-10	s.c.
R frontal pole	3.74	4	58	-2	s.c.
R frontal medial cortex	3.68	14	52	-4	s.c.
R cingulate gyrus	3.66	6	38	-4	s.c.
Paracingulate gyrus	3.4	0	38	-8	s.c.
Lifetime alcohol use					
L frontal orbital cortex	4.57	-16	14	-12	853
R frontal orbital cortex	3.89	24	18	-12	s.c.
L paracingulate gyrus	3.72	-6	32	6	s.c.
L subcallosal cortex	3.63	-8	20	-10	s.c.
R cingulate gyrus	3.55	2	38	-6	s.c.
R frontal orbital cortex	3.45	20	22	-10	s.c.

Discussion

We investigated the association between testosterone, amygdala-OFC functional connectivity and alcohol use in a sample of 173 typically developing adolescents. In agreement with our hypothesis, increased testosterone levels were related to increased alcohol consumption, through a mediation effect on reduced intrinsic amygdala-OFC connectivity, but only in boys. In other words, higher testosterone levels in boys were associated with reduced connectivity between the amygdala and the OFC, and reduced amygdala-OFC connectivity in turn was related to increased alcohol use.

Amygdala-OFC connectivity and risk taking behavior

The results of this study fit well with prior research reporting reduced amygdala-OFC connectivity with increased testosterone using task-based functional connectivity in adults (Bos, Panksepp, Bluthe, & van Honk, 2012; van Wingen et al., 2010) and adolescents (Spielberg et al., 2014), as well as a prior study that found reduced task-based amygdala-OFC connectivity after alcohol ingestion in adult heavy social drinkers (Gorka et al., 2013). These results suggest that testosterone and alcohol use are linked to similar brain mechanisms. Because our sample included

a relatively large age range, we also performed these analyses in smaller age ranges. The findings indicated that the effects were most pronounced in mid-adolescence (14-16.7 years). In addition, we tested how amygdala-OFC connectivity changed with age. We did not find a correlation between amygdala-OFC connectivity and age, but there was a negative correlation in boys between amygdala-OFC connectivity and testosterone, controlling for age (but see also Gabard-Durnam et al., 2014).

The amygdala and the OFC are hypothesized to work together during the processing of both negative and appetitive emotional stimuli (Baxter & Murray, 2002). That is, it has been argued that the amygdala detects the valence of affective stimuli, while the OFC guides decisions and goal-directed behavior in reaction to affective and rewarding stimuli (Bechara, Damasio, & Damasio, 2000; Kringelbach, 2005). Prior studies showed that increased task-related connectivity between the amygdala and the OFC is associated with better emotion regulation and self-control (Banks, Eddy, Angstadt, Nathan, & Phan, 2007; Lee, Heller, van Reekum, Nelson, & Davidson, 2012) which is in turn associated with less risk taking behavior.

However, further research is warranted to understand the link between task-related versus intrinsic brain activity on behavior. Raichle (2010) argues that brain function should be seen as mostly intrinsic, rather than reflexive (task-related) in nature. In terms of energy allocation, this fits with the finding that task-related brain activity -compared to intrinsic activation- only uses 5% additional energy (Raichle, 2010). It is possible that decreased intrinsic connectivity between the amygdala and the OFC biases a person towards acting in a less top-down controlled and more risky way.

On the other hand, a recent study comparing a group of 18 risk taking individuals with 18 non risk taking individuals (DeWitt, Aslan, & Filbey, 2014) did not find evidence for reduced intrinsic amygdala-OFC connectivity in a group of risk taking individuals, but instead found increased connectivity between the amygdala and left cingulate gyrus, left precuneus, right middle frontal gyrus and right inferior parietal lobule. Several factors might explain these opposing findings: i) DeWitt et al. (2014) compared a group of risk taking individuals with a group of non risk taking individuals (based on an assessment of many different forms of risk taking), rather than only alcohol use as a real-life form of risk taking measured on a continuous scale. ii) DeWitt et al.'s sample was relatively small compared to the current study, in an age range of 12 to 17 with both boys and girls. Future studies should unravel the heightened versus reduced amygdala connectivity in relation to increased risk taking within separate networks of intrinsic brain activity.

Note that with the functional connectivity analyses used in this study, it remains unclear whether there is less 'top-down' connectivity from the OFC to the amygdala or more 'bottom-up' connectivity from the amygdala to the OFC (i.e., the direction of connectivity). We examined functional connectivity using the amygdala as a seed region because of its high density of androgen receptors (Simerly et al., 1990), the correlation between amygdala activation and

testosterone (Derntl et al., 2009; Hermans et al., 2008; Manuck et al., 2010; Stanton et al., 2009), as well as the relation between amygdala activation and alcohol (Gilman et al., 2012, 2008; McBride, 2002; Sripada et al., 2011). Future research should further explore effects of alcohol and testosterone on functional brain connectivity, by –for instance- focusing on the OFC as a seed region. As a result, a broader network of brain regions might be revealed in relation to testosterone and alcohol use.

In sum, our findings suggest that a decreased coupling between the amygdala and the OFC which relates to increased testosterone, may be instrumental in explaining risk taking behavior in boys, such as adolescent alcohol use through the inability to assert top-down control over behavioral approach tendencies.

Indirect effect of testosterone on alcohol use

Although we found evidence for an indirect relationship between testosterone and alcohol consumption (through amygdala-OFC connectivity) in boys, a direct effect of testosterone on alcohol use was not significant. From a statistical point of view, a significant direct effect is not necessary for a significant mediation effect (MacKinnon et al., 2000; MacKinnon et al., 2002; Shrout and Bolger, 2002; Hayes, 2009; Rucker et al., 2011; Zhao et al., 2010). For instance, it is possible that multiple indirect effects are influencing alcohol use, which can explain why a direct effect was not found (Hayes, 2009). The lack of a direct effect was, however, contrary to our expectations as prior studies indicated an effect of advanced pubertal maturation on alcohol intake (Biehl et al., 2007; Bratberg et al., 2005; Westling et al., 2008).

A possible explanation can be found in the fact that testosterone can either have slow effects through genetic mechanims, such as on the structural organization and functional coupling of brain pathways, or fast nongenomic effects in the order of minutes to seconds, such as activitation of brain areas after the administration of testosterone (Bos, Panksepp, Bluthe, & van Honk, 2012). Possibly, endogeneous and gradual increases of testosterone during adolescence might not have influenced a person's immediate decision about consuming alcohol, but might have indirectly influenced brain mechanisms involved in less behavioral control and increased approach-related behavior. Indeed, prior studies in adolescents showed that adolescent testosterone relates to gray matter and white matter tracts within limbic brain areas (Herting, Maxwell, Irvine, & Nagel, 2012; Paus et al., 2010; Peper, Koolschijn, et al., 2013).

No relation between testosterone, amygdala-OFC connectivity and alcohol use in girls

It is interesting that we only found a relation between testosterone, brain connectivity and alcohol intake in boys, but not in girls. Prior research also showed different effects of testosterone in boys than in girls for risk taking behavior (Peper, Koolschijn, et al., 2013) and brain development (Bramen et al., 2011; Nguyen et al., 2013). Note that we did not find any gender difference in alcohol use in our sample, partly similar to prior studies which reported no gender difference in

alcohol intake until the age of about 17, after which men tend to drink more than women (Witt, 2007).

However, Van Wingen et al (2010) and Bos et al. (2012) did show an effect of testosterone administration on task-based amygdala-OFC connectivity in (adult) women. The relatively low endogenous levels of testosterone in adolescent girls compared to boys were possibly not able to reliably affect functional connectivity during a resting state scan.. In addition, although we corrected for menstrual cycle in girls at the time of saliva collection (all post-menarchal girls collected saliva at day 7 within their cycle), it was not feasible to also plan the MRI scan on the 7th day. Consequently, testosterone levels at the time of scanning in girls could have been different from the testosterone levels in the samples we collected.

Limitations and directions for future research

It is important to note that the assessment of alcohol intake relied on self-report. Self-report may lead to both over- and underestimations of actual alcohol intake. Especially for lifetime alcohol use, participants may have difficulty estimating how much alcohol they consumed in their lives. On the other hand, lifetime and recent alcohol use were highly correlated, and prior studies have shown that self-report measures of alcohol use are reliable if confidentiality is ensured (Brener et al., 2002; Sobell & Sobell, 1990), which was the case in our study. Although we excluded participants who were diagnosed with a psychiatric disorder, we did not include a measure for alcohol abuse. In future studies, it is important to additionally investigate age of onset, criteria for alcohol abuse and binge drinking behavior, as binge drinkers are especially at risk for negative consequences related to alcohol use (Wechsler & Nelson, 2001).

Another possible direction for future research would be a large-scale longitudinal study. The current study was cross-sectional, therefore we could only evaluate whether there was an association between testosterone and alcohol use, but it was not possible to assess whether future alcohol use can be predicted based on current testosterone levels. Future longitudinal studies should unravel factors predicting alcohol use. An important advantage of longitudinal studies is that they might eventually lead to early interventions for adolescents at high-risk for excessive alcohol use.

Conclusion

In conclusion, this is the first large-scale study in adolescence on the interplay between pubertal hormones, intrinsic brain connectivity and alcohol use as a measure of risk taking behavior. We found evidence for an indirect effect of testosterone on alcohol consumption, through reduced amygdala-OFC connectivity in boys. This provides important insights into the mechanisms behind alcohol consumption, and may contribute to the development of prevention work aimed at reducing the chance of the transition from normative into abnormal forms of risk taking.

Chapter 9

Amygdala-orbitofrontal connectivity predicts alcohol use two years later

This chapter is based on:

Peters, S., Peper, J.S., Van Duijvenvoorde, A.C.K., Braams, B.R. & Crone, E.A. Amygdala-orbitofrontal connectivity predicts alcohol use two years later: A longitudinal neuroimaging study on alcohol use in adolescence (in revision, 2015).

Abstract

This study tested the relation between cortical-subcortical functional connectivity and alcohol consumption in adolescents using an accelerated longitudinal design. Participants (N = 299 at T1 and N = 254 at T2) between ages 8 and 27 completed resting state neuroimaging scans at two time points separated by two years. In addition, participants between ages 12 and 27 reported on recent and lifetime alcohol use. Resting state connectivity analyses focused on amygdala-orbitofrontal connectivity given prior research linking reduced coupling between these regions to alcohol use. The results indicated that amygdala-orbitofrontal connectivity at the first time point predicted alcohol use two years later. There was no evidence for the reversed relation, suggesting that brain connectivity measures precede explorative risk taking behavior in adolescence, possibly because decreased subcortical-frontal connectivity biases towards more explorative or risky behavior.

Introduction

Adolescence is a developmental period that is associated with increased risk taking behavior (Steinberg, 2008). One of the most prevalent forms of risk taking in adolescence is alcohol consumption (Hibell et al., 2012). There is considerable evidence that alcohol use increases sharply in adolescence and has negative consequences for cognitive functioning and school performance (Zeigler et al., 2005). Despite the presumed relations between alcohol use and brain development (Peeters et al., 2015), surprisingly little is known about how longitudinal changes in alcohol use in normally developing adolescents are related to changes in brain function over time. The current study addressed this question with an assessment at two time points for alcohol use and brain function in an accelerated longitudinal design with participants between 8-27 years old.

A well-suited approach to address this question is by using resting state analyses to measure changes in brain connectivity over time. This technique involves measuring connectivity between brain regions at rest, i.e. during the absence of a specific task, which makes it suitable for testing longitudinal questions as performance differences can be excluded as a confounding factor (Dosenbach et al., 2010). Here we focused on connectivity between subcortical and cortical systems. It has been hypothesized that during adolescence there is an imbalance between the relative maturity of subcortical brain regions (including the amygdala and ventral striatum), and prefrontal cortex regions that exert control over subcortical brain regions, possibly explaining the increased incidence of risk taking in adolescence (Ernst et al., 2006; Somerville & Casey, 2010). We recently demonstrated that decreased connectivity between the amygdala and orbitofrontal cortex (OFC) was related to increased alcohol use in adolescents (Peters, Jolles, Van Duijvenvoorde, Crone, & Peper, 2015). This effect was modulated by testosterone levels: higher testosterone production was related to lower brain connectivity and increased alcohol use. These data support the hypothesis that the amygdala-OFC brain network is shaped by pubertal hormones and is related to risk taking behavior as measured by consumption of alcohol. These results are also consistent with task-based fMRI studies in adults, which show a crucial role for the amygdala in alcohol use. For instance, an attenuated amygdala response to emotional faces has been demonstrated after alcohol ingestion (Gilman et al., 2012, 2008; Sripada et al., 2011) and reduced coupling between the amygdala and the OFC during an emotional face processing task after alcohol ingestion (Gorka et al., 2013). Animal studies have also shown an important role for the amygdala in the context of alcohol use in multiple ways, such as by mediating the locomotor stimulating effects of alcohol, and the finding that receptors in the amygdala appear to contribute to regulation of alcohol use (for a review see McBride, 2002).

However, relatively little is known about the direction of the longitudinal relationship between brain connectivity and alcohol consumption. That is, it is unclear whether alcohol use affects subsequent brain development, or whether aberrant brain connectivity precedes an individual's propensity to alcohol use. Support for the hypothesis that alcohol influences subsequent

brain development in adolescence comes from numerous animal studies and neuroimaging studies in human participants, which showed that substance use is linked to abnormalities in white matter, gray matter volume and abnormal activation during cognitive tasks (for a review see Squeglia, Jacobus, & Tapert, 2009). On the other hand, it is also possible that aberrant connectivity between subcortical and cortical areas biases adolescents towards risk taking behavior. In this study, we investigated the directionality of the relationship between alcohol use and amygdala-OFC connectivity, by using a longitudinal approach. We examined a large sample of adolescent participants between 8-27 years old who underwent resting state MRI scanning, and who filled out questionnaires on recent and lifetime alcohol use at two time points with a two-year interval. This large-scale longitudinal sample allowed us to elucidate whether changes in functional connectivity between amygdala and OFC precede or follow from alcohol use at the first time point.

Methods

Participants

This study was part of a larger project on cognitive and affective development (e.g., Braams, van Duijvenvoorde, Peper, & Crone, 2015; Peper, Koolschijn, et al., 2013; Peters et al., 2014). Participants (8-25 years old at the first time point (T1)) were recruited through local schools and advertisements (N = 299). Ages were between 8.01 and 25.95 at T1 (M = 14.06, SD = 3.61). Participants were recruited from different schools in the Netherlands to ensure that the sample reflected the general population. IQ was estimated with two subtests of the WAIS-III or WISC-III (Similarities and Block Design). IQ ranged between 80 and 143 (M = 109.72, SD = 10.52).

The follow-up measurement (time point 2 (T2)) was approximately two years later (Mean time between T1 and T2: 2.01 years, SD = 0.20) (N = 254). Ages were between 10.02 and 26.62 at T2 (M = 15.90, SD = 3.50). IQ was estimated again using the WAIS-III and WISC-III subtests Picture Completion and Vocabulary, and at T2 ranged between 80 and 147.50 (M = 108.28, SD = 10.34).

At both time points, adults (18 years and older) received payment (60 euros) for participation, and children received presents and their parents received 30 euros (for 12-17 year old children) or 25 euros (for 8-11 year old children) for travel reimbursement. The study was approved by the Institutional Review Board at the Leiden University Medical Center. The participants (or in case of minors, participant's parents) signed a written informed consent. All anatomical MRI scans were reviewed and cleared by a radiologist. None of the participants reported neurological or psychiatric disorders or current use of psychotropic medication at T1.

Complete MRI data at T1 was collected for 295 participants (4 of the 299 participants did not complete the MRI scan), but there was data of sufficient quality for 274 participants. Reasons for exclusion were: > 2 mm movement on the fMRI scan (n = 11), > 10 % of volumes affected by

micromovements (see criteria in the fMRI analysis section) (n = 14), a psychiatric diagnosis disclosed after participation (n = 1), and insufficient quality data (n = 2). At T2, 13 of the 299 initial participants could not or did not want to participate a second time. At T2, a further 32 participants could not participate in the MRI session due to braces, resulting in complete MRI data at T2 for 254 participants. There was sufficient quality data for 231 participants for resting state fMRI (exclusions: movement > 2 mm: n = 5; > 10 % of volumes affected by micromovements: n = 9).

The alcohol questionnaire was only administered to participants who were 12 years or older. This resulted in 193 participants at T1 and 244 participants at T2. All analyses were conducted in a pairwise manner, i.e. using all available data for each particular analysis. See Table 1 for an overview of the number of participants in each analysis.

*Table 1: overview of the number of participants for each variable. MRI data were collected for all participants who took part in the study. *Alcohol self-report data were only collected in case participants were 12 years or older.*

	N		Age Range	
	T1	**T2**	**T1**	**T2**
Participation	299	286	8-25	10-27
MRI scan of sufficient quality	274	231	8-25	10-27
Alcohol data*	193	244	12-25	12-27

Alcohol Questionnaire

Participants filled out an on-line questionnaire at home on recent and lifetime alcohol use (Ames et al., 2007; de Water et al., 2013; Peters et al., 2015; Thush et al., 2008). The instructions explicitly stated that participant's answers were confidential and would not be disclosed to anyone. Participants were instructed to fill out the questionnaire at a time as close as possible to the MRI scan. Lifetime alcohol use was reported as the lifetime amount of glasses consumed on an 11-point scale (0, 1–10, 11–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80, 81–90, and > 90). Bottles and cans were counted as 1.5 glasses, because these contain more of the beverage than a standard glass in the Netherlands (Thush et al., 2008). Recent alcohol use was reported as the number of glasses of alcohol participants had consumed over the past 30 days on a 10-point scale (0, 1–2, 3–4, 5–6, 7–10, 11–15, 16–20, 21–30, 31–50, and > 50). To create a scale variable, the ordinal data on quantity of alcohol use were converted by calculating the mean of the answer (for > 50 and > 90, 51 and 91 were used, respectively). On average, participants had consumed 28.65 glasses of alcohol in their lives (SD = 37.68) and 6.35 glasses in the last month (SD = 12.36), at T1, and had consumed 36.00 glasses in their lives at T2 (SD = 39.21) and 9.25 in the past month (SD = 14.48).

MRI data Acquisition

Scans were acquired with a Philips 3T MRI scanner. The same scanner and settings were used at T1 and T2. Functional scans were acquired with T2*-weighted echo-planar imaging (EPI). The first two volumes were discarded to allow for equilibration of T1 saturation effects. The following scan parameters were used: 140 volumes; 38 slices; sequential acquisition; TR = 2200 ms, TE = 30 ms; flip angle = 80°; FOV = 220x220x114.67 mm; slice thickness = 2.75 mm. A high-resolution anatomical scan (T1-weighted; 140 slices; TR = 9.76 ms; TE = 4.59 ms; flip angle = 8°; FOV = 224×177.33×168 mm; in-plane resolution = 0.875x0.875 mm; slice thickness = 2 mm) and a high-resolution T2*-weighted gradient echo EPI scan (84 slices; TR = 2200 ms; TE = 30 ms; flip angle = 80°; FOV = 220x220x168 mm; in-plane resolution = 1.96x1.96; slice thickness = 2 mm) were acquired after the resting state scan. Participants were instructed to close their eyes during the resting state scan. Before the MRI scan, participants were accustomed to the MRI environment and sounds with a mock scanner.

FMRI data preprocessing

FMRI preprocessing was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.98, part of FSL (www.fmrib.ox.ac.uk/fsl). These steps were used: motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002); non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 5mm; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering of 100 s (Gaussian-weighted least-squares straight line fitting, with sigma = 50 s). The resting state scan was registered with FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001) to the high resolution T2*-weighted scan, which was registered to the T1-weighted scan, and the T1-weighted scan was registered to the 2 mm MNI-152 standard image.

FMRI data analysis

In keeping with the prior cross-sectional study (Peters et al., 2015), left and right amygdala were selected for a seed-based correlation approach (Fox & Raichle, 2007) to test for functional connectivity with the OFC. Amygdala masks were obtained using atlas-based masks of the amygdala (Automatic Anatomical Labeling; see Figure 1). Amygdala masks in MNI-space were transformed to native space (each individual's resting state scan) with a binary threshold of 0.5. Next, mean time courses were extracted from each individual's amygdala, i.e. all voxels located within the amygdala mask. These mean time courses were entered as regressors in a GLM (separately for left and right amygdala), with nuisance regressors for white matter and CSF signal (obtained from a bilateral 4 mm sphere in white matter (left: x = 54, y = 44, z = 44; right x = 35, y = 44, z = 44) and CSF (left: x = 59, y = 55, z = 50; right: x = 30, y = 55, z = 50), global signal, and six motion parameters (rigid body: three translations and three rotations). For participants with excessive micromovements (> .05 mm) between volumes, we included additional regressors (binary for all

volumes with movement > .05) to remove specific volumes where micromovements occurred from the analysis. Participants where more than 10 % of volumes were affected by micromovements (> .05 mm) were excluded from further analyses.

Statistical analyses

We issued a region-of-interest (ROI) approach to specifically investigate amygdala-OFC connectivity using an OFC anatomical mask (based on AAL: Medial Orbital Frontal Gyrus) with left and right OFC combined. OFC masks were transformed to native space with a binary threshold of 0.5. Next, we extracted Z-scores for amygdala connectivity with the OFC. To confirm that the amygdala and OFC were functionally connected, whole-brain analyses were performed for visual inspection. Left and right amygdala showed positive functional connectivity with the OFC at both T1 and T2 (Figure 1). The ROI results were further analyzed with SPSS 19 and R 3.1.1.

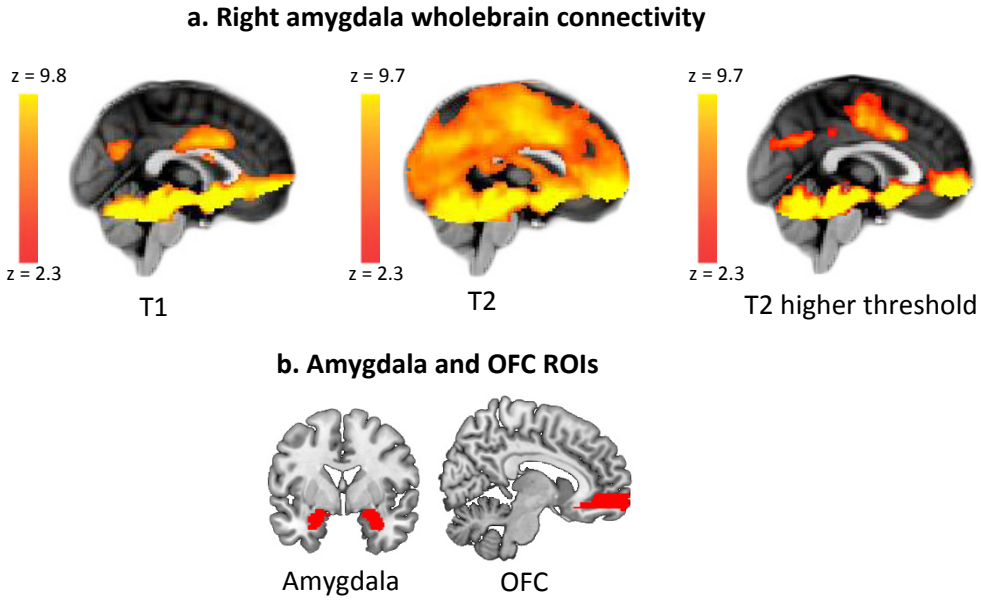


Figure 1a: Positive whole-brain connectivity with the right amygdala as seed (cluster-thresholded at 2.3, p < .05). The threshold at T2 was manually set to intensity 7 out of 9.62 for visual inspection. Figure 1b: Amygdala and orbitofrontal cortex anatomical ROIs

Prediction analyses for alcohol use and brain connectivity

Correlation analyses were performed to examine whether there was consistency between T1 and T2 for alcohol use and amygdala-OFC connectivity. Next, prediction analyses were performed to examine the direction of the relationship between alcohol use and amygdala-OFC connectivity. We performed regressions with alcohol use (recent and lifetime in separate analyses) at T2 as dependent variable, age at T1 as first predictor and amygdala-OFC connectivity at T1 (left and right amygdala-OFC connectivity in separate analyses) as second step. In addition, we tested for the reverse direction, with amygdala-OFC connectivity at T2 as dependent variable, age at T1 as first predictor and alcohol at T1 as second predictor. These analyses were also performed with baseline alcohol use/amygdala-OFC connectivity at T1 entered as additional (control) step.

Age effects: mixed model analyses

As a final goal, we assessed how all measures changed as a function of age. To model developmental trajectories (linear, quadratic or cubic shapes) for alcohol use and brain connectivity, we used mixed model analyses (Braams et al., 2015; Ordaz et al., 2013). We tested a linear effect of age (i.e., monotonic development), a quadratic (i.e., adolescent-specific effect) and a cubic effect (i.e., adolescent-emergent pattern). These analyses are a more advanced version of multiple regression, but take the longitudinal nature of the data into account. That is, both absolute (i.e., the intercept) and change values for each individual were analyzed, and it was not necessary to calculate change scores. The analyses were performed with the NLME package in R (Pinheiro et al., 2007). Models were compared using the Akaike Information Criterion (AIC) with lower values indicating a better model fit. We additionally tested with log-likelihood-tests whether changes in AIC model fit were significant. These model-building steps were used: First, we tested for each variable (left and right amygdala-OFC connectivity, recent and lifetime alcohol use; at two time points) which pattern best described the developmental trajectory. The base model consisted of a fixed and a random intercept, describing variation in starting points (intercepts) of individuals. Next, we tested with polynomials (Braams et al., 2015) whether a model with age as a linear effect resulted in a better fit compared to the base model without age. Then, a model including a linear and quadratic term for age was compared to the linear model, and finally, we tested if a combined linear, quadratic and cubic model predicted the data better than a combined linear and quadratic model. For the best age model, we tested whether age as an effect with a random slope resulted in a better fit, which would indicate that the age effect differs for each individual. We did not find evidence for significant random slopes and did not report this further in the results section.

Results

The results section is organized along the following lines: First, consistency for all measures between T1 and T2 was calculated. Next, prediction analyses were performed to examine the directionality of the relationship between alcohol use and amygdala-OFC connectivity. As a last step, we assessed developmental trajectories for alcohol use and amygdala-OFC connectivity.

Consistency between T1 and T2

Correlation analyses showed that for alcohol use, both recent (r = .71, p < .001) and lifetime alcohol use (r = .79, p < .001) were highly correlated between T1 and T2. In addition, amygdala-OFC connectivity was modestly correlated over time, for both left (r = .14, p = .042) and right amygdala (r = .15, p = .023).

Prediction analyses for amygdala-OFC connectivity and alcohol use: direction of the effect

The next set of analyses addressed the question whether current alcohol use can be predicted from amygdala-OFC connectivity at an earlier time point, or whether current amygdala-OFC connectivity can be predicted from earlier alcohol usage. In the analyses reported below, we corrected for age differences in alcohol use.

First, we investigated whether amygdala-OFC connectivity at T1 predicted alcohol use at T2. A hierarchical regression with alcohol use at T2 as dependent variable, age as first predictor and amygdala-OFC connectivity at T1 as second predictor, showed a significant effect of left amygdala-OFC connectivity on alcohol use two years later, for both lifetime (β = -.13, p = .002) and recent alcohol use (β = -.10, p = .042). That is, less positive connectivity at T1 was associated with increased alcohol use at T2. The relation between left amygdala-OFC connectivity at T1 and lifetime alcohol use at T2 remained significant when adding lifetime alcohol use at T1 as a second predictor above age (β = -.10, p = .024). These analyses showed that less positive connectivity between the amygdala and the OFC predicts alcohol use two years later, and that amygdala-OFC connectivity explains lifetime alcohol use two years later even when controlling for baseline alcohol use at T1.

To test for the reversed direction, we investigated whether alcohol use at T1 predicted amygdala-OFC connectivity at T2, but no significant results were found. Together, these analyses suggest that brain connectivity precedes alcohol use, but we found no evidence for the reverse direction, i.e. alcohol use preceding brain connectivity.

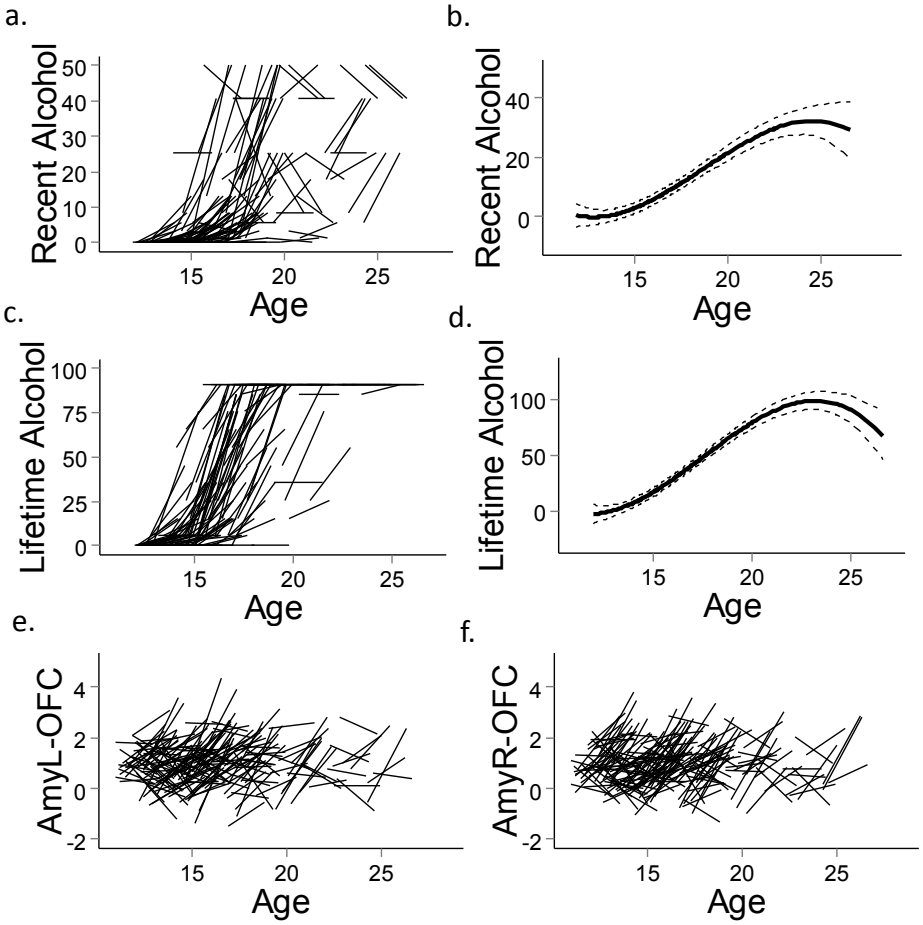


Figure 2: Predicted values (2b, 2d) and raw data (2a, 2c) for the cubic relationship between recent and lifetime alcohol use and age. Figure 2e and 2f depict raw data for left and right amygdala-OFC connectivity, which revealed no age effect.

Age effects on alcohol use and amygdala-OFC connectivity

We additionally investigated how alcohol use and amygdala-OFC connectivity changed as a function of age. Mixed models were used to test the longitudinal pattern of development (linear, quadratic or cubic). These analyses revealed that both lifetime and recent alcohol use were best described by cubic patterns for age (i.e., rising quickly in mid adolescence and leveling off in early adulthood, Figure 2; Table 2). For recent alcohol use, a combined linear and cubic pattern best

described the data, whereas for lifetime alcohol use, the best fitting function was a combination of a linear, quadratic and cubic function.
For amygdala-OFC connectivity, mixed linear modeling revealed that a model without age was the best fit to the data, suggesting no significant age-related change over time (Table 2).

Table 2: AIC and loglikelihood p-values for a base model (without age), linear, quadratic and cubic age pattern. The best-fitting model is highlighted in bold font.

	Base	**Linear**		**Quadratic**		**Cubic**	
	AIC	**AIC**	***p***	**AIC**	***p***	**AIC**	***p***
Lifetime alcohol	4321	3972	<.001	3956	<.001	3920	<.001
Recent alcohol	3435	3252	<.001	3254	.917	3235	<.001
Left amygdala-OFC	1348	1348	.273	1351	.870	1353	.727
Right amygdala-OFC	1384	1386	.615	1386	.173	1388	.396

Discussion

In this study, our goal was to investigate the longitudinal relationship between alcohol use and amygdala-OFC connectivity. In particular, our aim was to investigate whether amygdala-OFC connectivity could be predicted from earlier alcohol use, or instead, whether alcohol use could be predicted from amygdala-OFC connectivity two years earlier. The results indicated that amygdala-OFC connectivity at the first time point predicted alcohol use two years later, but there was no evidence for the reverse direction. The results are described in more detail in the following sections.

Longitudinal relationship between amygdala-OFC connectivity and alcohol use
In our prior study based on cross-sectional comparisons we reported a correlation between reduced amygdala-OFC connectivity and increased alcohol use (Peters et al., 2015). Our main goal in the current study was to investigate the directionality of the relationship between amygdala-OFC connectivity and alcohol consumption using longitudinal data on two time points. We tested whether reduced amygdala-OFC connectivity preceded alcohol use (suggesting vulnerability to alcohol use due to reduced coupling of prefrontal and subcortical brain systems), or whether increased alcohol use preceded reduced amygdala-OFC connectivity (suggesting a 'damaging' effect of alcohol use on amygdala-OFC connectivity). The results indicated that amygdala-OFC connectivity preceded alcohol use two years later, but we found no evidence for the reverse direction. This effect was found for both lifetime and recent alcohol consumption, and was specific for

left amygdala-OFC connectivity. Importantly, the prediction of lifetime alcohol use from left-amygdala OFC connectivity remained significant when controlling for alcohol use at the first time point, suggesting that brain connectivity explains unique variance in future alcohol use over and beyond behavioral assessments.

These findings are in line with the idea that subcortical-prefrontal connectivity is important for top-down control over behavioral approach tendencies. For instance, prior studies showed that increased connectivity between the amygdala and the OFC was associated with improved emotion regulation and behavioral control (Banks et al., 2007; Lee et al., 2012). This suggests that increased connectivity is protective against risk taking, which fits with the current findings that decreased amygdala-OFC connectivity predicts increased alcohol use.

No evidence was found for the reverse direction, i.e. alcohol use preceding reduced connectivity between the amygdala and the OFC. Although prior studies reported that alcohol consumption can affect brain structure and function (Squeglia et al., 2009), this is the first longitudinal study specifically investigating amygdala-OFC connectivity during resting state. Our findings suggest that, with regard to the specific connectivity between the amygdala and the OFC, increased alcohol use does not affect coupling between these regions. This highlights the importance of longitudinal designs to determine the direction of a cross-sectional association between brain connectivity and behavior.

Stability and change of alcohol use and amygdala-OFC connectivity over a two-year period

In addition to these main analyses, we assessed the level of stability and age-related changes in alcohol use and amygdala-OFC connectivity within a two-year period. All measures showed significant correlations between T1 and T2, confirming that they are valid indices of individual variation. Alcohol use showed relatively high stability over time. The correlation of amygdala-OFC connectivity over two time points was modest but significant. It should be noted that a limitation of this study was the relatively short assessment time for resting state analyses. That is, prior studies have argued that resting state connectivity is a reliable measure of brain function, but this appears to be mostly the case for scans of relatively long duration (i.e. > 9-12 minutes), compared to our acquisition time (6 minutes) (Birn et al., 2013). Nonetheless, the study resulted in consistent patterns over time.

Next to this substantial level of individual stability, we investigated whether alcohol use and amygdala-OFC connectivity showed age-related changes during adolescence. Consistent with prior studies, we observed a strong increase in alcohol use with increasing age (Hibell et al., 2012). With mixed model analyses for longitudinal data, we assessed the shape of developmental trajectories for alcohol use (linear, quadratic or cubic patterns). These analyses indicated that the developmental trajectory for alcohol use was best described by a cubic effect of age. That is, alcohol use was relatively stable in young adolescents, then showed a steep increase in mid-adolescence, and leveled off again towards young adulthood. These cubic age-effects were found

for both lifetime consumption and recent alcohol use (over the past month). It should be noted that the index of lifetime alcohol use reached a ceiling effect (i.e., the maximum amount of glasses that could be chosen in the questionnaire was '91 or more glasses') which makes the last phase less reliable, but the same pattern was found for recent alcohol use (and see also Chassin, Pitts, & Prost, 2002; White, Xie, Thompson, Loeber, & Stouthamer-Loeber, 2001).

With regard to developmental patterns in amygdala-OFC connectivity, we found no linear, quadratic or cubic effect of age using longitudinal mixed models on amygdala-OFC connectivity. These results do not concur with an earlier cross-sectional study in a smaller-scale task-based study (Gee et al., 2013), who reported a shift from positive to negative connectivity with increasing age, and a prior cross-sectional resting state study (Gabard-Durnam et al., 2014), which reported an age-related increase in connectivity, suggesting that cross-sectional and longitudinal studies, as well as task-based vs. resting state studies may reveal different findings when studying connectivity during adolescent development. Future studies should investigate age-related changes in amygdala-prefrontal connectivity in more detail, with more optimized acquisition times (Birn et al., 2013). The current results suggest that amygdala-OFC connectivity may be a developmental marker that is predictive for future explorative or risk taking behavior.

Limitations and future directions

There are several limitations to this study that should be taken into account. First, although our large-scale longitudinal data could be used to find support for the direction of the relation between alcohol use and amygdala-OFC connectivity, such studies in human participants still cannot provide true causal evidence. Individuals who consume relatively large amounts of alcohol may differ from peers who consume less alcohol in other aspects which could not be controlled for in this study. Second, the alcohol measures in this study were based on self-report, which may lead to overestimations or underestimations of actual alcohol consumption. However, prior studies showed that self-report measures of alcohol can be reliable if confidentiality of answers is ensured (Brener et al., 2002; Sobell & Sobell, 1990).

Conclusion

In conclusion, this large-scale longitudinal study provided evidence that future alcohol use can be predicted from amygdala-OFC connectivity. These results have important implications for understanding the onset and progression of alcohol use in particular, and more generally, the link between subcortical-cortical connectivity and risk taking behavior in adolescence. Possibly, relatively reduced subcortical-cortical connectivity in early to mid-adolescence creates a vulnerable window for starting alcohol use (Ernst et al., 2006; Somerville & Casey, 2010). Eventually, these results may inform early interventions aimed at adolescents with relatively more sensitivity to exploration and risk taking.

Chapter 10

Summary and discussion

Summary and discussion

Adolescence is an important developmental period that is often characterized as a period of slowly emerging self-control and increased propensity to risk taking behaviors (Steinberg, 2008). It is important to investigate the mechanisms at play during adolescence in order to understand how cognitive control develops and how reckless behavior can be prevented. Several authors have constructed brain-based models to explain the phenomena of steadily increasing cognitive control from childhood to adulthood and the adolescent peak in risk taking. The most influential of these models, known as e.g. imbalance models (e.g. Ernst et al., 2006; Somerville & Casey, 2010; Steinberg, 2008) theorized that increased risk taking in adolescence can be explained in terms of relative immaturity of frontoparietal control regions in the brain, combined with a heightened sensitivity of affective brain systems in adolescence.

Although these imbalance models have sparked an increasing amount of research attention focusing on brain development in adolescence, so far relatively few large-scale and longitudinal studies have been performed. Cross-sectional studies provide a proxy of development, but no information about true developmental trajectories. Possibly as a result of this, several inconsistencies have arisen in the literature regarding development of neural activity, such as studies reporting both age-related increases as well as age-related decreases in frontal control regions (Crone & Dahl, 2012). In this thesis, I reported data from a large cross-sectional and longitudinal sample of children, adolescents and adults ranging between 8 and 27 years old. The main questions addressed in this thesis were 1) how brain regions for cognitive control develop during adolescence, and 2) how connections between affective and cognitive brain regions influence the propensity to engage in risk taking behaviors.

Development of cognitive control

In the first part of this thesis (Chapters 2-7), I examined the development of brain regions important for cognitive control using both cross-sectional and longitudinal samples. In Chapter 2, an extensive literature review of the development of cognitive flexibility was described. Cognitive flexibility is posited as one of the three main cognitive functions aside from working memory and inhibition (Diamond, 2013; Miyake et al., 2000), and refers to the ability to flexibly switch between different behaviors and the ability to adapt to a continually changing environment. Conceptually, experimental paradigms measuring cognitive flexibility can be divided into instructed flexibility and adaptive flexibility. In instructed flexibility paradigms, participants are explicitly instructed to switch behaviors (e.g., 'now only press the button when you see a green instead of a red figure'). In paradigms for adaptive flexibility, participants are not explicitly asked to switch response patterns, but instead have to infer the new correct answer based on negative feedback for a response that previously resulted in positive feedback. The studies described in this review showed that performance on both types of flexibility improved with age. However, there were differences

between the two types when examining patterns of neural activity. For instructed flexibility, there appeared to be a linear increase in recruitment of brain regions for cognitive control with age. For adaptive flexibility, the pattern was more complex: regions for cognitive control became increasingly specialized in the processing of feedback with informative value. That is, brain regions for cognitive control did show activity in response to feedback in children, but these regions did not distinguish as clearly between informative and uninformative forms of feedback compared to the pattern observed in adults.

Brain mechanisms underlying feedback learning

In Chapter 3, I examined learning from positive and negative feedback in a sample of 32 adult participants (18-25 years). The goal of this study was twofold: first, to validate a paradigm to investigate neural activity in frontoparietal control regions; second, to investigate whether regions in the frontoparietal network were mostly activated after negative feedback (as reported in several earlier studies), or whether instead, these regions are better characterized as being sensitive to informative value of feedback. To test this, participants performed a feedback learning task in a 3T MRI scanner where they viewed three empty boxes with a picture presented underneath. Participants were instructed to determine which picture (out of three possible options) belonged in which of the three boxes, which could be inferred from positive and negative performance feedback. To analyze neural sensitivity to valence and/or informative value of feedback, a distinction was made between the learning phase and the application phase of the experiment. We defined the learning phase as the trials where the correct location was not yet known; therefore feedback during this phase was still informative for learning. The application phase started when participants were simply repeating correct responses they already provided before, and feedback was therefore no longer informative for learning. The neuroimaging results indicated that all regions of interest in the frontoparietal network (pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC), dorsolateral prefrontal cortex (DLPFC) and superior parietal cortex (SPC)) were sensitive to both informative value and valence. That is, all regions were more active after informative compared to uninformative feedback, and more active after negative compared to positive feedback (possibly indicating an increased need for adaptation after receiving negative feedback). The degree to which neural regions were sensitive to informative value (i.e., the difference in activation after informative compared to uninformative feedback) correlated with behavioral performance. In addition, we used computational modeling analyses to distinguish groups based on differences in strategy use. We found that participants who employed more efficient strategies that enabled faster learning showed more activity in the frontoparietal network.

Development of the frontoparietal network in adolescence

In Chapter 4, I described a large-scale cross-sectional study in 268 participants between 8 and 25 years old, who performed the same feedback learning paradigm as described in Chapter 3. The

main goal was to investigate how activity in the frontoparietal network changes across development, and importantly, whether children were also able to recruit these regions (which would argue against a strict interpretation of the frontal immaturity hypothesis in imbalance models). Several findings emerged from this study. First, I demonstrated that young children could in fact recruit regions in the frontoparietal network, albeit in different circumstances than adults. That is, activity in the frontoparietal network after receiving negative feedback increased with advancing age, as would be expected from imbalance models (which hypothesize that control regions become increasingly 'on-line' with development). However, children actually showed more activity than adults in parietal areas when they received positive feedback, supporting the notion that children can effectively recruit these areas. Second, I showed that neural reactions for negative feedback increased with age, whereas neural activity for positive feedback remained constant between ages 8 and 25. In addition, I discovered that both behavioral performance and neural activity reached adult levels around age 14-15 years, which was consistent with prior literature but was not yet confirmed in a study with a large number of participants per age group. Finally, I showed that the pubertal hormones testosterone and estradiol as well as pubertal stage did not explain additional variance above age, suggesting that areas in the cognitive control network develop relatively independent from pubertal influences.

In Chapter 5, I analyzed the behavioral data from the cross-sectional study using a different method. During the feedback learning task, participants could use reasoning strategies with different levels of complexity (e.g., a complex statistical reasoning strategy would be: 'the lion did not belong in the left box, so I am now trying this box first for the elephant', but it was also possible to complete the task by only relying on simple memory skills with no reasoning required). Markov modeling procedures were used to search whether underlying groups could be detected which differed in strategy use. The results showed that there were four different strategy groups, which differed in how advanced their reasoning strategies were. Importantly, the complexity of reasoning did not correspond directly to age, e.g., a substantial proportion of young children used more complex reasoning strategies than some of the adults. I also demonstrated that the more advanced strategy groups showed more activity in the frontoparietal network during the task. An important question I aimed to answer was whether more efficient strategy use was an explanation for the fact that older participants showed more activity in frontoparietal areas. Mediation analyses revealed that age effects on neural activity could partly be explained by strategy use, but a substantial portion of variance was still explained by age alone.

In Chapter 6, I investigated longitudinal aspects of development in frontoparietal activity. This chapter was based on data from the first time point described in Chapters 3-5, combined with data from a follow-up measurement two years later. In this study I assessed development of the frontoparietal network within individuals, rather than comparing individuals from different ages at one time point. I demonstrated that key regions in the frontoparietal network showed distinct developmental trajectories: Development of activity in the medial prefrontal cortex (pre-

SMA/ACC) was best characterized by a linear increase with age, but the SPC and DLPFC were best described by quadratic trajectories, showing a peak in adolescence or leveling off towards adulthood. An increase in behavioral performance explained additional variance above age alone for the DLPFC and SPC, whereas cortical thickness explained additional variance in pre-SMA/ACC.

Relevance for learning in real-world settings

In Chapter 7 I examined the important question whether laboratory-based learning tasks (such as the feedback learning task) which are often used to study the process of learning in tightly controlled settings, are actually predictive of learning in real-world settings such as in school. Reading and mathematics performance were used as proxies to assess real-world learning. I demonstrated that both performance and neural activity during the feedback learning task predicted reading and mathematics performance two years later. This effect remained significant even when correcting for age, working memory capacity and IQ, suggesting that feedback learning captures a unique aspect of variation in reading and mathematics performance. In addition, I showed that neural activity can predict additional variance in reading and mathematics above behavioral testing alone, highlighting the importance of neural measures in addition to behavioral measures when attempting to predict future school performance.

Risk taking in adolescence: connectivity between affective and cognitive brain regions

To get a more complete view on adolescent development I also examined affective aspects of development given that adolescent behavior is characterized by both increases in cognitive control as well as increased sensitivity to affective cues (Steinberg, 2008). To this end, I investigated whether connectivity between affective and cognitive regions is related to risk taking behavior, as would be predicted from imbalance models. In Chapter 8 and 9, I described two studies on the influence of connectivity between subcortical and cortical regions and testosterone on alcohol use in adolescents.

In Chapter 8, the goal was to investigate the relationship between alcohol use, functional brain connectivity and testosterone. To this end, resting state fMRI scans were acquired from the same participants described in Chapters 3-7. In order to collect these scans, participants were asked to lie as still as possible in the MRI scanner for five minutes with their eyes closed. Resting state functional MRI is a method for evaluating regional interactions that occur when a subject is not performing an explicit task. It is thought that individual differences in the strength of connectivity between regions in these resting state networks are related to behavior and personality traits. In this study, I specifically investigated connectivity between the amygdala and the orbitofrontal cortex. Prior studies have linked connectivity between these areas to both testosterone and alcohol use, but this was not yet investigated in a single sample nor in adolescents. In this study I showed that for boys, relatively high levels of testosterone were associated with decreased con-

nectivity between the amygdala and the OFC. This decreased connectivity in turn was associated with increased alcohol use. In Chapter 9, I further demonstrated using longitudinal data from two time points, that amygdala-OFC connectivity at baseline predicted alcohol use two years later. Interestingly, I found no evidence for the reverse direction, i.e. alcohol use predicting later brain connectivity. This study highlighted the importance of longitudinal studies, as well as providing support for the notion that aberrant subcortical-cortical connectivity can increase the propensity for risk taking behaviors.

Discussion and future directions

Nonlinear development of the frontoparietal network

Taken together, the studies described in this thesis revealed several important findings. First, I demonstrated that recruitment of the frontoparietal network cannot be explained by a simple linear increase with age. For instance, although regions in the frontoparietal network did show an age-related increase in activity after receiving negative feedback, a number of these regions were actually more activated in children than in adults after receiving positive feedback (Chapter 4). Furthermore, the results from a longitudinal study presented in this thesis (Chapter 6) revealed that frontoparietal activity increases in a nonlinear way with age in the DLPFC and SPC. Most prior studies used cross-sectional comparisons, relatively small sample sizes and/or discrete age groups (e.g. 8-12, 13-17 years) rather than assessing the whole age range from late childhood to early adulthood, which possibly explains contradictory findings in the literature regarding neurocognitive development. The work described in this thesis demonstrates the benefits and importance of using large sample sizes and longitudinal designs to effectively determine neurodevelopmental trajectories.

The findings of nonlinear development with age are not entirely consistent with simple models of frontal immaturity as predicted by imbalance models. The results from Chapter 4 and 6 showed that frontoparietal regions are not simply 'offline' or underdeveloped in childhood, but instead that they are involved during different processes compared to adults. In Chapter 6, I found support for a quadratic pattern with an adolescent peak or leveling off in late adolescence in behavioral performance and recruitment of SPC and DLPFC. A similar peak in behavioral performance for 17-year olds during executive functions tasks was reported previously (Taylor, Barker, Heavey, & McHale, 2013). These results fit better with the model of adolescent brain development posited by Crone and Dahl (2012), who hypothesized that adolescence can be characterized as a period of increased flexibility for recruitment of cognitive control regions. That is, the framework predicts that the range of possible behavioral and neural outcomes may be larger in adolescents than in adults depending on motivation or salience. For instance, in a 'hot context' when motivation is high (e.g. when peers are present or when monetary incentives are offered), performance and neural activity may be relatively more enhanced in adolescents than in adults.

Although most prior research has focused on peer presence effects on risk taking behavior (e.g. Chein, Albert, O'Brien, Uckert, & Steinberg, 2011), it has been shown that the presence of peers also influences performance on cognitive tasks such as relational reasoning (Wolf, Bazargani, Kilford, Dumontheil, & Blakemore, 2015). Also, recent research showed that adolescents recruit the frontoparietal network to a greater extent than adults for decisions associated with a greater monetary reward (Teslovich et al., 2014) and during relatively complex paradigms such as creative divergent thinking (Kleibeuker et al., 2013). However, relatively few studies have explored increased flexibility of adolescent cognition, thus further research is necessary to confirm this intriguing hypothesis.

Unraveling the mechanisms underlying cognitive development is especially important given the potential relevance for learning in school settings. In Chapter 7, I demonstrated that both behavioral performance and neural activity during a feedback learning task were predictive of reading and mathematics performance two years later. An interesting discussion that is relevant for school interventions is whether children learn better from positive compared to negative feedback. In Chapter 4, I found no evidence at the behavioral level that children learn relatively more from positive than from negative feedback: Children learned less overall from feedback compared to adults, but there was no developmental difference in the ratio of learning from positive compared to negative feedback (but see van Duijvenvoorde et al., 2008). However, the finding that frontoparietal activity for negative feedback increased with age, whereas activity for positive feedback remained constant with age, may indicate that processing negative feedback requires more effort as indicated by increased recruitment of cognitive control regions by adults. Negative feedback may be inherently more difficult than positive feedback, because an additional inference has to be made which demands more from the cognitive system (e.g. if response A received negative feedback, the child must infer that a switch to another answer is necessary). This particular feedback learning paradigm was relatively easy, but when the frontoparietal system is increasingly taxed, such as when multiple tasks are performed at the same time, a developmental difference in the relative learning rate from positive and negative feedback may become apparent. Encouraging teachers to focus more on positive rather than negative feedback may improve children's performance because positive feedback demands less from the developing cognitive system, but this needs to be confirmed in further research.

With regard to the adolescent period, I argue that it is important to focus not only on adolescent deficiencies in executive functioning, but also on the unique possibilities of the adolescent brain given the right incentives. New models of adolescent brain development (Casey, 2015; Crone & Dahl, 2012; Johnson, 2011) and preliminary evidence (Teslovich et al., 2014; Wolf et al., 2015) suggest that adolescents may benefit relatively more than other age groups from for instance social feedback from peers and affective rewards in relation to cognitive performance. Pursuing this research further may have important implications for educational practice and may lead to programs that are better suited to the specific challenges of the adolescent brain.

Individual differences influencing frontoparietal recruitment

Aside from age-related changes in recruitment of the frontoparietal network, I found evidence for other factors influencing the level of frontoparietal recruitment during a feedback learning task. This is an important line of research given that age-related changes interpreted in prior studies, may in fact be explained by individual differences in behavioral or neural measures rather than by development. In this thesis, I demonstrated that increased behavioral performance (defined as the percentage of feedback correctly used in further trials) corresponded to increased frontoparietal activation, even when correcting for age. The findings from Chapter 5 furthermore indicated that a portion of age-related increases in neural activity during a feedback learning task was mediated by differences in strategy use. That is, not only increased age, but also more efficient strategy use was related to increased recruitment of the frontoparietal network. Therefore, it was possible that young children who used relatively advanced reasoning strategies, showed more frontoparietal activity than adults who used less advanced strategies. Other individual differences I investigated in relation to cognitive development were cortical thickness (Chapter 4 and 6), working memory (Chapter 6), sex hormones and pubertal development (Chapter 4). I demonstrated that cortical thickness explained variance above age in pre-SMA/ACC activity, but working memory and sex hormones did not explain additional variance in frontoparietal recruitment.

It is important that new models of adolescent brain development theorize on the contribution of individual differences to developmental changes in neural activity. The individual differences I tested in thesis, i.e. task performance, working memory, cortical thickness and pubertal development, are so intricately linked with development that it is currently not certain whether they should be studied separately from age. It is an interesting theoretical discussion whether these individual differences should be regarded as a crucial part of development, or whether these are factors that should be corrected for so that pure age-related differences remain. Related to this issue, in Chapter 4 I found no relation between sex hormones and cognitive performance or frontoparietal activity. This is consistent with prevailing developmental theories which hypothesized that cognitive brain systems develop relatively independently from hormonal influences (Nelson et al., 2005; Steinberg, 2008). This may indicate that advancing age and pubertal development are separable processes that should be studied independently.

With regard to the relevance of studying individual differences for learning in school settings, it is intriguing that I found a relation between feedback learning performance and neural activity for feedback that is informative compared to feedback that is not informative for learning. Therefore it is possible that, the larger the difference between neural activity for learning vs. applying rules, the more sensitive these neural areas are to learning signals. This neural difference may be an interesting brain-based index of learning capacity that should be explored in further research, especially given the relation I found between this measure and future reading and mathematics performance (Chapter 7). In addition, my research provides preliminary reasons for

educators to focus more on teaching children different cognitive strategies, as the results in this thesis demonstrated that more advanced reasoning strategies were associated with increased frontoparietal recruitment. Prior research also suggested that development of feedback learning performance is better characterized by abrupt increases in performance due to a strategy switch, rather than slow, incremental changes (Schmittmann et al., 2012).

Adolescent risk taking and subcortical-cortical functional connectivity

In another line of research described in this thesis (Chapters 8 and 9), I investigated connectivity between cortical and subcortical regions and their relevance for risk taking behavior. I discovered that there is a relation between reduced amygdala-orbitofrontal cortex connectivity during resting state and increased alcohol intake. Longitudinal analyses showed that brain connectivity preceded alcohol use two years later, suggesting that aberrant brain connectivity influences an individual's future propensity to engage in risk taking behavior.

Although connectivity research is a relatively new topic of scientific interest, it is especially important to investigate connectivity during adolescence, given that imbalance models posit that risk taking can be explained in terms of an imbalance between emotional and cognitive brain systems, i.e. an inability of the cognitive system to 'put the brakes on' the affective system. This fits with the findings I reported in Chapter 8 and 9, which showed that decreased connectivity between the amygdala (seen as part of the affective brain system) and the medial frontal cortex was associated with increased alcohol consumption. Possibly, decreased connectivity corresponds to decreased (frontal) control over (subcortical) approach tendencies. On the other hand, increased connectivity between these regions may protect against risky decisions such as the choice to consume alcohol, for instance by increased top-down frontal control over more 'emotional' brain systems. Other research in line with imbalance accounts showed that with development, there is a decrease in the strength of local short-range connections (e.g. cortical-cortical, subcortical-subcortical connections) and an increase in more distal long-range connections (e.g. subcortical-cortical) in number and strength (Fair et al., 2009). However, with regard to the specific coupling between the amygdala and the OFC, it should be noted that I did not find evidence for age-related changes in connectivity using a large-scale longitudinal assessment over two time points.

Another interesting discussion surrounding imbalance models is whether risk taking should be seen as purely a societal problem, or also as a normal and perhaps even adaptive aspect of development. For alcohol specifically, it could be argued that alcohol use is in fact normative during adolescence, given the high prevalence during this developmental period (Hibell et al., 2012). It has also been shown that alcohol use is associated with popularity (Balsa, Homer, French, & Norton, 2011), indicating a socially rewarding and adaptive effect for alcohol use. For risk taking in general, several authors have argued that it is both a normal and adaptive characteristic of adolescence (Casey, 2015; Crone & Dahl, 2012). That is, it has been posited that a certain level of explorative behavior and increased independence from parents is both a necessary and beneficial

aspect of adolescence. Consistent with the idea that exploration/risk taking during adolescence is adaptive, is the fact that other animals also experience a juvenile period characterized by increased exploration, more fighting with parents and increased interest in same-aged peers (Casey, 2015).

Limitations and future directions for longitudinal research

Although the studies described in this thesis make an important contribution to the literature on adolescent brain development, several limitations need to be taken into account when interpreting these findings. Below I make several recommendations for future research which I believe will be crucial in order to advance our understanding of the developing brain.

The studies described in the first part of this thesis fit with recent models of adolescent brain development which emphasize increased flexibility of the frontoparietal control system during adolescence (Crone & Dahl, 2012; Johnson, 2011). Thus, adolescents may be more sensitive to manipulations of motivational context compared to children or adults. However, it should be noted that for the studies described in this paper, it was not possible to assess motivational salience nor was this explicitly manipulated with e.g. peer presence or monetary incentives. An interesting direction for future research would be to investigate influences of motivational salience on cognitive performance and frontoparietal recruitment in adolescents. Moreover, this hypothesis can also be tested by developing motivation-based interventions in schools that are tailored to the specific sensitivities of the adolescent brain.

In future studies, it will also be important to not only examine neural activity per se, but also connectivity between regions and its relation to cognitive development. For instance, Casey (2015) recently argued in favor of a circuit-based account of adolescent development, which takes into account connectivity between e.g. subcortical and cortical regions and more extensive brain networks. In Chapters 8 and 9 I investigated how neural connectivity between two brain regions is related to risk taking behavior. However, the relation to cognitive performance was not yet investigated in this thesis. It should also be noted that with the resting state methods employed in these chapters, it was not possible to determine whether activity in the OFC down-regulated amygdala activity or whether amygdala activity modulated OFC activity. Techniques such as dynamic causal modeling may unravel the direction of subcortical-cortical connectivity in relation to adolescent risk taking.

Another important step forward would be to assess participants at more than two time points, in order to pinpoint developmental trajectories in even more detail. This will allow for higher-order polynomial age functions to be tested within individuals, which is not possible with two time points. In addition, a broader selection of paradigms to measure cognitive functioning and risk taking behavior is needed in order to validate that the results presented in this thesis not only hold for this specific performance monitoring paradigm or to this particular questionnaire for alcohol consumption. With regard to hormonal measurements, it should be tested whether the

finding that cognitive development occurs independently from hormonal influences can be replicated. Other factors may also explain the lack of an effect, such as the timing of data collection for post-menarcheal girls (on the 7th day of the menstrual cycle rather than the day of the scan).

Taken together, future research should further test new models of adolescent brain development taking into account the complex environmental influences on adolescent behavior (Casey, 2015; Crone & Dahl, 2012) using similar large-scale longitudinal samples but a broader variety of measurements and more advanced techniques.

Conclusion

In conclusion, this thesis provides a comprehensive overview of both cognitive and affective aspects of development in relation to neural maturation and pubertal hormones. The results speak to a debate about imbalance models of adolescence, and provide evidence that a more nuanced description of development in frontoparietal control regions is needed. The results have important implications for constructing new theoretical frameworks and may eventually contribute to the advancement of educational interventions that are better tailored to the challenges and possibilities of the teenage brain.

Nederlandse Samenvatting

Hersenontwikkeling in de adolescentie

De adolescentie is een belangrijke periode in de ontwikkeling van kind tot volwassene. Niet alleen op lichamelijk vlak, maar ook in het gedrag treden er grote veranderingen op in de adolescentie. Zo zijn adolescenten meer geneigd tot risicovol gedrag, zoals alcohol- en drugsgebruik, roekeloos gedrag in het verkeer, gokken, gewelddadig gedrag en andere vormen van impulsiviteit (Steinberg, 2008). Een belangrijke bevinding uit eerder onderzoek is dat de cognitieve vermogens van kinderen, zoals zelfcontrole, het vermogen te redeneren en na te denken over de toekomst, zich nog tot ver in de adolescentie door ontwikkelen (Huizinga, Dolan, & van der Molen, 2006). De belangrijkste hersengebieden die bij deze processen betrokken zijn, zijn de prefrontale cortex en de parietale cortex, die deel uit maken van het frontoparietale netwerk (Niendam et al., 2012). Uit onderzoek naar de ontwikkeling van de structuur van het brein is gebleken dat de ontwikkeling van grijze stof (de hersencellen) nog doorgaat tot na het twintigste levensjaar (Giedd, 2004; Koolschijn & Crone, 2013). Belangrijk hierbij is dat de ontwikkeling niet even snel gaat in alle hersengebieden. De ontwikkeling gaat het traagst in de frontale en parietale gebieden (Giedd et al., 2009), juist de gebieden die belangrijk zijn voor cognitieve controle (Niendam et al., 2012).

Modellen voor hersenontwikkeling in de adolescentie

Verschillende onderzoekers hebben daarom modellen opgesteld om risicogedrag in de adolescentie te verklaren vanuit de hersenen (Ernst, Pine, & Hardin, 2006; Somerville & Casey, 2010; Steinberg, 2008). Deze modellen worden vaak 'imbalance' modellen genoemd, omdat ze uitgaan van een disbalans tussen twee belangrijke hersensystemen: de cognitieve gebieden en de affectieve gebieden. Binnen deze modellen wordt gedacht dat affectieve hersengebieden, dat wil zeggen de gebieden die betrokken zijn bij emotionele processen zoals gevoeligheid voor beloningen, op relatief jonge leeftijd volwassen zijn of zelfs een piek in activiteit laten zien in de adolescentie. Aan de andere kant wordt gedacht dat cognitieve gebieden juist relatief lang door ontwikkelen (Somerville & Casey, 2010). Dit leidde tot de hypothese dat de cognitieve hersengebieden in de adolescentie nog niet ver genoeg ontwikkeld zijn om de impulsieve neigingen vanuit de affectieve hersengebieden 'onder controle' te houden. Deze imbalance modellen hebben veel nieuwe onderzoeken geïnspireerd en hebben daarnaast een belangrijke populair-wetenschappelijke en maatschappelijke impact gehad.

Recent hebben verschillende auteurs echter beargumenteerd dat imbalance modellen een te gesimplificeerde weergave zijn van het complexe proces van hersenontwikkeling (Crone & Dahl, 2012; Pfeifer & Allen, 2012). Zo zijn er inderdaad veel studies die een toename in activiteit met leeftijd laten zien in de frontale en parietale cortex, zoals te verwachten op basis van imbalance modellen. Aan de andere kant zijn er ook studies die juist een afname in activiteit met leeftijd laten zien in deze gebieden (Crone & Dahl, 2012). Tot op heden is er nog geen consensus over wat er precies gebeurt in de puberhersenen, ondanks de vaak vergaande claims in populair-

wetenschappelijke publicaties. Dit komt mede doordat er weinig grootschalige onderzoeken onder grote groepen adolescenten zijn uitgevoerd. Daarnaast waren de meeste onderzoeken cross-sectioneel van aard, wat inhoudt dat adolescenten van verschillende leeftijden op één tijdspunt met elkaar werden vergeleken om ontwikkeling te onderzoeken. In longitudinale onderzoeken worden jongeren daadwerkelijk over langere tijd gevolgd, om ontwikkeling *binnen* personen te kunnen onderzoeken.

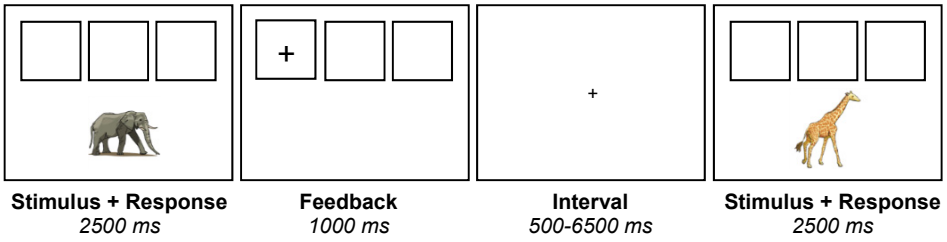
In dit proefschrift werden zowel cognitieve als affectieve aspecten van de ontwikkeling onderzocht in een grootschalig en longitudinaal onderzoek bij adolescenten. De volgende vragen stonden centraal: 1) hoe ontwikkelen hersengebieden voor cognitieve controle zich gedurende de adolescentie? 2) hoe beïnvloeden hersenverbindingen tussen affectieve en cognitieve gebieden de neiging tot risicogedrag?

Literatuuronderzoek naar ontwikkeling van hersengebieden voor cognitieve controle

In hoofdstuk 2 werd een literatuuroverzicht gegeven van verschillende onderzoeken naar de ontwikkeling van cognitieve controle, en cognitieve flexibiliteit in het bijzonder. Cognitieve flexibiliteit wordt gezien als een van de drie 'executieve functies', en wordt omschreven als het vermogen om je flexibel aan te passen aan een omgeving die continu verandert. In dit hoofdstuk hebben we een onderscheid gemaakt tussen twee verschillende manieren waarop cognitieve flexibiliteit kan worden gemeten: geïnstrueerde flexibiliteit en adaptieve flexibiliteit. Bij geïnstrueerde flexibiliteit wordt de participanten expliciet verteld dat ze moeten switchen in hun gedrag (bv. 'druk nu op de knop voor groene in plaats van rode figuren'). Bij adaptieve flexibiliteit wordt dit niet expliciet verteld, maar ontvangt de deelnemer negatieve feedback voor een gedrag dat eerder positieve feedback opleverde. De onderzoeken besproken in dit literatuuroverzicht laten zien dat voor zowel geïnstrueerde als adaptieve flexibiliteit een toename in de prestatie te zien is met toenemende leeftijd. Wanneer wordt gekeken naar hersenactiviteit tijdens het uitvoeren van deze taken, is er wel een verschil te zien: Voor geïnstrueerde flexibiliteit is er een toename te zien in cognitieve controlegebieden, wat overeenkomt met de voorspelling vanuit imbalance modellen dat het cognitieve systeem langzaam 'online' komt. Studies naar adaptieve flexibiliteit vonden echter dat er eerder sprake is van een specialisatie van cognitieve hersengebieden en niet simpelweg een toename in activatie met leeftijd. De resultaten lieten zien dat die hersengebieden met toenemende leeftijd steeds verder gespecialiseerd raken in het verwerken van feedback met informatieve waarde. Dat wil zeggen, de hersenen van jonge kinderen reageren nog op allerlei vormen van feedback, ook wanneer die feedback zo gemanipuleerd is dat die niet nuttig of informatief was voor de prestatie. Naarmate kinderen ouder werden, gingen de hersenen steeds meer onderscheid maken tussen informatieve en niet-informatieve vormen van feedback. Dit literatuuronderzoek vormde de basis voor de empirische studies die in dit proefschrift werden besproken.

Hersengebieden voor cognitieve controle: het frontoparietale netwerk

In het derde hoofdstuk werd een studie beschreven naar het leren van positieve en negatieve feedback bij 32 volwassenen participanten (18-25 jaar). Het doel van deze studie was om te onderzoeken welke hersengebieden betrokken zijn tijdens een feedback leertaak (een goede maat voor cognitieve controle) en of die gebieden voornamelijk actief worden na negatieve feedback (zoals gevonden in veel eerdere studies) en dus gevoelig zijn voor valentie, of dat deze gebieden meer gevoelig zijn voor de informatieve waarde van feedback. De deelnemers deden een taak in de MRI scanner waarbij steeds drie lege hokjes werden gepresenteerd met een figuur eronder. Ze werden geïnstrueerd om te ontdekken welke figuur (uit drie mogelijkheden) bij welk van de drie hokjes thuis hoorde (zie Figuur 1).



Figuur 1: Grafische weergave van de feedback leertaak.

In het begin moesten de deelnemers gokken in welk hokje het figuur thuishoorde. Na elke keuze kregen zij positieve of negatieve feedback te zien. Aan de hand van deze feedback kon uiteindelijk het juiste hokje voor elk figuur worden gevonden. Ondertussen heb ik gekeken naar de reacties van de hersenen op deze positieve en negatieve feedback. Hiervoor werd eerst onderscheid gemaakt tussen de 'leerfase' en 'toepasfase' van de taak. Tijdens de leerfase wisten de deelnemers nog niet in welk hokje het figuur thuis hoorde, en maakten zij dus nog daadwerkelijk gebruik van de feedback die zij ontvingen. De toepasfase begon als ze het figuur al een keer goed hadden geplaatst. De verdere feedback die zij in de toepasfase ontvingen had geen informatieve waarde meer, omdat de deelnemers de juiste locatie al wisten. De fMRI-resultaten lieten zien dat gebieden in het frontoparietale netwerk gevoelig waren voor zowel valentie (i.e., de gebieden waren meer actief na negatieve feedback dan na positieve feedback tijdens de leerfase, dus gecontroleerd voor informatieve waarde) als informatieve waarde (i.e., de gebieden waren meer actief gedurende de leerfase vergeleken met de toepasfase).

Ontwikkeling van het frontoparietale netwerk in de adolescentie

In hoofdstuk 4 werd een grootschalige cross-sectionele studie beschreven met 268 deelnemers tussen de 8 en 25 jaar oud. De deelnemers voerden dezelfde taak uit als beschreven in hoofdstuk 3. Het belangrijkste doel van deze studie was om de ontwikkeling van neurale activiteit in het

frontoparietale netwerk te beschrijven. Verschillende bevindingen kwamen naar voren uit dit onderzoek. Ten eerste werd gevonden dat neurale activiteit na negatieve feedback toenam met leeftijd, terwijl activiteit na positieve feedback niet veranderde gedurende de ontwikkeling. Interessant is dat de jongste kinderen wel degelijk activiteit in het frontoparietale netwerk lieten zien, maar in andere situaties dan volwassenen: Jonge kinderen lieten inderdaad minder activiteit dan volwassenen zien in de frontale en parietale cortex na negatieve feedback, zoals te verwachten op basis van imbalance modellen die voorspellen dat de frontoparietale gebieden nog niet ontwikkeld zijn. In precies hetzelfde gebied in de parietale cortex lieten kinderen echter juist *meer* activiteit zien dan volwassenen na positieve feedback. Dit betekent dat jonge kinderen de hersengebieden voor cognitieve controle wel degelijk kunnen gebruiken, maar op een andere manier dan volwassenen. Daarnaast werd gevonden dat de prestatie op de cognitieve taak een volwassen niveau bereikte rond de leeftijd van 14 jaar. De neurale reacties lieten een vergelijkbaar patroon zien: activiteit in de frontale cortex liet een toename zien tot de leeftijd van ongeveer 14 jaar, en verschilde daarna niet meer van de activiteit voor volwassenen. Voor de parietale cortex werd een volwassen niveau bereikt rond 15 jaar. De hormonen testosteron en oestradiol hadden geen invloed op activiteit tijdens deze cognitieve taak, wat past bij het idee dat puberteitshormonen voornamelijk affectieve maar niet cognitieve processen beïnvloeden.

In hoofdstuk 5 werd de data van dezelfde cross-sectionele studie op een andere manier geanalyseerd. In de feedback leertaak was het mogelijk om het leerproces te versnellen door verschillende redeneerstrategieën te gebruiken (bv. 'de leeuw hoorde niet thuis in het linker hokje, dus probeer ik de olifant als eerste in het linker hokje'). Met rekenmodellen werd gezocht naar latente groepen die van elkaar verschilden in strategiegebruik. De resultaten lieten zien dat er vier verschillende onderliggende groepen gedefinieerd konden worden. Hoewel leeftijd een groot deel van het strategiegebruik bepaalde, waren er bijvoorbeeld ook kinderen die betere strategieën gebruikten dan sommige volwassenen. Uit de fMRI-resultaten bleek dat er verschillen tussen de strategiegroepen waren in de mate waarin het frontoparietale netwerk werd geactiveerd tijdens de taak. Hoe beter het strategiegebruik, hoe meer activiteit in het frontoparietale netwerk werd waargenomen. Tot slot heb ik met een mediatieanalyse gekeken of strategiegebruik een verklaring was voor het feit dat oudere kinderen vaak meer activiteit in het frontoparietale netwerk laten zien dan jongere kinderen. Dit is belangrijk omdat eerdere onderzoekers vaak hebben beargumenteerd dat leeftijd de belangrijkste bepaler is voor toenames in activiteit met leeftijd, terwijl het ook mogelijk is dat dit verschil niet aan leeftijd an sich kan worden toegeschreven, maar meer aan verbeteringen in strategiegebruik. Deze hypothese werd bevestigd door de mediatie-analyse. Een deel van de toename in activiteit met leeftijd kon worden verklaard doordat oudere kinderen over het algemeen meer geavanceerde strategieën gebruikten.

In hoofdstuk 6 werd een longitudinale studie naar ontwikkeling in het frontoparietale netwerk beschreven. Deze studie was gebaseerd op de eerste meting besproken in hoofdstuk 3-5, maar daarnaast is de data gebruikt van een tweede meting bij dezelfde participanten na een fol-

low-up periode van twee jaar. In deze studie heb ik laten zien hoe de ontwikkeling van het frontoparietale netwerk verloopt binnen personen. De mediale prefrontale cortex liet een lineaire toename in activiteit met leeftijd zien, maar de laterale prefrontale cortex en parietale cortex een toename die meer kwadratisch van aard was, met een piek of afvlakking in de late adolescentie. Een toename in prestatie verklaarde extra variantie bovenop leeftijd voor veranderingen in activiteit in de dorsolaterale prefrontale en parietale cortex, terwijl corticale dikte (een maat voor de structuur van de grijze stof) juist extra variantie verklaarde in de mediale prefrontale cortex. Dit is een van de eerste studies die de longitudinale leeftijdspatronen in het frontoparietale netwerk in de adolescentie heeft bestudeerd, en evidentie voor non-lineaire ontwikkeling in dit netwerk heeft gevonden. Deze resultaten benadrukken het belang van het gebruiken van longitudinale designs waarin daadwerkelijk naar de ontwikkeling binnen personen wordt gekeken, omdat eerdere cross-sectionele studies mogelijk niet genoeg statistische power hadden om non-lineaire patronen in de data te ontdekken.

Voorspellen van leerprocessen in het echte leven

Hoofdstuk 7 is het laatste hoofdstuk over cognitieve aspecten van de ontwikkeling. In dit hoofdstuk werd gekeken of 'laboratorium' taken die worden gebruikt om het leerproces te bestuderen in gecontroleerde situaties (zoals de feedback leertaak), leerprocessen in het echte leven kunnen voorspellen. In dit onderzoek heb ik gekeken of prestatie en neurale activiteit tijdens de feedback leertaak voorspellend was voor lees- en rekenvaardigheid twee jaar later. De resultaten lieten zien dat prestatie op de feedback leertaak inderdaad kon voorspellen hoe goed een individu twee jaar later kon lezen en rekenen. Dit effect bleef zelfs bestaan na correctie voor leeftijd, werkgeheugen en IQ. Daarnaast was het mogelijk om aan de hand van neurale activiteit tijdens de feedback leertaak te voorspellen hoe goed iemand twee jaar later kon lezen en rekenen, en zelfs beter dan wanneer je alleen gedragsmaten van feedback leren zou gebruiken. Dit laat zien dat neurale activiteit belangrijk is om te onderzoeken bij voorspellen van latere reken- en leesvaardigheid, omdat het extra variantie verklaart die je niet met alleen gedragstesten kunt verklaren.

Risicogedrag in de adolescentie: connectiviteit tussen affectieve en cognitieve gebieden

In hoofdstuk 8 en 9 beschrijf ik twee studies waarin ik meer affectieve aspecten van de ontwikkeling heb onderzocht, namelijk de invloeden van testosteron en connectiviteit tussen affectieve en cognitieve gebieden op het alcoholgebruik van jongeren. In hoofdstuk 8 wordt een studie beschreven in dezelfde groep participanten als de eerdere hoofdstukken. Het doel van deze studie was om de relatie tussen hersenconnectiviteit, alcoholgebruik en testosteron te onderzoeken. Hiervoor werden 'resting-state' scans gemaakt. Om deze scans te maken werd de deelnemers gevraagd om 5 minuten lang stil te liggen met de ogen dicht. Er zijn verschillende netwerken in het brein actief die zelfs in 'rust' met elkaar communiceren. Onderzoekers hebben laten zien dat individuele verschillen in de mate waarin bepaalde gebieden met elkaar verbonden zijn in rust,

samenhangen met allerlei vormen van gedrag. In dit onderzoek hebben wij specifiek gekeken naar connectiviteit tussen de amygdala en de orbitofrontale cortex, omdat eerdere onderzoeken hebben laten zien dat deze verbinding gerelateerd is aan zowel testosteron als alcoholgebruik. Dit was echter nog niet in één studie onderzocht en tot nu toe alleen in volwassen participanten. De resultaten lieten zien dat voor jongens, een relatief hoge hoeveelheid testosteron samen hing met een verslechterde verbinding tussen de amygdala en de orbitofrontale cortex. Die verslechterde verbinding bleek weer samen te hangen met een verhoogd alcoholgebruik, zowel voor recent alcoholgebruik in de afgelopen maand als voor een schatting van het alcoholgebruik gedurende het hele leven van adolescenten.

In hoofdstuk 9 werd vervolgens de richting van dit effect onderzocht met een longitudinaal onderzoek. Ik heb gekeken of een verslechterde verbinding tussen de amygdala en de orbitofrontale cortex een gevolg was van alcoholgebruik, of dat alcoholgebruik juist een gevolg was van een verslechterde hersenverbinding. In deze studie heb ik laten zien dat connectiviteit tussen de amygdala en de orbitofrontale cortex voorspelt hoeveel alcohol een adolescent twee jaar later gebruikt. Hoe slechter de verbinding, hoe meer alcohol er werd gebruikt. Mogelijk zorgt een verslechterde verbinding voor verminderde controle van cognitieve hersengebieden over emotionele hersengebieden, wat kan leiden tot een toename van risicovol gedrag. Voor de omgekeerde richting werd geen bewijs gevonden: Meer alcoholgebruik voorspelde niet een verslechterde verbinding tussen de amygdala en de orbitofrontale cortex twee jaar later. Voor het specifieke pad tussen de amygdala en de orbitofrontale cortex tijdens rust heb ik dus geen bewijs voor een beschadigend effect van alcohol gevonden.

Conclusie

De onderzoeken in dit proefschrift geven een overzicht van zowel cognitieve als affectieve aspecten van de ontwikkeling, in relatie tot het brein en puberteitshormonen. De resultaten zijn belangrijk in de discussie rondom imbalance modellen van hersenontwikkeling in de adolescentie, en laten zien dat een meer genuanceerde beschrijving van ontwikkeling in hersengebieden voor cognitieve controle nodig is. Het cognitieve hersensysteem kan wel degelijk gebruikt worden door jonge kinderen en adolescenten, maar vaak in andere situaties dan bij volwassenen. Het populaire idee dat de prefrontale cortex achterloopt in de ontwikkeling ten opzichte van emotionele hersengebieden moet dus worden aangepast op basis van deze resultaten en bevindingen van andere onderzoekers. Mogelijk kan het cognitieve systeem meer flexibel worden ingezet in de adolescentie als de sociale of affectieve motivatie hoog genoeg is, maar dit idee moet bevestigd worden in toekomstig onderzoek. De bevindingen hebben belangrijke implicaties voor het opstellen van nieuwe theoretische modellen van de adolescentie die in de toekomst kunnen leiden tot onderwijs- en voorlichtingsinterventies die beter aansluiten bij de tekortkomingen maar vooral ook de mogelijkheden van het puberbrein.

References

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*, 20–29.

Althaus, M., Groen, Y., Wijers, A. A., Mulder, L. J., Minderaa, R. B., Kema, I. P., … Hoekstra, P. J. (2009). Differential effects of 5-HTTLPR and DRD2/ANKK1 polymorphisms on electrocortical measures of error and feedback processing in children. *Clinical Neurophysiology*, *120*, 93–107.

Ames, S. L., Grenard, J. L., Thush, C., Sussman, S., Wiers, R. W., & Stacy, A. W. (2007). Comparison of indirect assessments of association as predictors of marijuana use among at-risk adolescents. *Experimental Clinical Psychopharmacology*, *15*, 204–218.

Andersen, L. M., Visser, I., Crone, E. A., Koolschijn, P. C. M. P., & Raijmakers, M. E. J. (2014). Cognitive strategy use as an index of developmental differences in neural responses to feedback. *Developmental Psychology*, *50*, 2686–96.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+ 2= 4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*, *54*, 2382–2393.

Balsa, A. I., Homer, J. F., French, M. T., & Norton, E. C. (2011). Alcohol use and popularity: Social payoffs from conforming to peers' behavior. *Journal of Research on Adolescence*, *21*, 559–568.

Banks, S. J., Eddy, K. T., Angstadt, M., Nathan, P. J., & Phan, K. L. (2007). Amygdala-frontal connectivity during emotion regulation. *Social Cognitive and Affective Neuroscience*, *2*, 303–312.

Barcelo, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia*, *40*, 349–356.

Barnea-Goraly, N., Eliez, S., Menon, V., Bammer, R., & Reiss, A. L. (2005). Arithmetic ability and parietal alterations: a diffusion tensor imaging study in velocardiofacial syndrome. *Cognitive Brain Research*, *25*, 735–740.

Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., … Reiss, A. L. (2005). White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. *Cerebral Cortex*, *15*, 1848–1854.

Baxter, M. G., & Murray, E. A. (2002). The amygdala and reward. *Nature Reviews Neuroscience*, *3*, 563–573.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307.

Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *NeuroImage*, *20*, 1052–1063.

Bennett, C. M., Wolford, G. L., & Miller, M. B. (2009). The principled control of false positives in neuroimaging. *Social Cognitive and Affective Neuroscience*, *4*, 417–22.

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex*, *19*, 1583–96.

Biehl, M. C., Natsuaki, M. N., & Ge, X. J. (2007). The influence of pubertal timing on alcohol use and heavy drinking trajectories. *Journal of Youth and Adolescence*, *36*, 153–167.

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology*, *20*, 821–843.

Birn, R. M., Molloy, E. K., Patriat, R., Parker, T., Meier, T. B., Kirk, G. R., … Prabhakaran, V. (2013). The effect of scan length on the reliability of resting-state fMRI connectivity estimates. *NeuroImage*, *83*, 550–8.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647–663.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology*, *28*, 561–571.

Boettiger, C. A., & D'Esposito, M. (2005). Frontal networks for learning and executing arbitrary stimulus-response associations. *Journal of Neuroscience*, *25*, 2723–2732.

Booth, J. R., Burman, D. D., Meyer, J. R., Trommer, B. L., Davenport, N. D., Parrish, T. B., … Mesulam, M. M. (2004). Brain-behavior correlation in children depends on the neurocognitive network. *Human Brain Mapping*, *23*, 99–108.

Bos, P. A., Hermans, E. J., Ramsey, N. F., & van Honk, J. (2012). The neural mechanisms by which testosterone acts on interpersonal trust. *NeuroImage*, *61*, 730–737.

Bos, P. A., Panksepp, J., Bluthe, R. M., & van Honk, J. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: a review of single administration studies. *Frontiers in Neuroendocrinology*, *33*, 17–35.

Braams, B. R., Peters, S., Peper, J. S., Güroğlu, B., & Crone, E. A. (2014). Gambling for self, friends, and antagonists: differential contributions of affective and social brain regions on adolescent reward processing. *NeuroImage*, *100*, 281–9.

Braams, B. R., van Duijvenvoorde, A. C. K., Peper, J. S., & Crone, E. A. (2015). Longitudinal changes in adolescent risk-taking: a comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. *Journal of Neuroscience*, *35*, 7226–7238.

Bramen, J. E., Hranilovich, J. A., Dahl, R. E., Forbes, E. E., Chen, J., Toga, A. W., … Sowell, E. R. (2011). Puberty influences medial temporal lobe and cortical gray matter maturation differently in boys than girls matched for sexual maturity. *Cerebral Cortex*, *21*, 636–646.

Bratberg, G. H., Nilsen, T. I. L., Holmen, T. L., & Vatten, L. J. (2005). Sexual maturation in early adolescence and alcohol drinking and cigarette smoking in late adolescence: a prospective study of 2,129 Norwegian girls and boys. *European Journal of Pediatrics*, *164*, 621–625.

Brener, N. D., Kann, L., McManus, T., Kinchen, S. A., Sundberg, E. C., & Ross, J. G. (2002). Reliability of the 1999 youth risk behavior survey questionnaire. *Journal of Adolescent Health*, *31*, 336–342.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *NeuroImage*, *16*, S497.

Brovelli, A., Laksiri, N., Nazarian, B., Meunier, M., & Boussaoud, D. (2008). Understanding the neural computations of arbitrary visuomotor learning through fMRI and associative learning theory. *Cerebral Cortex*, *18*, 1485–1495.

Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. *Human Brain Mapping*, *25*, 35–45.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*, 205–228.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*, 273–293.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, *33*, 301–311.

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology*, *17*, 243–50.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31.

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive Affective & Behavioral Neuroscience*, *7*, 367–379.

Case, R. (1992). The role of the frontal lobes in the regulation of cognitive development. *Brain and Cognition*, *20*, 51–73.

Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology*, *66*, 295–319.

Casey, B. J., Davidson, M. C., Hara, Y., Thomas, K. M., Martinez, A., Galvan, A., … Tottenham, N. (2004). Early development of subcortical regions involved in non-cued attention switching. *Developmental Science*, *7*, 534–542.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, *54*, 241–257.

Casey, B. J., Jones, R. M., & Somerville, L. H. (2011). Braking and accelerating of the adolescent brain. *Journal of Research on Adolescence*, *21*, 21–33.

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. C. (2001). Changes in executive control across the life span: examination of task-switching performance. *Developmental Psychology*, *37*, 715–730.

Chamberlain, S. R., Blackwell, A. D., Fineberg, N. A., Robbins, T. W., & Sahakian, B. J. (2005). The neuropsychology of obsessive compulsive disorder: the importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neuroscience & Biobehavioral Reviews*, *29*, 399–419.

Chassin, L., Pitts, S. C., & Prost, J. (2002). Binge drinking trajectories from adolescence to emerging adulthood in a high-risk sample: predictors and substance abuse outcomes. *Journal of Consulting and Clinical Psychology*, *70*, 67–78.

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science*, *14*, F1–10.

Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting test. *Journal of Clinical and Experimental Neuropsychology*, *8*, 219–228.

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: effect of cue transparency on children's cognitive flexibility. *Developmental Psychology*, *45*, 782–797.

Choudhury, S. (2010). Culturing the adolescent brain: what can neuroscience learn from anthropology? *Social Cognitive and Affective Neuroscience*, *5*, 159–67.

Christakou, A., Halari, R., Smith, A. B., Ifkovits, E., Brammer, M., & Rubia, K. (2009). Sex-dependent age modulation of frontostriatal and temporo-parietal activation during cognitive control. *NeuroImage*, *48*, 223–236.

Cocosco, C. A., Kollokian, V., Kwan, R. K. S., & Evans, A. C. (1997). Online interface to a 3D MRI simulated brain database. *NeuroImage*, *5*, S425.

Cole, M. W., Pathak, S., & Schneider, W. (2010). Identifying the brain's most globally connected regions. *NeuroImage*, *49*, 3132–3148.

Colzato, L. S., Waszak, F., Nieuwenhuis, S., Posthuma, D., & Hommel, B. (2010). The flexible mind is associated with the catechol-O-methyltransferase (COMT) Val158Met polymorphism: evidence for a role of dopamine in the control of task-switching. *Neuropsychologia*, *48*, 2764–2768.

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research*, *166*, 210–222.

Cragg, L., & Chevalier, N. (2012). The processes underlying flexibility in childhood. *Quarterly Journal of Experimental Psychology*, *65*, 209–232.

Crone, E. A., Bunge, S. A., van der Molen, M. W., & Ridderinkhof, K. R. (2006). Switching between tasks and responses: a developmental study. *Developmental Science*, *9*, 278–287.

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*, 636–650.

Crone, E. A., Ridderinkhof, K. R., Worm, M., Somsen, R. J., & van der Molen, M. W. (2004). Switching between spatial stimulus-response mappings: a developmental study of cognitive flexibility. *Developmental Science*, *7*, 443–455.

Crone, E. A., Wendelken, C., Donohue, S. E., & Bunge, S. A. (2006). Neural evidence for dissociable components of task-switching. *Cerebral Cortex*, *16*, 475–486.

Crone, E. A., Zanolie, K., Van Leijenhorst, L., Westenberg, P. M., & Rombouts, S. A. (2008). Neural mechanisms supporting flexible performance adjustment during development. *Cognitive, Affective & Behavioral Neuroscience*, *8*, 165–177.

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, *8*, 109–114.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, *9*, 179–194.

Darki, F., & Klingberg, T. (2014). The role of fronto-parietal and fronto-striatal networks in the development of working memory: a longitudinal study. *Cerebral Cortex*, *25*, bht352.

de Water, E., Braams, B. R., Crone, E. A., & Peper, J. S. (2013). Pubertal maturation and sex steroids are related to alcohol use in adolescents. *Hormones and Behavior*, *63*, 392–397.

Derntl, B., Windischberger, C., Robinson, S., Kryspin-Exner, I., Gur, R. C., Moser, E., & Habel, U. (2009). Amygdala activity to fear and anger in healthy young males is associated with testosterone. *Psychoneuroendocrinology*, *34*, 687–693.

Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Human Brain Mapping*, *25*, 22–34.

DeWitt, S. J., Aslan, S., & Filbey, F. M. (2014). Adolescent risk-taking and resting state functional connectivity. *Psychiatry Research*, *222*, 157–64.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*, 135–168.

Dosenbach, N. U. F., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., … Schlaggar, B. L. (2010). Prediction of individual brain maturity using fMRI. *Science*, *329*, 1358–61.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Brain Research*, *9*, 103–109.

Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex*, *22*, 1078–1085.

Eliassen, J. C., Lamy, M., Allendorfer, J. B., Boespflug, E., Bullard, D. P., Smith, M. S., … Strakowski, S. M. (2012). Selective role for striatal and prefrontal regions in processing first trial feedback during single-trial associative learning. *Brain Research*, *1458*, 56–66.

Emerson, R. W., & Cantlon, J. F. (2014). Continuity and change in children's longitudinal neural responses to numbers. *Developmental Science*, *18*, 314–326.

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: evidence from ERPs. *Psychophysiology*, *46*, 1043–1053.

Ernst, M., Pine, D. S., & Hardin, M. (2006). Triadic model of the neurobiology of motivated behavior in adolescence. *Psychological Medicine*, *36*, 299–312.

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., … Petersen, S. E. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology*, *5*, e1000381.

Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, *29*, 581–593.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*, 11050–11055.

Fjell, A. M., Walhovd, K. B., Westlye, L. T., Ostby, Y., Tamnes, C. K., Jernigan, T. L., … Dale, A. M. (2010). When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. *NeuroImage*, *50*, 1376–1383.

Forbes, E. E., & Dahl, R. E. (2010). Pubertal development and behavior: hormonal activation of social and motivational tendencies. *Brain and Cognition*, *72*, 66–72.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, *33*, 636–47.

Fortier, M. S., Vallerand, R. J., & Guay, F. (1995). Academic motivation and school performance: Toward a structural model. *Contemporary Educational Psychology*, *20*, 257–274.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, *8*, 700–711.

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology*, *137*, 201–225.

Frost, J. A., Binder, J. R., Springer, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Language processing is strongly left lateralized in both sexes: Evidence from functional MRI. *Brain*, *122*, 199–208.

Gabard-Durnam, L. J., Flannery, J., Goff, B., Gee, D. G., Humphreys, K. L., Telzer, E., … Tottenham, N. (2014). The development of human amygdala functional connectivity at rest from 4 to 23 years: A cross-sectional study. *NeuroImage*, *95*, 193–207.

Galvan, A. (2013). The teenage brain: Sensitivity to rewards. *Current Directions in Psychological Science*, *22*, 88–93.

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of Neuroscience*, *26*, 6885–6892.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*, 31–60.

Geary, D. C. (2000). From infancy to adulthood: The development of numerical abilities. *European Child & Adolescent Psychiatry*, *9*, S11–S16.

Gee, D. G., Gabard-Durnam, L. J., Flannery, J., Goff, B., Humphreys, K. L., Telzer, E. H., … Tottenham, N. (2013). Early developmental emergence of human amygdala-prefrontal connectivity after maternal deprivation. *Proceedings of the National Academy of Sciences*, *110*, 15638–15643.

Geier, C. F., Garver, K., Terwilliger, R., & Luna, B. (2009). Development of working memory maintenance. *Journal of Neurophysiology*, *101*, 84–99.

Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Sciences*, *1021*, 77–85.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., … Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, *2*, 861–863.

Giedd, J. N., Lalonde, F. M., Celano, M. J., White, S. L., Wallace, G. L., Lee, N. R., & Lenroot, R. K. (2009). Anatomical brain magnetic resonance imaging of typically developing children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*, 465–70.

Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: What have we learned and where are we going? *Neuron*, *67*, 728–734.

Gilman, J. M., Ramchandani, V. A., Crouss, T., & Hommer, D. W. (2012). Subjective and neural responses to intravenous alcohol in young adults with light and heavy drinking patterns. *Neuropsychopharmacology*, *37*, 467–477.

Gilman, J. M., Ramchandani, V. A., Davis, M. B., Bjork, J. M., & Hommer, D. W. (2008). Why we like to drink: A functional magnetic resonance imaging study of the rewarding and anxiolytic effects of alcohol. *Journal of*

Neuroscience, *28*, 4583–4591.

Giorgio, A., Watkins, K. E., Chadwick, M., James, S., Winmill, L., Douaud, G., … James, A. C. (2010). Longitudinal changes in grey and white matter during adolescence. *NeuroImage*, *49*, 94–103.

Glahn, D. C., Lovallo, W. R., & Fox, P. T. (2007). Reduced amygdala activation in young adults at high risk of alcoholism: Studies from the Oklahoma family health patterns project. *Biological Psychiatry*, *61*, 1306–1309.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., … Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 8174–9.

Gorka, S. M., Fitzgerald, D. A., King, A. C., & Phan, K. L. (2013). Alcohol attenuates amygdala-frontal connectivity during processing social signals in heavy social drinkers A preliminary pharmaco-fMRI study. *Psychopharmacology*, *229*, 141–154.

Grant, J. D., Scherrer, J. F., Lynskey, M. T., Lyons, M. J., Eisen, S. A., Tsuang, M. T., … Bucholz, K. K. (2006). Adolescent alcohol use is a risk factor for adult alcohol and drug dependence: evidence from a twin design. *Psychological Medicine*, *36*, 109–118.

Haier, R. J., Karama, S., Leyba, L., & Jung, R. E. (2009). MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task. *BMC Research Notes*, *2*, 174.

Harms, M. P., Wang, L., Csernansky, J. G., & Barch, D. M. (2013). Structure-function relationship of working memory activity with hippocampal and prefrontal cortex volumes. *Brain Structure and Function*, *218*, 173–186.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Communication Monographs*, *76*, 408–420.

Hegarty, C. E., Foland-Ross, L. C., Narr, K. L., Townsend, J. D., Bookheimer, S. Y., Thompson, P. M., & Altshuler, L. L. (2012). Anterior cingulate activation relates to local cortical thickness. *Neuroreport*, *23*, 420–424.

Hermans, E. J., Ramsey, N. F., & van Honk, J. (2008). Exogenous testosterone enhances responsiveness to social threat in the neural circuitry of social aggression in humans. *Biological Psychiatry*, *63*, 263–270.

Herting, M. M., Maxwell, E. C., Irvine, C., & Nagel, B. J. (2012). The impact of sex, puberty, and hormones on white matter microstructure in adolescents. *Cerebral Cortex*, *22*, 1979–1992.

Hibell, B., Guttormsson, U., Ahlström, S., Balakireva, O., Bjarnason, T., Kokkevi, A., & Kraus, L. (2012). The 2011 ESPAD Report. Substance use among students in 36 European countries. Tukholma: The Swedish Council for Information on Alcohol and other Drugs, 2012.

Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., … Jo, B. (2007). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience*, *121*, 602.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*, 497–498.

Houdé, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: an fMRI meta-analysis of 52 studies including 842 children. *Developmental Science*, *13*, 876–885.

Huizinga, M., Dolan, C. V, & van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017–2036.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*, 309–334.

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology*, *30*, 190–200.

Jansen, B. R. J., van Duijvenvoorde, A. C. K., & Huizenga, H. M. (2014). Developmental and gender related differences in response switches after nonrepresentative negative feedback. *Developmental Psychology*, *50*, 237–246.

Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, *95*, 719.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*, 825–841.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*, 143–156.

Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology*, *45*, 482–553.

Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*, 7–21.

Jolles, D. J., & Crone, E. A. (2012). Training the developing brain: A neurocognitive perspective. *Frontiers in Human Neuroscience*, *6*, 76.

Jolles, D. J., van Buchem, M. A., Crone, E. A., & Rombouts, S. A. R. B. (2011). A comprehensive study of whole-brain functional connectivity in children and young adults. *Cerebral Cortex*, *21*, 385–91.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*, 437.

Kann, L., Kinchen, S., Shanklin, S. L., Flint, K. H., Kawkins, J., Harris, W. A., … Zaza, S. (2014). Youth risk behavior surveillance--United States, 2013. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C. : 2002)*, *63 Suppl 4*, 1–168.

Kerns, J. G., Cohen, J. D., MacDonald 3rd, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*, 1023–1026.

Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: a meta-analysis. *Human Brain Mapping*, *33*, 130–142.

Kleibeuker, S. W., De Dreu, C. K. W., & Crone, E. A. (2013). The development of creative cognition across adolescence: distinct trajectories for insight and divergent thinking. *Developmental Science*, *16*, 2–12.

Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y., & Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science*, *318*, 1642–1645.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience*, *14*, 1–10.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: an fMRI study. *NeuroImage*, *25*, 1068–1076.

Konishi, S., Hayashi, T., Uchida, I., Kikyo, H., Takahashi, E., & Miyashita, Y. (2002). Hemispheric asymmetry in human lateral prefrontal cortex during cognitive set shifting. *Proceedings of the National Academy of Sciences*, *99*, 7803–7808.

Koolschijn, P. C. M. P., & Crone, E. A. (2013). Sex differences and structural brain maturation from childhood to early adulthood. *Developmental Cognitive Neuroscience*, *5*, 106–118.

Koolschijn, P. C. M. P., Peper, J. S., & Crone, E. A. (2014). The influence of sex steroids on structural brain maturation in adolescence. *PloS One*.

Koolschijn, P. C. M. P., Schel, M. A., de Rooij, M., Rombouts, S. A., & 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. *Journal of Neuroscience*, *31*, 4204–4212.

Kringelbach, M. L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nature Reviews Neuroscience*, *6*, 691–702.

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). DMT en AVI: Wetenschappelijke verantwoording. CITO Instituut voor Testontwikkeling.

Landi, N., Frost, S. J., Mencl, W. E., Sandak, R., & Pugh, K. R. (2013). Neurobiological bases of reading comprehension: insights from neuroimaging studies of word-level and text-level processing in skilled and impaired readers. *Reading & Writing Quarterly*, *29*, 145–167.

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence*, *12*, 131–147.

Lebel, C., & Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *Journal of Neuroscience*, *31*, 10937–10947.

Lee, H., Heller, A. S., van Reekum, C. M., Nelson, B., & Davidson, R. J. (2012). Amygdala-prefrontal coupling underlies individual differences in emotion regulation. *NeuroImage*, *62*, 1575–1581.

Leisch, F. (2004). FlexMix: a general framework for finite mixture models and latent class regression in R. *Journal of Statistical Software*, *11*, 1–18.

Li, Y., Hu, Y., Wang, Y., Weng, J., & Chen, F. (2013). Individual structural differences in left inferior parietal area are associated with schoolchildrens' arithmetic scores. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00844

Lie, C. H., Specht, K., Marshall, J. C., & Fink, G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *NeuroImage*, *30*, 1038–1049.

Lu, L. H., Dapretto, M., O'Hare, E. D., Kan, E., McCourt, S. T., Thompson, P. M., … Sowell, E. R. (2009). Relationships between brain activation and brain structure in normally developing children. *Cerebral Cortex*, *19*, 2595–2604.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*, 1357–72.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*, 173–181.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83–104.

Madden, D. J., Costello, M. C., Dennis, N. A., Davis, S. W., Shepler, A. M., Spaniol, J., … Cabeza, R. (2010). Adult age differences in functional connectivity during executive control. *NeuroImage*, *52*, 643–657.

Manuck, S. B., Marsland, A. L., Flory, J. D., Gorka, A., Ferrell, R. E., & Hariri, A. R. (2010). Salivary testosterone and a trinucleotide (CAG) length polymorphism in the androgen receptor gene predict amygdala reactivity in men. *Psychoneuroendocrinology*, *35*, 94–104.

Mars, R. B., Coles, M. G., Grol, M. J., Holroyd, C. B., Nieuwenhuis, S., Hulstijn, W., & Toni, I. (2005). Neural dynamics of error processing in medial frontal cortex. *NeuroImage*, *28*, 1007–1013.

Marshall, W. A., & Tanner, J. M. (1970). Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood*, *44*, 291–303.

Maurer, U., Bucher, K., Brem, S., Benz, R., Kranz, F., Schulz, E., … Brandeis, D. (2009). Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Biological Psychiatry*, *66*, 341–348.

McBride, W. J. (2002). Central nucleus of the amygdala and the effects of alcohol and alcohol-drinking behavior in rodents. *Pharmacology Biochemistry and Behavior*, *71*, 509–515.

McLachlan, G. J., & Peel, D. (2000). *Finite Mixture Models*. New York: Wiley.

Milner, B. (1963). Effects of different brain lesions on card sorting - Role of frontal lobes. *Archives of Neurology*, *9*, 90–&.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and

diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, *21*, 7733–7741.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*, 134–140.

Morton, J. B. (2010). Understanding genetic, neurophysiological, and experiential influences on the development of executive functioning: the need for developmental models. *Wiley Interdisciplinary Reviews-Cognitive Science*, *1*, 709–723.

Morton, J. B., Bosma, R., & Ansari, D. (2009). Age-related changes in brain activation associated with dimensional shifts of attention: an fMRI study. *NeuroImage*, *46*, 249–256.

Nelson, E. E., Leibenluft, E., McClure, E. B., & Pine, D. S. (2005). The social re-orientation of adolescence: a neuroscience perspective on the process and its relation to psychopathology. *Psychological Medicine*, *35*, 163–74.

Nguyen, T. V, McCracken, J., Ducharme, S., Botteron, K. N., Mahabir, M., Johnson, W., … Karama, S. (2013). Testosterone-related cortical maturation across childhood and adolescence. *Cerebral Cortex*, *23*, 1424–1432.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective & Behavioral Neuroscience*, *12*, 241–68.

Nieuwenhuis, S., Slagter, H. A., von Geusau, N. J., Heslenfeld, D. J., & Holroyd, C. B. (2005). Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *The European Journal of Neuroscience*, *21*, 3161–3168.

Nunez, S. C., Dapretto, M., Katzir, T., Starr, A., Bramen, J., Kan, E., … Sowell, E. R. (2011). fMRI of syntactic processing in typically developing children: structural correlates in the inferior frontal gyrus. *Developmental Cognitive Neuroscience*, *1*, 313–323.

Nyffeler, T., Muri, R. M., Bucher-Ottiger, Y., Pierrot-Deseilligny, C., Gaymard, B., & Rivaud-Pechoux, S. (2007). Inhibitory control of the human dorsolateral prefrontal cortex during the anti-saccade paradigm - a transcranial magnetic stimulation study. *European Journal of Neuroscience*, *26*, 1381–1385.

O'Shaughnessy, T. E., Lane, K. L., Gresham, F. M., & Beebe-Frankenberger, M. E. (2003). Children placed at risk for learning and behavioral difficulties implementing a school-wide system of early identification and intervention. *Remedial and Special Education*, *24*, 27–35.

Ordaz, S. J., Foran, W., Velanova, K., & Luna, B. (2013). Longitudinal growth curves of brain function underlying inhibitory control through adolescence. *Journal of Neuroscience*, *33*, 18109–18124.

Ozyurt, J., Rietze, M., & Thiel, C. M. (2012). Prefrontal neural activity when feedback is not relevant to adjust performance. *PloS One*, *7*, e36509.

Paus, T., Nawaz-Khan, I., Leonard, G., Perron, M., Pike, G. B., Pitiot, A., … Pausova, Z. (2010). Sexual dimorphism in the adolescent brain: Role of testosterone and androgen receptor in global and local volumes of grey and white matter. *Hormones and Behavior*, *57*, 63–75.

Peeters, M., Janssen, T., Monshouwer, K., Boendemaker, W., Pronk, T., Wiers, R., & Vollebergh, W. (2015). Weaknesses in executive functioning predict the initiating of adolescents' alcohol use. *Developmental Cognitive Neuroscience*.

Peper, J. S., & Dahl, R. E. (2013). The teenage brain: Surging hormones - Brain-behavior interactions during puberty. *Current Directions in Psychological Science*, *22*, 134–139.

Peper, J. S., Koolschijn, P. C. M. P., & Crone, E. A. (2013). Development of risk taking: contributions from adolescent testosterone and the orbito-frontal cortex. *Journal of Cognitive Neuroscience*, *25*, 2141–2150.

Peper, J. S., Mandl, R. C. W., Braams, B. R., de Water, E., Heijboer, A. C., Koolschijn, P. C. M. P., & Crone, E. A. (2013). Delay Discounting and Frontostriatal Fiber Tracts: A Combined DTI and MTR Study on Impulsive Choices in Healthy Young Adults. *Cerebral Cortex*, *23*, 1695–1702.

Peters, S., Braams, B. R., Raijmakers, M. E. J., Koolschijn, P. C. M. P., & Crone, E. A. (2014). The neural coding of feedback learning across child and adolescent development. *Journal of Cognitive Neuroscience*, *26*, 1705–1720.

Peters, S., & Crone, E. A. (2014). Cognitive flexibility in childhood and adolescence. In *Task Switching and Cognitive Control*. USA: Oxford University Press.

Peters, S., Jolles, D. J., Van Duijvenvoorde, A. C. K., Crone, E. A., & Peper, J. S. (2015). The link between testosterone and amygdala–orbitofrontal cortex connectivity in adolescent alcohol use. *Psychoneuroendocrinology*, *53*, 117–126.

Peters, S., Koolschijn, P. C. M. P., Crone, E. A., Van Duijvenvoorde, A. C. K., & Raijmakers, M. E. J. (2014). Strategies influence neural activity for feedback learning across child and adolescent development. *Neuropsychologia*, *62*, 365–374.

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*, 117–33.

Pfeifer, J. H., & Allen, N. B. (2012). Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. *Trends in Cognitive Sciences*, *16*, 322–9.

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2007). NLME: Linear and nonlinear mixed effects models. R package version 3.1-122, http://CRAN.R-project.org/package=nlme.

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*, 1746–1758.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717–731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavioral Research Methods*, *40*, 879–891.

Rabiner, L. R. (1989). A tutorial on Hidden Markov-Models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*, 257–286.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110–122.

Raichle, M. E. (2010). Two views of brain function. *Trends in Cognitive Sciences*, *14*, 180–90.

Raijmakers, M. E. J., Dolan, C. V, & Molenaar, P. C. M. (2001). Finite mixture distribution models of simple discrimination learning. *Memory & Cognition*, *29*, 659–677.

Ravizza, S. M., & Carter, C. S. (2008). Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia*, *46*, 2924–2935.

Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G. L., Greenstein, D., … Giedd, J. N. (2011). How does your cortex grow? *Journal of Neuroscience*, *31*, 7174–7.

Reuter, M., & Fischl, B. (2011). Avoiding asymmetry-induced bias in longitudinal image processing. *NeuroImage*, *57*, 19–21.

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: a robust approach. *NeuroImage*, *53*, 1181–1196.

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*, 1402–1418.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*, 443–447.

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources*, 313–328.

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E., & Brammer, M. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Human Brain Mapping*, *27*, 973–993.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in Social Psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, *5*, 359–371.

Scherf, K. S., Smyth, J. M., & Delgado, M. R. (2013). The amygdala: An agent of change in adolescent neural networks. *Hormones and Behavior*, *64*, 298–313.

Schmittmann, V. D., van der Maas, H. L. J., & Raijmakers, M. E. J. (2012). Distinct discrimination learning strategies and their relation with spatial memory and attentional control in 4-to 14-year-olds. *Journal of Experimental Child Psychology*, *111*, 644–662.

Schmittmann, V. D., Visser, I., & Raijmakers, M. E. (2006). Multiple learning modes in the development of performance on a rule-based category-learning task. *Neuropsychologia*, *44*, 2079–2091.

Schulz, K. P., Bedard, A. C. V, Czarnecki, R., & Fan, J. (2011). Preparatory activity and connectivity in dorsal anterior cingulate cortex for cognitive control. *NeuroImage*, *57*, 242–250.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

Schweinsburg, A. D., Nagel, B. J., & Tapert, S. F. (2005). fMRI reveals alteration of spatial working memory networks across adolescence. *Journal of the International Neuropsychological Society*, *11*, 631–44.

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*, 1060–1075.

Sells, C. W., & Blum, R. W. (1996). Morbidity and mortality among US adolescents: An overview of data and trends. *American Journal of Public Health*, *86*, 513–519.

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., … Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience*, *28*, 3586–94.

Shaw, P., Malek, M., Watson, B., Greenstein, D., de Rossi, P., & Sharp, W. (2013). Trajectories of cerebral cortical development in childhood and adolescence and adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *74*, 599–606.

Shirtcliff, E. A., Dahl, R. E., & Pollak, S. D. (2009). Pubertal development: correspondence between hormonal and physical development. *Child Development*, *80*, 327–337.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychol Ogical Methods*, *7*, 422–445.

Simerly, R. B., Chang, C., Muramatsu, M., & Swanson, L. W. (1990). Distribution of androgen and estrogen receptor mRNA-containing cells in the rat brain: an in situ hybridization study. *The Journal of Comparative Neurology*, *294*, 76–95.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*, 143–155.

Sobell, L. C., & Sobell, M. B. (1990). Self-report issues in alcohol abuse - State-of-the-art and future directions. *Behavioral Assessment*, *12*, 77–90.

Somerville, L. H., & Casey, B. J. (2010). Developmental neurobiology of cognitive control and motivational systems. *Current Opinion in Neurobiology*, *20*, 236–241.

Somerville, L. H., Jones, R. M., & Casey, B. J. (2010). A time of change: behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and Cognition*, *72*, 124–33.

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science*, *24*, 1554–62.

Speekenbrink, M., Lagnado, D. A., Wilkinson, L., Jahanshahi, M., & Shanks, D. R. (2010). Models of probabilistic

category learning in Parkinson's disease: Strategy use and the effects of L-dopa. *Journal of Mathematical Psychology*, *54*, 123–136.

Spielberg, J. M., Forbes, E. E., Ladouceur, C. D., Worthman, C. M., Olino, T. M., Ryan, N. D., & Dahl, R. E. (2014). Pubertal testosterone influences threat-related amygdala-orbitofrontal coupling. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsu062

Squeglia, L. M., Jacobus, J., & Tapert, S. F. (2009). The influence of substance use on adolescent brain development. *Clinical EEG Neuroscience*, *40*, 31–38.

Squeglia, L. M., McKenna, B. S., Jacobus, J., Castro, N., Sorg, S. F., & Tapert, S. F. (2013). BOLD response to working memory not related to cortical thickness during early adolescence. *Brain Research*, *1537*, 59–68.

Sripada, C. S., Angstadt, M., McNamara, P., King, A. C., & Phan, K. L. (2011). Effects of alcohol on brain responses to social signals of threat in humans. *NeuroImage*, *55*, 371–380.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*, 745–759.

Stanton, S. J., Wirth, M. M., Waugh, C. E., & Schultheiss, O. C. (2009). Endogenous testosterone levels are associated with amygdala and ventromedial prefrontal cortex responses to anger faces in men but not women. *Biological Psychology*, *81*, 118–122.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*, 78–106.

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. *Developmental Psychology*, *44*, 1764–78.

Stelzel, C., Basten, U., & Fiebach, C. J. (2011). Functional connectivity separates switching operations in the posterior lateral frontal cortex. *Journal of Cognitive Neuroscience*, *23*, 3529–3539.

Stevens, M. C. (2009). The developmental cognitive neuroscience of functional connectivity. *Brain and Cognition*, *70*, 1–12.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453–489.

Talairach, J., & Tourneaux, P. (1988). *Co-planar stereotactic atlas of the human brain.* New York: Thieme.

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 1231–8.

Tamnes, C. K., Østby, Y., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., & Walhovd, K. B. (2010). Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cerebral Cortex*, *20*, 534–548.

Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: A review of developmental changes in the error-related negativity and brain maturation. *Developmental Cognitive Neuroscience*, *6C*, 1–13.

Taylor, S. J., Barker, L. A., Heavey, L., & McHale, S. (2013). The typical developmental trajectory of social and executive functions in late adolescence and early adulthood. *Developmental Psychology*, *49*, 1253–65.

Tchanturia, K., Davies, H., Roberts, M., Harrison, A., Nakazato, M., Schmidt, U., … Morris, R. (2012). Poor cognitive flexibility in eating disorders: examining the evidence using the Wisconsin Card Sorting Task. *PloS One*, *7*, e28331.

Teslovich, T., Mulder, M., Franklin, N. T., Ruberry, E. J., Millner, A., Somerville, L. H., … Casey, B. J. (2014). Adolescents let sufficient evidence accumulate before making a decision when large incentives are at stake. *Developmental Science*, *17*, 59–70.

Thomas, L. A., Hall, J. M., Skup, M., Jenkins, S. E., Pine, D. S., & Leibenluft, E. (2011). A developmental neuroimaging investigation of the change paradigm. *Developmental Science*, *14*, 148–161.

Thomason, M. E., Race, E., Burrows, B., Whitfield-Gabrieli, S., Glover, G. H., & Gabrieli, J. D. E. (2009). Development of spatial and verbal working memory capacity in the human brain. *Journal of Cognitive Neuroscience*, *21*, 316–332.

Thush, C., Wiers, R. W., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2008). Interactions between implicit and explicit cognition and working memory capacity in the prediction of alcohol use in at-risk adolescents. *Drug and Alcohol Dependence*, *94*, 116–124.

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. A. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience*, *18*, 1029–1043.

Uddin, L. Q., Supekar, K. S., Ryali, S., & Menon, V. (2011). Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. *Journal of Neuroscience*, *31*, 18578–18589.

Ullman, H., Almeida, R., & Klingberg, T. (2014). Structural maturation and brain activity predict future working memory capacity during childhood development. *Journal of Neuroscience*, *34*, 1592–1598.

Ullsperger, M., & von Cramon, D. Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *The Journal of Neuroscience*, *23*, 4308–4314.

van den Bos, Güroğlu, B., van den Bulk, B. G., Rombouts, S. a R. B., & Crone, E. a. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience*, *3*, 52.

van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. a. (2012). Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral Cortex*, *22*, 1247–1255.

van der Maas, H. L. J., & Straatemeier, M. (2008). How to detect cognitive strategies: commentary on "Differentiation and integration: guiding principles for analyzing cognitive change." *Developmental Science*, *11*, 449–453.

Van der Sluis, S., De Jong, P. F., & Van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology*, *87*, 239–266.

Van Duijvenvoorde, A. C. K., Jansen, B. R. J., Griffioen, E. S., Van der Molen, M. W., & Huizenga, H. M. (2013). Decomposing developmental differences in probabilistic feedback learning: A combined performance and heart-rate analysis. *Biological Psychology*, *93*, 175–183.

van Duijvenvoorde, A. C. K., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *Journal of Neuroscience*, *28*, 9495–9503.

van Holstein, M., Aarts, E., van der Schaaf, M. E., Geurts, D. E., Verkes, R. J., Franke, B., … Cools, R. (2011). Human cognitive flexibility depends on dopamine D2 receptor signaling. *Psychopharmacology*, *218*, 567–578.

Van Leijenhorst, L., Gunther Moor, B., Op de Macks, Z. A., Rombouts, S. A. R. B., Westenberg, P. M., & Crone, E. A. (2010). Adolescent risky decision-making: neurocognitive development of reward and control regions. *NeuroImage*, *51*, 345–55.

van Schouwenburg, M. R., den Ouden, H. E., & Cools, R. (2010). The human basal ganglia modulate frontal-posterior connectivity during attention shifting. *Journal of Neuroscience*, *30*, 9910–9918.

van Veen, V., Holroyd, C. B., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2004). Errors without conflict: implications for performance monitoring theories of anterior cingulate cortex. *Brain and Cognition*, *56*, 267–276.

van Wingen, G., Mattern, C., Verkes, R. J., Buitelaar, J., & Fernandez, G. (2010). Testosterone reduces amygdala-orbitofrontal cortex coupling. *Psychoneuroendocrinology*, *35*, 105–113.

Velanova, K., Wheeler, M. E., & Luna, B. (2008). Maturational changes in anterior cingulate and frontoparietal recruitment support the development of error processing and inhibitory control. *Cerebral Cortex*, *18*, 2505–22.

Visser, I. (2011). Seven things to remember about hidden Markov models: A tutorial on Markovian models for time series. *Journal of Mathematical Psychology*, *55*, 403–415.

Visser, I., & Speekenbrink, M. (2010). DepmixS4: An R Package for Hidden Markov Models. *Journal of Statistical Software*, *36*, 1–21.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2005). Frontomedian activation depends on both feedback validity and valence: fMRI evidence for contextual feedback evaluation. *NeuroImage*, *27*, 564–571.

Von Der Heide, R. J., Skipper, L. M., Klobusicky, E., & Olson, I. R. (2013). Dissecting the uncinate fasciculus: disorders, controversies and a hypothesis. *Brain*, *136*, 1692–1707.

Wager, T. D., Jonides, J., & Reading, S. (2004). Neuroimaging studies of shifting attention: a meta-analysis. *NeuroImage*, *22*, 1679–1693.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*, *3*, 255–274.

Wechsler, H., & Nelson, T. F. (2001). Binge drinking and the American college student: what's five drinks? *Psychology of Addictive Behaviors*, *15*, 287–91.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative developmental study of executive function - A window on prefrontal function in children. *Developmental Neuropsychology*, *7*, 131–149.

Wendelken, C., O'Hare, E. D., Whitaker, K. J., Ferrer, E., & Bunge, S. A. (2011). Increased functional selectivity over development in rostrolateral prefrontal cortex. *Journal of Neuroscience*, *31*, 17260–17268.

Westling, E., Andrews, J. A., Hampson, S. E., & Peterson, M. (2008). Pubertal timing and substance use: The effects of gender, parental monitoring and deviant peers. *Journal of Adolescent Health*, *42*, 555–563.

White, H. R., Xie, M., Thompson, W., Loeber, R., & Stouthamer-Loeber, M. (2001). Psychopathology as a predictor of adolescent drug use trajectories. *Psychology of Addictive Behaviors*, *15*, 210–8.

Witt, E. D. (2007). Puberty, hormones, and sex differences in alcohol abuse and dependence. *Neurotoxicology and Teratology*, *29*, 81–95.

Wolf, L., Bazargani, N., Kilford, E., Dumontheil, I., & Blakemore, S. (2015, May 21). The audience effect in adolescence depends on who's looking over your shoulder. *Journal of Adolescence*. Elsevier.

Woolrich, M. (2008). Robust group analysis using outlier inference. *NeuroImage*, *41*, 286–301.

Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*, 1732–1747.

Zanolie, K., Van Leijenhorst, L., Rombouts, S. A., & Crone, E. A. (2008). Separable neural mechanisms contribute to feedback processing in a rule-learning task. *Neuropsychologia*, *46*, 117–126.

Zanto, T. P., Pa, J., & Gazzaley, A. (2014). Reliability measures of functional magnetic resonance imaging in a longitudinal evaluation of mild cognitive impairment. *NeuroImage*, *84*, 443–452.

Zeigler, D. W., Wang, C. C., Yoast, R. A., Dickinson, B. D., McCaffree, M. A., Robinowitz, C. B., … Assoc, A. M. (2005). The neurocognitive effects of alcohol on adolescents and college students. *Preventive Medicine*, *40*, 23–32.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*, 297–301.

Zhao, X. S., Lynch, J. G., & Chen, Q. M. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*, 197–206.

Zuo, X. N., & Xing, X. X. (2014). Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: A systems neuroscience perspective. *Neuroscience & Biobehavioral Reviews*, *45C*, 100–118.

List of publications

Peters, S., Peper, J.S., Van Duijvenvoorde, A.C.K., Braams, B.R. & Crone, E.A. (in revision, 2015). Amygdala-orbitofrontal connectivity predicts alcohol use two years later: A longitudinal neuroimaging study on alcohol use in adolescence.

Peters, S., Van der Meulen, M., Zanolie, C.K.K. & Crone, E.A. (in revision, 2015). Predicting reading and mathematics from neural activity for feedback learning: A longitudinal study.

Peters, S., Van Duijvenvoorde, A. C. K., Koolschijn, P.C.M.P. & Crone, E.A. (in revision, 2015). Longitudinal development of neural activity in the frontoparietal network: contributions of age, performance, working memory and brain structure.

Van Duijvenvoorde, A. C. K., Achterberg, M., Braams, B. R., **Peters, S.**, & Crone, E. A. (in press). Testing a dual-systems model of adolescent brain development using resting-state connectivity analyses. *NeuroImage*.

Braams, B.R., Peper, J.S., Van der Heide, D., **Peters, S**., & Crone, E.A. (in revision, 2015). Nucleus accumbens response to rewards and testosterone levels are related to alcohol use in adolescents.

Peters, S., Jolles, D. J., van Duijvenvoorde, A. C. K., Crone, E. A., & Peper, J. S. (2015). The link between testosterone and amygdala-orbitofrontal cortex connectivity in adolescent alcohol use. *Psychoneuroendocrinology*, 53, 117-126.

Braams, B.R., **Peters, S.**, Peper, J.S., Güroglu, B., Crone, E.A. (2014). Gambling for self, friends, and antagonists: Differential contributions of affective and social brain regions on adolescent reward processing. *Neuroimage,* 100, 281-289.

Peters, S., Braams, B.R., Raijmakers, M.E., Koolschijn, P.C.M.P.*, Crone, E.A* (2014). The neural coding of feedback learning across child and adolescent development. *Journal of Cognitive Neuroscience*, 26, 1705-1720. *shared last authorship.

Peters, S., & Crone, E.A. (2014). Cognitive flexibility in childhood and adolescence. In J. A. Grange, & G. Houghton (Eds.), Task Switching and Cognitive Control. Oxford University Press, USA.

Peters, S., Koolschijn, P.C.M.P., Crone, E.A., van Duijvenvoorde, A.C.K., Raijmakers, M.E.J. (2014). Strategies influence neural activity for feedback learning across child and adolescent development. *Neuropsychologia*, 62, 365-374.

Klapwijk, E. T., **Peters, S.**, Vermeiren, R. R., & Lelieveld, G. J. (2013). Emotional reactions of peers influence decisions about fairness in adolescence. *Frontiers in Human Neuroscience*, 7, 745.

Peters, S., Cleare, A. J., Papadopoulos, A., & Fu, C. H. (2011). Cortisol responses to serial MRI scans in healthy adults and in depression. *Psychoneuroendocrinology*, 36, 737-741.

Curriculum Vitae

Sabine Peters was born on July 12th 1987 in Leiderdorp, the Netherlands. After graduating from secondary school (Stedelijk Gymnasium Leiden), Sabine obtained her Bachelor's degree in Psychology (cum laude) in 2008 and her Research Master's degree in Brain and Cognitive Sciences (cum laude) in 2010 from the University of Amsterdam. During her studies, she worked as an fMRI assistant in the Cognitive Neuroscience Group, where she gained experience with many different techniques and research topics. Sabine also completed two research internships in which she investigated neural aspects of sleep and insomnia (Netherlands Institute for Neuroscience) and hippocampal volume and cortisol levels in depressed individuals (King's College London). In 2011, she started her PhD project at the Brain & Development Lab in Leiden University under supervision of Eveline Crone. Sabine explored cognitive and affective aspects of adolescent development in relation to brain maturation and sex hormones. She will continue her work on adolescent brain development as a post-doctoral researcher in the Brain & Development Lab in Leiden.