
Learning and Plasticity in Adolescence

Thesis submitted for the degree of
Doctor of Philosophy

Delia Ute Dorothea Fuhrmann

University College London
Institute of Cognitive Neuroscience

Declaration

I, Delia Ute Dorothea Fuhrmann, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: [REMOVED FROM E-THESIS DEPOSIT]

Date: 28/11/2017

Acknowledgements

The work presented in this thesis would not have been possible without Sarah-Jayne Blakemore's expert supervision. Her insightful guidance, dedication and optimism have been incredibly inspiring and provided the foundation for my research. I would also like to thank Maarten Speekenbrink, my secondary supervisor, for his discerning advice and assistance in all things statistics.

This work was completed in close collaboration with the following, extremely talented, scientists: Caroline Casey, Kathrin Cohen Kadosh, Lisa Knoll, Jovita Leung, Ashok Sakhardande, Susanne Schweizer and Fabian Stamp. I would also like to thank all my other past and present colleagues at UCL and Birkbeck for their advice, help and inspiration, including Annie Brookman-Byrne, Iroise Dumontheil, Lucy Foulkes, Emily Garrett, Anne-Lise Goddings, Cait Griffin, Emma Kilford, Lucía Magis-Weinberg, Kate Mills, Camilla Nord, Stefano Palminteri, Cat Sebastian, Geert-Jan Will and Laura Wolf.

I would also like to thank my family and friends, particularly David Christensen, Herbert Fuhrmann and Dorit Neumann, for much support and encouragement.

Last but not least, I am immensely grateful to all the teachers and individuals who have taken part in our studies; for volunteering their time, minds and brains to science.

My PhD was generously funded by the UCL Department of Psychology and Language Science, the Cusanuswerk and a grant to Sarah-Jayne Blakemore by the Jacob's Foundation. My research projects were also realized with grants to Sarah-Jayne Blakemore by the Royal Society, Nuffield Foundation and Jacob's Foundation.

Abstract

Adolescence is the period of life between puberty and relative independence. It is a time during which the human brain undergoes protracted changes - particularly in the frontal, parietal and temporal cortices. These changes have been linked to improvements in cognitive performance; and are thought to render adolescence a period of relatively high levels of plasticity, during which the environment has a heightened impact on brain development and behaviour. This thesis investigates learning and plasticity in adolescence in four experimental studies. Study 1 examined age differences in face cognition, a key component of social cognition, by testing face perception and face memory performance in 661 participants aged 11 - 33. Study 2 tested whether the effects of social exclusion are age-dependent by measuring cognitive performance after social exclusion in 99 participants between ages 10 - 38. For Study 3, 663 participants aged 11 - 33 were asked to complete 20 days of cognitive training to probe whether the effects of cognitive training are also age-dependent. Study 4 investigated the neural correlates of academic diligence in 40 girls aged 14 - 15, using functional and structural neuroimaging. The research in this thesis demonstrates protracted development of cognitive functions in adolescence, consistent with previous studies. It highlights adolescence as a window of opportunity for learning but also as a vulnerable phase during which the brain is particularly susceptible to harmful effects of social exclusion. Finally, it highlights that individual variability in self-control and underlying frontal systems may be related to academic diligence, and thus educational outcomes.

Table of Contents

Declaration	2
Acknowledgements	3
Abstract	4
Table of Contents.....	5
List of Figures	10
List of Tables	11
Abbreviations.....	12
1. Chapter 1: Introduction	13
1.1. Adolescent Development.....	14
1.1.1. Brain development	14
1.1.2. Cognitive development.....	22
1.1.3. Social development.....	27
1.2. Plasticity and Sensitive Periods of Development.....	34
1.2.1. Evolutionary perspectives on sensitive periods	35
1.2.2. Early sensitive periods	38
1.3. Adolescence as a Sensitive Period of Development.....	40
1.3.1. What evidence would be consistent with adolescence being a sensitive period?	41
1.3.2. Adolescence as a sensitive period for memory and other complex cognitive skills	43
1.3.3. Adolescence as a sensitive period for the effects of stress.....	45
1.3.4. Adolescence as a sensitive period for the effects of drug use	49
1.4. Learning, Plasticity and Education.....	52
2. Chapter 2: Design Issues in Developmental Studies.....	55
2.1. Types of Developmental Designs	55
2.1.1. Cross-sectional designs.....	56
2.1.2. Longitudinal designs	56

2.1.3.	Combination designs	57
2.2.	Validity Issues	58
2.2.1.	Instrumentation effects.....	58
2.2.2.	Missing data and selective attrition.....	60
2.3.	Design Issues in Training Studies.....	63
2.3.1.	Temporal design.....	63
2.3.2.	Adaptive designs	64
2.3.3.	Active vs. passive control groups	64
2.3.4.	Randomized-control studies.....	65
2.3.5.	Transfer tasks	66
2.4.	Statistical Design and Analysis	67
2.4.1.	Generalized Linear Mixed Models (GLMMs)	67
2.4.2.	Modelling age.....	70
3.	Chapter 3: Perception and Recognition of Faces in Adolescence	73
3.1.	Introduction	74
3.2.	Methods.....	76
3.2.1.	Participants.....	76
3.2.2.	Experimental design	78
3.2.3.	Testing procedure	79
3.2.4.	Face memory task.....	80
3.2.5.	Face perception task.....	81
3.2.6.	Statistical analysis	84
3.3.	Results.....	85
3.3.1.	The development of face memory and face perception	85
3.3.2.	Gender differences in face memory and face perception.....	87
3.3.3.	Development of face aspect perception	87
3.4.	Discussion	91

4. Chapter 4: Cognitive Performance after Social Exclusion in Adolescence	95
4.1. Introduction	95
4.2. Methods.....	98
4.2.1. Participants.....	98
4.2.2. Materials.....	99
4.2.3. Procedure	103
4.2.4. Design and analysis.....	103
4.3. Results.....	104
4.3.1. Working memory performance	104
4.3.2. Mood ratings	110
4.4. Discussion	111
5. Chapter 5: Age-Dependent Effects of Cognitive Training	117
5.1. Introduction	117
5.2. Methods.....	121
5.2.1. Participants.....	121
5.2.2. Experimental design	123
5.2.3. Testing procedure	123
5.2.4. Training procedure	124
5.2.5. Numerosity discrimination task	125
5.2.6. Relational reasoning task.....	127
5.2.7. Face perception task.....	128
5.2.8. Backward digit span task	130
5.2.9. Face memory task.....	131
5.2.10. Statistical analysis	132
5.3. Results.....	133
5.3.1. General training effects	133
5.3.2. Age-dependent training effects.....	138

5.3.3.	Transfer effects	141
5.3.4.	Potential confounds.....	143
5.4.	Discussion	149
6.	Chapter 6: The Neurocognitive Correlates of Academic Diligence	154
6.1.	Introduction	154
6.2.	Methods.....	158
6.2.1.	Participants.....	158
6.2.2.	Behavioural task.....	159
6.2.3.	fMRI task.....	160
6.2.4.	Imaging data acquisition and pre-processing.....	163
6.2.5.	Functional analyses.....	164
6.2.6.	Connectivity analyses.....	165
6.2.7.	Structural analysis	165
6.2.8.	Regression Models.....	166
6.3.	Results.....	167
6.3.1.	Behaviour in the Academic Diligence Task.....	167
6.3.2.	Behaviour in the go-no-go task	167
6.3.3.	fMRI results	168
6.3.4.	Connectivity results	171
6.3.5.	Structural results.....	171
6.4.	Discussion	171
7.	Chapter 7: Discussion	176
7.1.	Summary of Findings.....	177
7.1.1.	Are there age-related differences in face cognition between adolescence and adulthood?	177
7.1.2.	Do the effects of social exclusion on cognitive performance differ between age groups?	179

7.1.3.	Do some age groups benefit more from cognitive training than others?.....	181
7.1.4.	What are the neurocognitive correlates of academic diligence?	183
7.2.	Implications for Policy and Practise.....	185
7.2.1.	Adolescence as a key period for social development	185
7.2.2.	Cognitive ability is not fixed.....	187
7.2.3.	Earlier is not always better for learning.....	188
7.2.4.	Transfer effects are limited	190
7.2.5.	Individual differences matter	191
7.2.6.	Specificity of developmental trajectories	194
7.3.	Methodological Considerations.....	195
7.3.1.	Cross-sectional versus longitudinal designs	195
7.3.2.	Gender differences	196
7.3.3.	The need for real-life measures	197
7.3.4.	Modelling development	199
7.4.	Directions for Future Research.....	200
7.4.1.	Lifespan perspectives.....	201
7.4.2.	Learning beyond plasticity.....	202
7.4.3.	Latent variable and multivariate models	205
7.5.	Conclusion	207
8.	References.....	208
	Appendix 2.1: R script	261

List of Figures ¹

Figure 1.1. Changes in Cortical Volume over Development.	16
Figure 1.2. The Dual-Systems Hypothesis.	18
Figure 1.3. Raven’s Progressive Matrices Task.	25
Figure 1.4. Social Influence on Risk Perception.	33
Figure 1.5. Models of Plasticity in Adolescence.	42
Figure 1.6. Age of Onset of Psychiatric Disorders.	46
Figure 1.7. Fear Extinction Learning in Mice and Humans across Development.	48
Figure 2.1. The Regression Approach.	68
Figure 3.1. Cambridge Face Memory Task.	81
Figure 3.2. Face Perception Task.	83
Figure 3.3. Face Cognition Performance.	86
Figure 3.4. Face Perception Performance by Face Aspect.	90
Figure 4.1. Cyberball Game.	100
Figure 4.2. N-back Accuracy after Inclusion and Exclusion.	106
Figure 4.3. Dot-Matrix Response Times after Inclusion and Exclusion.	110
Figure 4.4. Mood Ratings after Inclusion and Exclusion.	111
Figure 5.1. Timeline of the Study.	120
Figure 5.2. Performance by Training Group.	135
Figure 5.3. Performance by Age Group.	139
Figure 5.4. Training Gains Compared between Age Groups.	140
Figure 6.1. The Academic Diligence Task and Emotional Go-No-Go Task.	162
Figure 6.2. Ventral-Striatum Activation during the Emotional Go-No-Go Task.	170
Figure 6.3. Inferior Frontal Gyrus ROI and Correlation with Diligence.	170
Figure 7.1. Individual Differences in Brain Maturation.	193
Figure 7.2. A Model of Lifespan Changes in Cognitive Performance and Learning.	201

¹ Please note that all figures without specific permissions information are original.

List of Tables ²

Table 3.1. Demographic Information.....	77
Table 3.2. Comparisons of Relational Reasoning Accuracy between Age Groups.....	78
Table 3.3. Face Cognition Performance Overall	87
Table 3.4. Comparisons of Face Perception Accuracy between Age Groups and Face Aspects	89
Table 3.5. Comparisons of Face Perception Speed between Age Groups and Face Aspects.	90
Table 4.1. Participant Characteristics	99
Table 4.2. Overall Performance in the N-Back and Dot-Matrix Task	105
Table 4.3. Cognitive Performance and Mood Compared between Exclusion and Inclusion within Age Groups	107
Table 4.4. Cognitive Performance and Mood Compared between Exclusion and Inclusion between Age Groups	108
Table 4.5. Mean Accuracy in the 0-Back Task.....	108
Table 5.1. Participant Numbers	122
Table 5.2. Numerosity Discrimination Performance Change after Training	136
Table 5.3. Relational Reasoning Performance Change after Training.....	137
Table 5.4. Face Perception Performance Change after Training	138
Table 5.5. Change in Backward Digit Span Performance after Training.....	142
Table 5.6. Change in Cambridge Face Memory Task Performance after Training.....	143
Table 5.7. Confounds with Age.....	144
Table 5.8. Contrasts that Became Non-Significant after Excluding Adults' Data	146
Table 5.9. Contrasts that Became Non-Significant after Including Covariates for Group Size and Spacing between Test Sessions	147
Table 5.10. Contrasts that Became Non-Significant after Excluding Data from Individuals whose Test Sessions Were Split over Several Days	149
Table 6.1. Participant Characteristics	158
Table 6.2. Results of the Cluster-Corrected Whole-Brain Analysis during No-Go Compared to Go Trials.	168
Table 6.3. Results of Models Predicting Functional Activation of the IFG, DS and VS	169
Table 6.4. Results of a Model Predicting Connectivity between the IFG and DS	171

² Please note that all tables without specific permissions information are original.

Abbreviations

AFC	alternative forced choice
CFMT	Cambridge Face Memory Task
DS	dorsal striatum
FD	frame-wise displacement
FDR	false discovery rate
FP	face perception
GAM	Generalized Additive Model
GLM	General Linear Model
GLMM	Generalized Linear Mixed Model
HDAC	histone-deacetylase
IFG	inferior frontal gyrus
IQR	interquartile range
LMM	Linear Mixed Effects Model
MAR	missing at random
MCAR	missing completely at random
ND	numerosity discrimination
NMAR	not missing at random
PND	post-natal day
PPI	psycho-physiological interaction
RCT	randomized-control trial
ROI	region of interest
RR	relational reasoning
SEM	structural equation modelling
SES	socio-economic status
T1	test session one
T2	test session two
T3	test session three
TIV	total intracranial volume
UCL	University College London
VS	ventral striatum

1. Chapter 1: Introduction ³

Educational policy tends to focus on early childhood. However, research over the past 20 years has demonstrated that the human brain and mind undergo protracted changes beyond childhood. Adolescence, in particular, has been shown to be characterized by continued changes in brain structure; accompanied by the development of high-level cognitive functions relevant to education. These changes in brain structure have also been linked to the protracted development of social skills necessary to navigate life inside and outside the classroom. These findings have led to the suggestion that adolescence is a period of relatively high levels of plasticity, during which time the environment has a heightened impact on brain development and behaviour. This chapter reviews evidence for protracted development in adolescence; extracts general characteristics of sensitive periods from well-researched sensitive periods in early development and discusses evidence for high levels of plasticity of cognitive functions relevant to education during adolescence.

³ Parts of this chapter have been published as:

Fuhrmann, D., Knoll, L.J., Blakemore, S.-J. (2015). Adolescence as a sensitive period of brain development. *Trends in Cognitive Sciences*, 19 (10), 558-566. doi:10.1016/j.tics.2015.07.008

Fuhrmann, D.* , Knoll, L.J.* , Sakhardande, A., Stamp, F., Speekenbrink, M. & Blakemore, S-J. (2016). A window of opportunity for cognitive training in adolescence. *Psychological Science*, 27(12), 1620-1631. doi: 10.1177/0956797616671327 *Joint first authors.

Fuhrmann, D., Knoll, L.J., Sakhardande, A., Speekenbrink, M., Cohen Kadosh, K. & Blakemore, S-J. (2016). Perception and recognition of faces in adolescence. *Scientific Reports*, 6(33497), doi:10.1038/srep33497

1.1. Adolescent Development

Adolescence is the period of life between the onset of puberty and the point at which we attain a stable, independent role in society (Damon, 2004). It is a time of protracted changes in brain structure across the cortex, but particularly in the frontal, parietal and temporal lobe (Giedd et al., 1999; Tamnes et al., 2017; Tamnes, Walhovd, Dale, et al., 2013). Concomitantly, processes requiring high levels of cognitive control continue to develop during adolescence, leading to improvements in cognitive skills such as reasoning and memory (Bazargani, Hillebrandt, Christoff, & Dumontheil, 2014; Dumontheil, Houlton, Christoff, & Blakemore, 2010; Tamnes, Walhovd, Grydeland, et al., 2013). Adolescence is also a time of social maturation during which peers becoming increasingly important (Blakemore, 2008; Blakemore & Mills, 2014; Steinberg, 2008).

1.1.1. Brain development

Until 20 years ago, it was thought that the brain had more or less reached maturity after childhood. This view was based on post-mortem studies, showing that total brain volume increases rapidly during the first years of life, and levels off after 5 or 6 years of age (Dobbing & Sands, 1973; Giedd et al., 1999; Reiss, Abrams, Singer, Ross, & Denckla, 1996). However, advances in in-vivo neuroimaging techniques over the past two decades have resulted in a wealth of data showing that regional structural changes continue to occur throughout childhood, adolescence and into adulthood (Giedd et al., 1999; Lenroot & Giedd, 2006; Tamnes et al., 2017; Tamnes, Walhovd, Dale, et al., 2013).

Protracted cortical development

Brain maturation during adolescence is characterized by an overall increase in white matter (Brain Development Cooperative Group, 2012; Giedd et al., 1999) and decrease in grey matter volume and cortical thickness (Tamnes et al., 2017). White matter has been shown to increase linearly up until mid-to-late adolescence. It still increases in volume thereafter but the rate of change decelerates (Mills et al., 2016). Grey matter volume, in contrast, is highest around 8 years of age and decreases throughout adolescence (Mills et al., 2016).

The increase in white matter during adolescence is thought to largely reflect an increase in myelination and axon diameter (Grydeland, Walhovd, Tamnes, Westlye, & Fjell, 2013; Miller et al., 2012). Myelin increases the speed of signal transmission along the axon and regulates timing of information transmission (Fields, 2014; Lillard & Erisir, 2011).

Grey matter, in contrast, consists mainly of neural soma, dendrites and synapses. It has been posited that the decrease in grey matter during adolescence is due to synaptic pruning, that is, the active loss of synapses that are not used (Giorgio et al., 2010; Selemon, 2013). However, synaptic boutons only comprise a small fraction of cortical volume - 1.5% in macaque monkeys (Bourgeois & Rakic, 1993). It therefore seems unlikely that grey matter volume reduction, which amounts to approximately 17% in the prefrontal cortex between late childhood and early adulthood (Mills, Goddings, Clasen, Giedd, & Blakemore, 2014), is solely due to pruning. Other potential cellular mechanisms underlying grey matter volume reduction include increasing myelination encroaching on grey matter (Mills & Tamnes, 2014; Paus, 2005).

Changes in brain structure during adolescence are not uniform across the cortex. More posterior regions of the cortex mature before more anterior regions (Figure 1.1) (Tamnes et al., 2017; Tamnes, Walhovd, Dale, et al., 2013). In late childhood, volume changes are most prominent in the occipital and parietal lobes, while in late adolescence, volume changes are more pronounced in the frontal lobes and inferior temporal and parietal regions (Tamnes, Walhovd, Dale, et al., 2013). The latter brain regions are implicated in a number of high-level cognitive functions, including adaptive cognitive performance and social cognition (Blakemore, 2008; Crone & Dahl, 2012; Milner, 1963; Tamnes, Walhovd, Grydeland, et al., 2013).

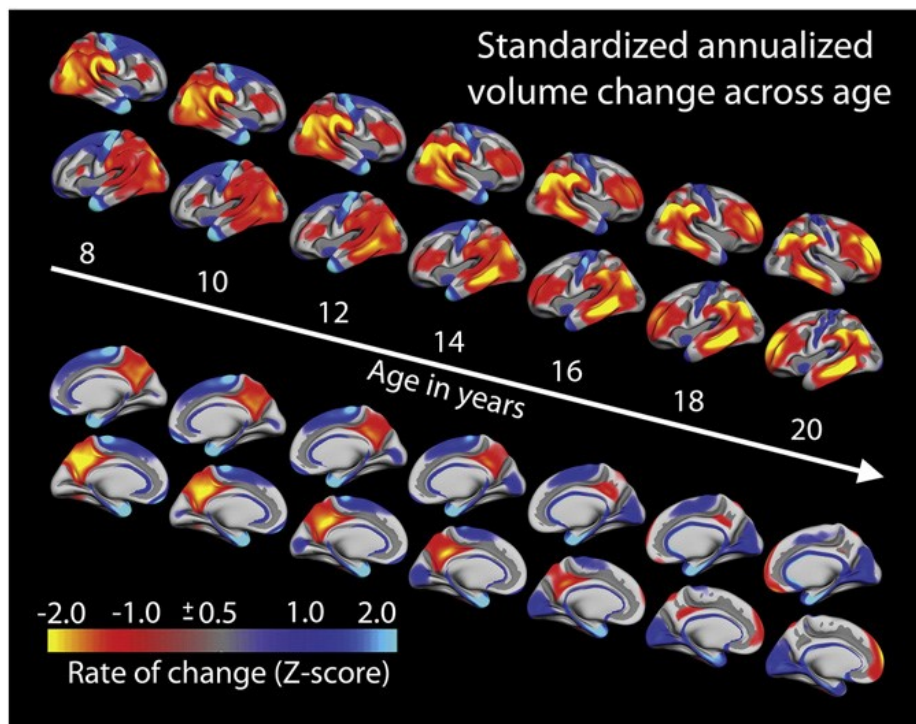


Figure 1.1. Changes in Cortical Volume over Development. Red-yellow regions show the largest reductions in volume, while blue regions show relatively smaller changes in volume (adapted from Tamnes, Walhovd, Dale, et al., 2013). Permission to reproduce this figure has been granted by Elsevier.

Cortical versus subcortical areas: The dual-systems hypothesis

Among the late-maturing cortical brain regions, the frontal lobe has been of particular interest to developmental science because its development is closely linked to improvements in executive functions and self-control (Blakemore & Robbins, 2012; Crone & Steinbeis, 2017; Tamnes, Walhovd, Grydeland, et al., 2013).

It has been proposed that, in contrast to frontal regions, subcortical regions mature earlier in adolescence (Casey, Getz, & Galván, 2008; Steinberg, 2008). This is thought to be particularly true for regions involved in reward processing, such as the ventral striatum and amygdala (Ernst et al., 2005; van Leijenhorst et al., 2010). This maturational mismatch between cortical control and subcortical reward regions has been linked to increased sensation-seeking in adolescence (Figure 1.2) (Steinberg et al., 2017).

This dual-systems hypothesis is one of the most influential models of adolescent development (Shulman et al., 2016; Strang, Chein, & Steinberg, 2013) and has been used to explain phenomena such as increased impulsivity and risk-taking in adolescence (Casey et al., 2008; Steinberg, 2008; van den Bos, Rodriguez, Schweitzer, & McClure, 2015). Recent reviews however, have critiqued the dual-systems hypothesis as overly simplistic and have highlighted heterogeneity in the development of frontal and striatal structures, task-related functional activation and individual trajectories (Casey, Galván, & Somerville, 2016; Crone & Dahl, 2012; Pfeifer & Allen, 2012; Romer, Reyna, & Satterthwaite, 2017).

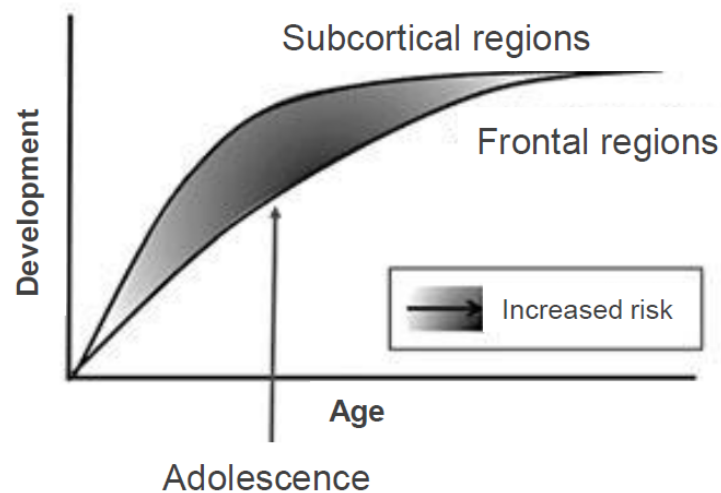


Figure 1.2. *The Dual-Systems Hypothesis*. This theoretical model illustrates the mismatch in brain maturation during adolescence; with subcortical regions being relatively mature during adolescence, whereas the frontal lobe does not reach similar levels of maturity until adulthood. The gap in maturity (shaded) is hypothesised to increase the risk for sensation-seeking behaviours during adolescence. Adapted from Mills et al. (2014). Permission to reproduce this figure has been granted by Karger.

Consistent with the dual-systems hypothesis, structural imaging studies have shown that the average rate of volume change during adolescence is higher in cortical than in subcortical regions (Brain Development Cooperative Group, 2012). Annual volume change in the cortex exceeds 1% whilst subcortical regions change approximately 0.5% per year during adolescence (Tamnes, Walhovd, Dale, et al., 2013).

Individual differences in maturational trajectories are pervasive however. A recent longitudinal study directly compared maturation of three relevant brain regions: the prefrontal cortex as well as two subcortical areas - the nucleus accumbens, a

part of the ventral striatum, and the amygdala (Mills et al., 2014). The volume of each of these three structures was measured on at least three occasions for each of the 33 participants in the study. Maturation was measured as the current regional volume compared to the regional volume at the last point of measurement. At the group level, the results were consistent with the dual-systems hypothesis. Volume in the prefrontal cortex changed by 17% from 7 - 30 years, while volume changed by only 7% each in the two sub-cortical regions. However, differences in individual developmental trajectories were large. 82% of participants presented with a mismatch between the maturation of the prefrontal cortex and the amygdala, while only about half the participants showed a mismatch between the maturity of the prefrontal cortex and the nucleus accumbens. Some participants showed no mismatch at all. There was also no systematic relationship between the individual extent of structural mismatch and self-reported risk-taking (Mills et al., 2014).

In line with the dual-systems hypothesis, functional neuroimaging studies have mainly shown age-related decreases in striatal activation during reward-processing tasks (Crone & Dahl, 2012; Pfeifer & Allen, 2012). However, some studies have found adolescent striatal hypoactivity rather than hyperactivity in response to rewards (Bjork, Smith, Chen, & Hommer, 2010; Geier, Terwilliger, Teslovich, Velanova, & Luna, 2010). Bjork and colleagues, for example, found that the nucleus accumbens was hypoactive or similarly activated in adolescents (aged 12 - 17) compared to adults (aged 22 - 42) when anticipating or receiving monetary rewards (Bjork et al., 2010). What is more, Pfeifer and colleagues found that increases in ventral striatum activity in response to affective facial displays

between ages 10 - 13 longitudinally predicted *less* susceptibility to risk-taking (Pfeifer et al., 2011). Some of these inconsistencies in striatal activation between studies may be due to the fact that different studies tend to assess different stages of reward processing, i.e. reward anticipation, reward receipt, reward assessment, etc. Nonetheless, the functional significance of striatal hyperactivity in adolescence remains unclear (Pfeifer & Allen, 2012).

Regarding frontal regions, many developmental neuroimaging studies have found age-related increases in frontal activation in cognitive control tasks (Crone & Dahl, 2012; Pfeifer & Allen, 2012). This has been interpreted as indicative of frontal regions coming increasingly online during adolescence and taken as evidence in favour of the dual-systems hypothesis (Shulman et al., 2016; Strang et al., 2013). However, patterns of activation are strongly task-dependent (Crone & Dahl, 2012; Pfeifer & Allen, 2012). Critically, the most complex executive function tasks, which require self-control as well as performance monitoring, were often not associated with clear differences in frontal activation between adolescents and adults (Siegel et al., 2014; van den Bos, Guroglu, van den Bulk, Rombouts, & Crone, 2009).

Some studies have even found that adolescents and young adults can be more self-controlled than older adults. A recent longitudinal study with 192 participants aged 8 - 26 showed that the ability to delay gratification follows a quadratic trend with a peak in the early twenties, rather than a dip in adolescence (Achterberg, Peper, van Duijvenvoorde, Mandl, & Crone, 2016). This challenges the idea of pervasive frontal immaturity during adolescence and indicates that self-control systems may already be online during this time of life (Crone & Dahl, 2012).

Overall, the existing literature thus highlights heterogeneity in the maturation of frontal and striatal structures and task-related functional activation, as well as individual differences in adolescent behaviour and brain development. There is also little and inconsistent data on how imaging data relates to real-world behaviour and individual differences thereof (Pfeifer & Allen, 2012).

The dual-systems hypothesis and education

Many studies investigating the dual-systems theory have focussed on risk-taking behaviours in adolescence (e.g. Braams, van Duijvenvoorde, Peper, & Crone, 2015; Casey et al., 2008; Steinberg, 2008). However, reward processing and self-control also affect many other phenomena including educational outcomes (Duckworth & Steinberg, 2015; Steinbeis & Crone, 2016). Personality traits such as diligence, conscientiousness or grit, all of which describe the ability to regulate behaviour in the service of goals, have also been shown to be related to educational attainment (Credé, Tynan, & Harms, 2016; Duckworth & Gross, 2014; Duckworth, Peterson, Matthews, & Kelly, 2007; Galla et al., 2014). The contribution of diligence to educational outcomes is thought to be dissociable from, and sometimes orthogonal to, IQ (Credé et al., 2016; Duckworth et al., 2007).

It has been proposed that diligence is the product of conflicting psychological processes – the exercise of will and the drive to seek immediate gratification (Duckworth & Steinberg, 2015). This rationale is similar to that of the dual-systems hypothesis, and therefore diligence might be hypothesized to correlate with front-striatal structure function. However, only a handful of studies have

investigated neural correlates of diligence and related constructs (DeYoung et al., 2010; Forbes et al., 2014; C. A. Myers, Wang, Black, Bugescu, & Hoefft, 2016; Nemmi, Nymberg, Helander, & Klingberg, 2016; S. Wang et al., 2016) and have produced results only partly consistent with a dual-systems account of diligence. Nemmi and colleagues (2016), for instance, found that grit correlates with nucleus accumbens grey matter density but not cortical thickness in the frontal lobe. There is also still a need for studies that combine structural, functional and connectivity data to provide a more holistic understanding of adolescent self-control (Kilford, Garrett, & Blakemore, 2016). This gap in the literature will be addressed in Chapter 6.

1.1.2. Cognitive development

Changes in brain structure and function during adolescence are accompanied by protracted changes in cognitive functions. Piaget conceptualized adolescence as the 'formal operational' period of development during which we increasingly rely on abstract thought (Inhelder & Piaget, 1958). Recent empirical evidence supports this proposition. Adolescence has been shown to be a time during which many cognitive skills relevant to education undergo rapid development. These skills include planning (Luciana, Collins, Olson, & Schissel, 2009), attention (Velanova, Wheeler, & Luna, 2008) and arithmetic (Rivera, Reiss, Eckert, & Menon, 2005). Here, the development of three non-social cognitive skills, investigated in experimental Chapters 3 - 6 of this thesis, is discussed: (I) working memory, (II) relational reasoning and (III) enumeration.

Working memory

Working memory describes the ability to store and manipulate information for ongoing cognitive processing (Baddeley & Hitch, 1974). Working memory predicts both fluid intelligence (Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004) and academic performance (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Gathercole, Brown, & Pickering, 2003).

The most established model of working memory assumes that working memory consists of a central executive managing two storage systems: a phonological loop and the visuo-spatial sketchpad (Baddeley, 2003; Baddeley & Hitch, 1974). The capacity of these two storage systems critically limits working memory performance and can be assessed using tests of verbal and visuo-spatial working memory. Verbal working memory is known to predict reading comprehension (Swanson, Howard, & Saez, 2006) and can be assessed, for example, by asking participants to memorize and repeat digits (de Haan, 2014). Visuo-spatial working memory is more predictive of mathematical performance than reading comprehension (Holmes & Adams, 2006) and can be measured with tasks that require participants to memorize and repeat spatial sequences (de Haan, 2014).

Each of these two storage systems can be assessed with simpler tasks, in which information need only be maintained over a delay (e.g. forward digit span), or more complex tasks, in which information has to be mentally manipulated (e.g. backward digit span) (de Haan, 2014). Generally speaking, complex working memory tasks with high processing demands that recruit frontal areas of the brain, show particularly protracted development throughout adolescence, while performance on simpler tasks plateaus before age 12 (Conklin, Luciana, Hooper, &

Yarger, 2007; Isbell, Fukuda, Neville, & Vogel, 2015; Luciana, Conklin, Hooper, & Yarger, 2005; Tamnes, Walhovd, Grydeland, et al., 2013). Luciana and colleagues, for instance, showed that simple aspects of visuo-spatial working memory, such as delayed spatial recall, reach maturity around ages 11 - 12 (Luciana et al., 2005). More complex working memory abilities, such as strategic self-guided spatial search, in contrast, continue to improve at least up to 16 - 17 years of age (Luciana et al., 2005) and possibly even up to the mid-twenties (Murre, Janssen, Rouw, & Meeter, 2013).

Relational reasoning

Relational reasoning is the ability to detect abstract relationships between groups of items (Krawczyk, 2012). Relational reasoning skills correlate with mathematics performance (Mackey, Whitaker, & Bunge, 2012), and are often assessed in tests of fluid intelligence, a strong predictor of educational outcomes (Chuderski, 2014).

Relational reasoning can be reliably scored using tests like Raven's Progressive Matrices Test (Raven, 1941). Raven's is a visual task in which a stimulus has to be identified, that completes a pattern (Figure 1.3). Difficulty is scaled by the number of dimensions (e.g. horizontal, vertical, colour, shape) participants need to take into account to arrive at the correct solution (Christoff et al., 2001). For instance, zero-relational problems are the easiest to solve and involve only simple visual matching, while one-relational problems require pattern matching along one dimension.

Relational reasoning performance has been shown to improve between childhood and adulthood (Crone et al., 2009; Richland, Morrison, & Holyoak, 2006). For instance, young adults aged 18 – 25 perform better than children aged 8 - 12 on two-relational problems but similarly on zero- and one-relational problems (Crone et al., 2009). These changes in relational reasoning performance during adolescence have been linked to the protracted development of frontal brain regions, particularly the rostralateral prefrontal cortex (Bazargani et al., 2014; Crone et al., 2009; Dumontheil, Houlton, et al., 2010).

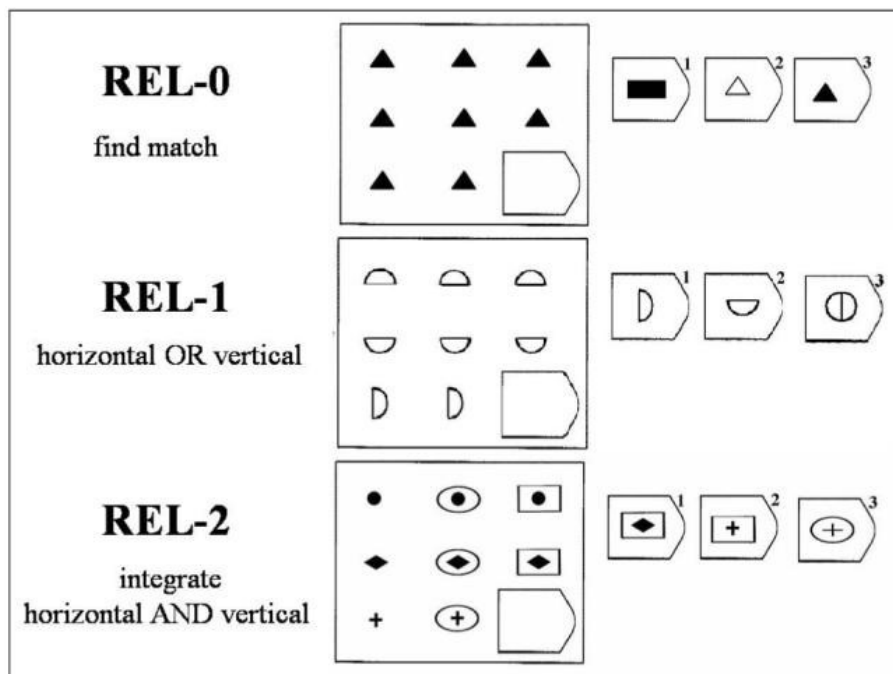


Figure 1.3. Raven's Progressive Matrices Task. The panel shows examples of a zero-relational problem (REL-0), a one-relational problem (REL-1) and a two-relational problem (REL-2). Participants are instructed to indicate the correct solution out of three options. Adapted from Crone et al. (2009). Permission to reproduce this figure has been granted by John Wiley and Sons.

Some studies have found non-linear trajectories with a dip in relational reasoning between ages 11 - 17, compared to younger and older participants (Dumontheil, Houlton, et al., 2010). Such non-linear changes may relate to the onset of puberty and can be observed in other domains as well, including in face processing and on match-to-sample tasks (Carey, Diamond, & Woods, 1980; Dumontheil, Houlton, et al., 2010; McGivern, Andersen, Byrd, Mutter, & Reilly, 2002).

Numerosity discrimination

A foundational skill in mathematics is numerosity discrimination, the ability to represent approximate numbers and compare quantities (e.g. which group of icons is bigger) (Piazza, 2010). This is sometimes also called the 'approximate number system' (Feigenson, Dehaene, & Spelke, 2004) or 'number acuity' (Halberda, Mazocco, & Feigenson, 2008).

Some sense of numerosity is present already in infancy (Piazza, 2010). Newborns can discriminate between quantities differing by a 1:3 ratio, 6-month old infants discriminate quantities differing by 2:1, and 9 month olds succeed at 2:3 ratios (Piazza, 2010). It has been proposed that symbolic mathematical skills such as learning numbers is achieved by mapping number words or digits onto pre-existing, approximate, quantitative representations. In this model, the approximate number system is a pre-requisite for acquiring symbolic mathematics skills (Wynn, 1992).

This view was supported by studies showing that numerosity discrimination correlates with and longitudinally predicts mathematics performance in children

and adults (Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Halberda et al., 2008; Starr, Libertus, & Brannon, 2013). However, the relationship may become bidirectional relatively early in development. Symbolic number representation abilities have been shown to longitudinally predict the refinement of the approximate number system in 4 - 5 year olds (Mussolin, Nys, Content, & Leybaert, 2014).

Number acuity undergoes protracted development past childhood. While earlier computational modelling suggested that number acuity is mature from the pre-teen years (Halberda & Feigenson, 2008), a more recent large-scale data ($N > 10,000$) showed that number acuity improves throughout the school-age years, peaking at around 30 years of age (Halberda, Ly, Wilmer, Naiman, & Germine, 2012).

1.1.3. Social development

In addition to the development and refinement of non-social cognitive skills, social cognition undergoes major changes in adolescence. A key developmental task for adolescents is transitioning from the relative dependence on caregivers during childhood to the relative social independence of adulthood. This requires the acquisition of social skills necessary for navigating life inside and outside the classroom (Blakemore, 2010).

This process is accompanied by marked changes in social cognition (Blakemore, 2008; Blakemore & Mills, 2014). Adolescence is characterised, for example, by maturation in perspective taking (Dumontheil, Apperly, & Blakemore, 2010;

Dumontheil, Kuster, Apperly, & Blakemore, 2010; Sebastian et al., 2012), emotion processing (Blakemore & Robbins, 2012; Goddings, Burnett Heyes, Bird, Viner, & Blakemore, 2012) and social learning (Cook & Bird, 2011). Here, evidence for developmental changes in two areas of social cognitive development, that are of particular interest for the experiments presented in this thesis, are reviewed: (I) face cognition and (II) peer-influence.

Face cognition

Faces are of unique importance in everyday life. Recognition of faces is fundamental to building and maintaining relationships (Behrmann & Avidan, 2005). Faces also provide social signals such as information about other people's mental states and emotions (Adolphs, 2003) and facilitate communication and social learning (Tomasello & Carpenter, 2007). Because faces are so central to our social life, it has been proposed that face cognition may serve as a model for broader changes in social cognition during adolescence (Scherf, Behrmann, & Dahl, 2012). There is, however, an ongoing controversy as to whether face cognition actually changes qualitatively past childhood (McKone, Crookes, Jeffery, & Dilks, 2012).

Theories of face cognition mainly distinguish between two sub-components: *face memory*, the ability to learn and recognize known faces, and *face perception*, the ability to discriminate facial features and configurations (Dolzycka, Herzmann, Sommer, & Wilhelm, 2014). Face memory and face perception are thought to be face-specific skills that are distinct from other abilities such as object cognition (Wilhelm, Herzmann, Kunina, & Sommer, 2007).

The ability to recognize people from their faces has its origins early in development, potentially even prenatally (Crookes & McKone, 2009; Reid et al., 2017), but face memory expertise follows a protracted course of development thereafter. Face memory has been shown to improve rapidly between 6 and 10 years and then level off, or even dip with the onset of puberty, to then rise again later during the teenage years (Diamond, Carey, & Back, 1983). Other studies have shown linear improvements between childhood and adulthood (Gur et al., 2012; Song, Zhu, Li, Wang, & Liu, 2015).

Proponents of the *late maturation account* argue that these developmental patterns demonstrate that face memory does not mature until at least 10 years of age and that it is likely driven by experience of faces - by one's 'face diet' (Maurer & Mondloch, 2011). Proponents of the *early maturation account*, in contrast, contend that experience has little effect on face memory development and that quantitative improvements in face memory after 3 - 5 years are due to improvements in general cognitive ability (McKone et al., 2012).

Similarly for face perception, early maturation accounts compete with late maturation accounts (Mondloch, Le Grand, & Maurer, 2002; Mondloch, Robbins, & Maurer, 2010). The disagreement here may in part be attributable to the fact that the perception of different face aspects such as identity, expression or gaze, develops at different rates (Cohen Kadosh, 2011). These face aspects are processed in different ways by the face perception network of the brain (Cohen Kadosh, Johnson, Henson, Dick, & Blakemore, 2013; Cohen Kadosh, Walsh, & Cohen Kadosh, 2010), and gaze perception generally matures earlier than identity or expression perception, with identity perception maturing last (Cohen Kadosh,

2011; Cohen Kadosh, Johnson, Henson, et al., 2013). It is thought that this order of maturation reflects the degree to which configural processing is required (Mondloch et al., 2002). Configural processing refers to the processing of the overall layout of the face, rather than features in isolation. This is thought to be a complex cognitive skill that requires much experience (Mondloch et al., 2002).

In order to investigate when in development face cognition matures, Chapter 3 explores age-related differences in the two aspects of face cognition, face memory and face perception. It also explores when different aspects of face perception mature.

Peer influence

Adolescence is thought to be a time during which peers become increasingly important (Crone & Dahl, 2012; Steinberg, 2008). A recent, eight-year longitudinal study investigated developmental changes time allocation in American girls and boys. Lam and colleagues followed children and adolescents aged 8 - 18 and showed that time with same-sex peers peaks around the age of 14, after which time is spent increasingly with opposite-sex peers (Lam, McHale, & Crouter, 2014). The quality of relationships with peers changes during adolescence too. While parents are perceived as the greatest source of support in late childhood (between the ages 9 - 12.5), by mid-adolescence (around 16 years), same-sex friends are the greatest source of support. At age 19, romantic partners are the most important source of support (Furman & Buhrmester, 1992).

Peers also exert a strong influence on opinions and behaviours during adolescence (Blakemore & Mills, 2014). For instance, young adolescents (aged

12 - 14) appear to be particularly susceptible to peer influence on risk perception, compared with other age groups (Figure 1.4) (Knoll, Magis-Weinberg, Speekenbrink, & Blakemore, 2015). Knoll and colleagues measured the degree of social influence on risk perception in different age groups and found that, while other age groups were more influenced by adults' opinions about risk, young adolescents were more influenced by the opinions of other adolescents. Mid-adolescents (aged 15 to 18) showed no difference in social influence between adults' and teenagers' opinions about risks, suggesting that this is a transitional stage in development.

Adolescents and young adults (aged 13 - 22) are also more likely to take driving risks in the presence of peers, whilst adults' (aged 24 and over) driving risks are unaffected by peers (Gardner & Steinberg, 2005). This peer influence effect is not restricted to risk-taking. When adolescents (aged 10 - 18) are observed by a friend rather than an adult, their performance on a reasoning task is reduced. Adults' (aged 21 and over) performance, in comparison, is unaffected by being observed by a friend or an adult (Wolf, Bazargani, Kilford, Dumontheil, & Blakemore, 2015).

Peer influence may be heightened in adolescence because peer acceptance or rejection strongly influences feelings of self-worth (Blakemore & Mills, 2014; Burke, McCormick, Pellis, & Lukkes; O'Brien & Bierman, 1988). Peer-rejection and social exclusion can be simulated experimentally using the Cyberball paradigm (Williams, Cheung, & Choi, 2000). Cyberball is an online ball-tossing game during which participants are ostensibly either included or excluded by two peers. In adults, Cyberball exclusion reliably lowers mood and induces a threat to four

fundamental psychological needs: self-esteem, belonging, control and a sense of meaningful existence (Williams, 2007; Williams et al., 2000). Such effects may be amplified in younger age groups. For example, young adolescent girls (aged 11 - 13) showed a reduction in mood and increase in anxiety after exclusion compared to baseline, whilst mid-adolescents (aged 14 - 15) showed reduced mood only, and adults (aged 22 - 47) showed no changes in either mood or anxiety (Sebastian, Viding, Williams, & Blakemore, 2010). Another study showed that Cyberball exclusion threatened psychological needs in adolescents (aged 13 - 17) and emerging adults (aged 18 - 22) more than it did in older adults (aged 22 to 27; Pharo, Gross, Richardson, & Hayne, 2011).

Social exclusion may affect not only mood and need-threat but also cognitive performance (Baumeister, Twenge, & Nuss, 2002). Studies using the Cyberball paradigm in adults have largely found negative effects of exclusion on cognitive functioning, particularly on executive functions such as inhibitory control and working memory. For instance, Cyberball exclusion is associated with reduced performance in the Flanker task (Themanson, Ball, Khatcherian, & Rosen, 2014) and the anti-saccade task (Jamieson, Harkins, & Williams, 2010) in adults. Cyberball exclusion has also been shown to disrupt cognitive performance in children. Hawes and colleagues showed that social exclusion disrupted cognitive performance in girls, but not boys, aged 8 - 12 (Hawes et al., 2012). To date, however, there is little experimental evidence on how social exclusion affects cognitive performance in adolescence. This question is explored in Chapter 4.

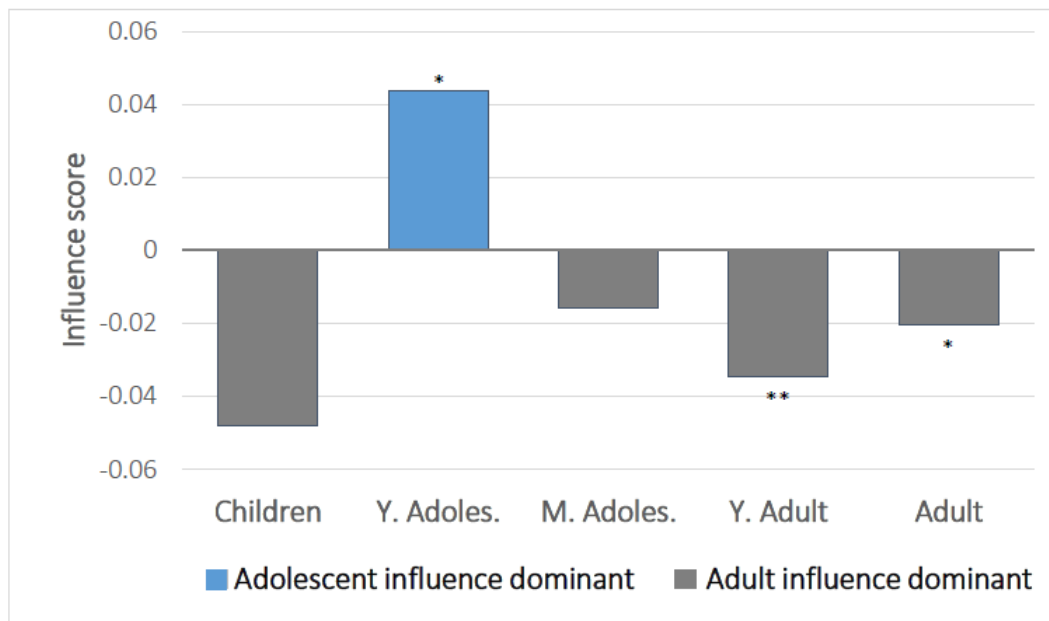


Figure 1.4. Social Influence on Risk Perception. A total of 563 participants rated the riskiness of everyday situations – before and after they were informed about the ratings of other people, either adults or teenagers. An index of conformity to other people's ratings is shown, depending on the origin of the social influence (adults or teenagers) across five age groups: children (aged 8 - 11), young adolescents (Y. Adoles., aged 12 - 14), mid-adolescents (M. Adoles., aged 15 - 18), young adults (Y. Adult, aged 19 - 25), and adults (aged 26 and over). *** $p < .001$, ** $p < .01$, * $p < .05$ significant difference in social influence effect between social influence origin (adults compared with teenagers) for each age group. Data published in Knoll et al. (2015). Reused from Fuhrmann et al. (2015) with permission from Elsevier.

In summary, adolescence is characterized by protracted changes in brain structure, cognitive function and social cognition. It has been proposed that these changes may make adolescence a time of high levels of plasticity or possibly even a sensitive period of brain development (Blakemore & Mills, 2014; Selemon, 2013; Spear, 2013; Steinberg, 2014).

1.2. Plasticity and Sensitive Periods of Development

In the 1960s, Hubel and Wiesel investigated the effect of monocular deprivation in kittens. Neurons in the corresponding visual cortex lost responsiveness to stimuli directed towards the previously deprived eye and started responding preferentially to the non-deprived eye (Wiesel & Hubel, 1963, 1965a). Monocular deprivation in the first three months of life was also associated with atrophy in cells in the thalamus receiving input from the deprived eye. Recovery from this atrophy was very limited, even after five years of light exposure. In contrast, monocular deprivation after three months of age produced virtually no physiological, morphological or behavioural effects (Hubel & Wiesel, 1970; Wiesel & Hubel, 1965b). The findings from these studies were taken as evidence that the first few months of life form a sensitive period for perceptual development, during which neuronal plasticity is heightened (Knudsen, 2004).

Plasticity describes the ability of the nervous system to adapt its structure and function in response to environmental demands, experiences and physiological changes (Pascual-Leone, Amedi, Fregni, & Merabet, 2005). It remains an underspecified concept, with usage varying between fields (Lövdén, Bäckman, Lindenberger, Schaeffe, & Schmiedek, 2010). Plasticity can be measured at the level of the neuron (changes in synaptic strength, dendritic branching, neurogenesis etc.; Selemon, 2013), tissue (changes in cortical thickness, grey matter volume etc.; Wenger, Brozzoli, Lindenberger & Lövdén, 2017), or by observing changes in behaviour (e.g. performance; Wenger et al., 2017, Fuhrmann et al., 2015).

The human brain retains a baseline level of plasticity throughout life – this is known as *experience-dependent plasticity* and underlies all learning (Greenough, Black, & Wallace, 1987). Motor skill learning is a well-described example of experience-dependent plasticity (Adams, 1987). New motor skills can be acquired compensate for injuries or adjust to new tasks, such as learning to play the piano or learning to juggle. Motor skill learning is accompanied by changes in white and grey matter of the motor cortex and is possible throughout life (Ungerleider, Doyon & Karni, 2002, Sampaio-Baptista et al., 2014). Plasticity during sensitive periods, on the other hand, is *experience-expectant* – the brain ‘expects’ to be exposed to a particular stimulus during this time (Greenough et al., 1987; Johnson, 2005).

Sensitive periods were originally referred to as ‘critical periods’. This term is used less now, as it has since become clear that some gain or recovery of function may be possible even outside the particular time window in question. In the case of monocular deprivation, research on monocular deprivation in kittens during the 70ies showed that animals can be trained to use the initially deprived eye after it is uncovered, and this can bring about a certain level of recovery (Dews & Wiesel, 1970).

1.2.1. Evolutionary perspectives on sensitive periods

Optimality models in evolutionary biology seek to understand a given phenotype in terms of its associated costs and benefits. Adaptive phenotypes are predicted to maximise the cost-benefit ratio (Parker & Smith, 1990). For instance, for

Escherichia coli, expressing Lac proteins is associated with costs (producing and maintaining the proteins) and benefits (the ability to digest lactose). The cost-benefit ratio of expressing Lac depends on the lactose content of the environment (Dekel & Alon, 2005). By experimentally manipulating this environmental constraint, Dekel and Alon (2005) showed that *E. coli* evolves to express optimal levels of Lac within a few hundred generations.

Optimality models have also provided an insight into costs and benefits of plasticity (Fawcett & Frankenhuis, 2015). Almost by definition, the main benefit of plasticity is that it allows an organism to adapt to new or changing environments. Plasticity is not cost-free, however. Plasticity requires energy and resources, as demonstrated by studies of *Drosophila melanogaster* (Mery & Kawecki, 2003, 2005). Fly larvae from strains selected for their high learning ability showed less competitive ability than larvae from low-learning strains (Mery & Kawecki, 2003).

Plasticity also introduces the possibility of error. Environmental cues, for instance, may be unreliable, irrelevant or interpreted incorrectly. Lorenz (1937) famously observed that young birds imprint on any moving object within their first few hours of life, and follow it, regardless of whether it is their mother, or not (Lorenz, 1937).

Two key factors influence the cost-benefit ratio of plasticity: (I) the quality and quantity of available environmental cues available, and (II) the degree of uncertainty in environments (Fawcett & Frankenhuis, 2015). Carroll and Corneli studied soapberry bugs (*Jadera haematoloma*) and found that plasticity of mating behaviour was related to the stochasticity of environmental conditions. In

Oklahoma, where sex ratios vary due to climatic fluctuations, mate-guarding in male bugs was plastic. In Florida, where sex ratios are more stable due to a constant climate, males engaged in a fixed amount of mate-guarding and were not able to adapt their behaviour variable sex-ratios imposed in the lab (Carroll & Corneli, 1995).

The quality of environmental cues and degree of uncertainty varies over ontogeny, which may produce changes in plasticity over the lifespan and predict sensitive periods of development (Fawcett & Frankenhuis, 2015; Panchanathan & Frankenhuis, 2016; Stamps & Krishnan, 2014). Simulation studies have highlighted that, under most circumstances, plasticity is expected to decline with age. Over ontogeny, the system accumulates more and more information. This reduces uncertainty, which in turn, is thought to reduce plasticity (Fawcett & Frankenhuis, 2015; Panchanathan & Frankenhuis, 2016). Later peaks in plasticity (for instance during adolescence) will be produced only if environmental conditions change drastically and contradict earlier estimates. The variable onset of puberty, as well as rapid changes in the social environment during adolescence, may be such conditions of uncertainty (Fawcett & Frankenhuis, 2015), but this suggestion is, at present, untested.

It should be noted that sensitive periods do not necessarily have to be adaptive. They could also be the by-product of other developmental programs and have no effect on fitness or even a negative effect (Laland & Brown, 2011; Michel & Tyler, 2005; Thomas & Johnson, 2008). Optimality models do not actually demonstrate that sensitive periods are optimal. Rather, they are a tool for understanding biological constraints on evolution (Parker & Smith, 1990).

1.2.2. Early sensitive periods

Most research on plasticity to date has focussed on early development. Early sensitive periods for visual, language and face perception development are particularly well described and highlight some key characteristics of sensitive periods in general.

Characteristics of early sensitive periods

Unlike translational work on sensitive periods of the visual system (Hubel & Wiesel, 1962, 1970; Wiesel & Hubel, 1963, 1965a, 1965b), studies in humans have relied on naturally occurring instances of visual deprivation in individuals born with cataracts, which occlude the lens of the eye. Sight may be regained after cataract reversal procedures. Cataract reversal studies indicate differences among sensitive periods for normal visual development, periods of sensitivity to deprivation and periods of recovery from deprivation (Lewis & Maurer, 2005). For visual acuity, for instance, the period of visually-driven typical development extends over the first 7 years of life, but individuals remain sensitive to deprivation up to 10 years of age and some recovery of function may be possible throughout life (Maurer & Lewis, 2012).

Language development, too, generally shows heightened plasticity in childhood (Kuhl, 2010; Sakai, 2005), although there is no single sensitive period for language. Different linguistic abilities are acquired by partly separable neural systems, and these differ in their response to deprivation and periods of heightened plasticity (Kuhl, 2004). Congenital deafness, for instance, is associated with altered processing of grammatical information while semantic processing appears to be

insensitive to auditory deprivation (Neville, Mills, & Lawson, 1992). This highlights the specificity of sensitive periods.

The timing of onset and offset of early sensitive periods is malleable. Studies with monkeys have demonstrated that the face sensitive period at the beginning of life can be extended by two or more years if infant monkeys are not exposed to face stimuli during this time. Face deprivation, therefore, delays the onset of the sensitive period (Sugita, 2008). The end of a sensitive period may in some cases be self-generated: learning may drive the commitment of neural structures, effectually reducing plasticity (Johnson, 2001, 2005). Face perception undergoes perceptual narrowing, for instance, during which individuals become better at processing the category of faces they are most exposed to, at the expense of categories they see less frequently, producing effects such as the own-race bias of face perception (Malpass & Kravitz, 1969; Scott, Pascalis, & Nelson, 2007; Tanaka & Pierce, 2009). Another explanation for the end of sensitive periods is that plasticity may not actually reduce but, instead, that frequency of or variation in environmental stimulation decreases (Johnson, 2005).

Cellular mechanism of early sensitive periods

Early development of the visual system has served as a model for cellular mechanisms of sensitive periods (Hensch, 2005). Evidence from this domain indicates that plasticity after sensitive periods is not lost but rather actively dampened by functional and structural 'brakes' (Takesian & Hensch, 2013). Functional brakes include neurotransmitters like serotonin and dopamine (Bavelier, Levi, Li, Dan, & Hensch, 2010), enzymes like histone-deacetylase (HDAC)

(Bieszczad et al., 2015) and proteins like myelin-derived Nogo (McGee, Yang, Fischer, Daw, & Strittmatter, 2005). Structural brakes inhibit further neuronal growth. Myelin, for example can hinder axonal sprouting (Bavelier et al., 2010).

That capacity for high levels of plasticity is not lost after early childhood, is evidenced by incidents of stroke or traumatic brain injury naturally re-triggering plasticity (Hubener & Bonhoeffer, 2014). It is also possible to artificially enhance plasticity beyond childhood by inhibiting functional brakes. Treatment with the serotonin reuptake inhibitor fluoxetine has been found to restore visual function in amblyopic adult rats, for example (Maya Vetencourt et al., 2008). Another example comes from music learning. Whilst absolute pitch cannot usually be acquired after 6 years of age, inhibition of HDAC using valproate has been shown to facilitate the acquisition of absolute pitch in human adults (Gervain et al., 2013).

In summary, early sensitive periods for visual and language development highlight some key facets of sensitive periods. Namely, that sensitive periods vary within and between domains, that the timing of sensitive periods is malleable and that plasticity after sensitive periods can be re-accessed. The evidence-base for sensitive periods beyond early childhood, however, is much sparser.

1.3. Adolescence as a Sensitive Period of Development

Adolescence, with its protracted changes in brain structure and function, has been posited to form a 'second period of heightened malleability' after early childhood (Steinberg, 2014, p. 9; see also Blakemore & Mills, 2014; Selemon,

2013). In this section three areas of adolescent development that have been proposed to be characterised by heightened plasticity, are explored: memory, the effects of stress, and the effects of drug use. The argument is made, that advances in developmental studies have yielded intriguing data that is consistent with heightened plasticity in adolescence. However, despite recent advances, concrete evidence for sensitive periods is mostly lacking.

1.3.1. What evidence would be consistent with adolescence being a sensitive period?

If adolescence is indeed a sensitive period, certain patterns of development ought to appear. First, the impact of a specific stimulus on brain and behaviour should be higher in adolescence than before or after. For that reason, studies are necessary that compare adolescents with other age groups. Only if multiple age groups are considered, can we assess whether adolescence is a stand-alone period of heightened plasticity (Model A, Figure 1.5), a continuous sensitive period with childhood (Model B, Figure 1.5) or not a sensitive period at all (Model C, Figure 1.5).

As a result of the differences in the timing of maturation of different brain regions and circuits (Tamnes, Walhovd, Dale, et al., 2013), considerable variation in the on- and off-set of sensitive periods for different domains would be expected. Just as early development is characterised by multiple sensitive periods (Kuhl, 2004; Lewis & Maurer, 2005), adolescence is not proposed to be a sensitive period per se; instead, it is proposed that there are certain periods in adolescence during which a specific input from the environment is expected.

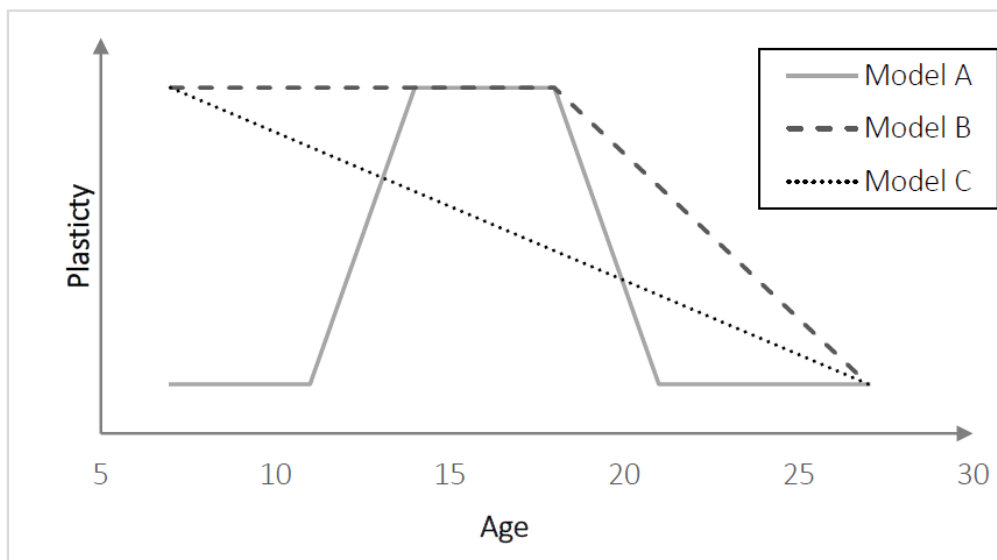


Figure 1.5. *Models of Plasticity in Adolescence*. Adolescence may be a stand-alone period of heightened plasticity (A) or form a continuous sensitive period with childhood (B). Alternatively, plasticity may decline continuously from childhood through adolescence and into adulthood (C). Adapted by Fuhrmann et al. (2015) for adolescence from Thomas (2012). Reused with permission from Elsevier.

If certain environmental stimuli do indeed have a heightened impact during this time, we would expect there to be enhanced learning, particularly of late-maturing skills. This is discussed in section 1.3.2. A lack of stimulation or aberrant stimulation would also be expected to have a disproportionate effect during this time, however. This feature of sensitive periods is discussed in the section 1.3.3.

Adolescent plasticity might differ from plasticity early in development because, unlike babies and young children, adolescents are more likely to actively choose the environmental stimuli they experience. During childhood, environments are usually more structured by parents or caregivers, while adolescents have more autonomy to choose what to experience and with whom (Lam et al., 2014; Larson

& Richards, 1991). We might thus expect a large degree of individual differences in sensitive periods in adolescence, and some sensitive periods may only ever be experienced by a subset of adolescents, as discussed in section 1.3.4.

1.3.2. Adolescence as a sensitive period for memory and other complex cognitive skills

At age 35, we are more likely to recall autobiographical memories from ages 10 - 30 than prior or subsequent to this period - a phenomenon referred to as the 'reminiscence bump' (Rubin & Schulkind, 1997). The reminiscence bump is remarkably robust and shows a similar pattern when tested with different mnemonic tests and in different cultures (Conway, Wang, Hanyu, & Haque, 2005; Rubin & Schulkind, 1997). In addition to autobiographical events, the recall of music, books, films and public events from adolescence is also superior compared to other periods of life (Janssen, Chessa, & Murre, 2007; Janssen, Murre, & Meeter, 2008). Even mundane events that happened in adolescence and early adulthood appear to be over-represented in memory, suggesting that mnemonic capacity in general is heightened during this time of life (Janssen & Murre, 2008). For example, a large-scale study showed a peak in visuo-spatial memory between 14 - 26 years of age (Murre et al., 2013).

While these data are suggestive of sensitive periods, studies are needed that can provide experimental evidence for plasticity across development. There is some evidence for plasticity of working memory from training studies. For children and young adolescents (mean age: 9 years), gains in n-back type working memory

training, but not knowledge-based training, was shown to transfer to improvements in fluid intelligence (Jaeggi, Buschkuhl, Jonides, & Shah, 2011). Improvements were sustained over a 3-month period during which time no further training was implemented. Working memory training may also be effective in adolescents aged 14 - 15 with poor executive functioning, as well as in typically-developing controls (Løhaugen et al., 2011).

Other complex cognitive functions like relational reasoning and numerosity discrimination can be trained as well. Relational reasoning and fluid intelligence were originally thought of as stable characteristics. Recent research has shown, however, that relational reasoning training can result in changes in fluid intelligence (Mackey, 2012). Relational reasoning training has been shown to increase IQ by 10 points in children aged 7 - 9 (Mackey, Hill, Stone, & Bunge, 2011). There is some evidence that training induces plasticity in white matter microstructure. When young adults (mean age: 21 years) trained relational reasoning for standardized law school admission tests, white matter microstructure was altered and fronto-parietal connections in relational reasoning networks were strengthened (Mackey, Miller Singley, & Bunge, 2013; Mackey et al., 2012).

Similarly, education and environment have been shown to influence numerosity discrimination. Access to schooling in indigenous South American participants aged 4 - 63 has been shown to predict number acuity (Piazza, Pica, Izard, Spelke, & Dehaene, 2013). Experimental data also suggested that training on approximate addition and subtraction of arrays of dots in adults selectively improves symbolic addition and subtraction (Park & Brannon, 2013).

In summary, there is some observational evidence for heightened memory capacities in adolescence and experimental evidence for plasticity of working memory, relational reasoning and numerosity discrimination in childhood and adulthood. However, little is known about training effects in adolescence and there are virtually no studies comparing training effects between age groups. To address this issue, Chapter 5 details the results of a large-scale training study with adolescent and adult age groups.

1.3.3. Adolescence as a sensitive period for the effects of stress

Many mental illnesses have their onset in adolescence and early adulthood (Figure 1.6) (Kessler et al., 2007; Kessler et al., 2005). A longitudinal study showed that 73.9% of adults with a mental disorder had a diagnosis before 18 years of age and 50.0% before 15 years of age (Kim-Cohen et al., 2003). It is thought that psychiatric disorders develop due to a combination of genetic predispositions and environmental stressors; and some may be triggered by the onset of puberty (Andersen & Teicher, 2008; Rosenthal, 1970). Social stress in particular may have a disproportionate impact during this time (Andersen & Teicher, 2008). The experience of acculturation stress attributable to migration, for example, longitudinally predicts internalising symptoms such as depression and anxiety between ages 16 - 18 (Sirin, Ryce, Gupta, & Rogers-Sirin, 2013). Of course, adolescence is not the only life stage during which social stress has adverse effects. Bullying in childhood (age 7 - 11), for instance, also has lasting effects on physical and mental health in adulthood (Takizawa, Maughan, & Arseneault, 2014).

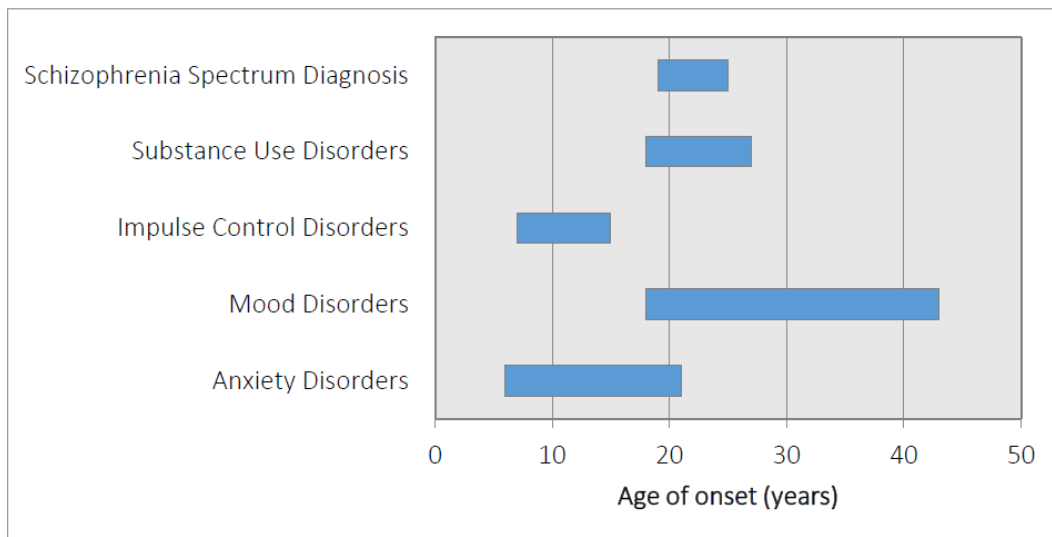


Figure 1.6. Age of Onset of Psychiatric Disorders. The interquartile ranges for the age of onset of selected psychiatric disorders are shown. Data for Schizophrenia Spectrum Diagnosis were adapted from the Early Psychosis Prevention and Intervention Centre in Melbourne, Australia, as reviewed by Kessler et al. (Kessler et al., 2007). Data for the remaining disorders stems from the National Comorbidity Survey Replication in the United States (Kessler et al., 2005).

Rodent studies provide an opportunity to experimentally manipulate exposure to social stress, and have offered valuable insights into the deleterious effects of stress across development (Marin, 2016). Adolescence in female rats lasts approximately from post-natal day (PND) 30 - 60 and from PND 40 - 80 in males. In female mice, adolescence lasts from PND 20 - 40 and from PND 25 - 55 in males (Schneider, 2013). It should be noted that there is considerable variation in terms of the age at which rodents are classified as adolescent or adult in the literature, however, as there is in humans (Schneider, 2013). Adolescent rats subjected to repeated defeat by a dominant individual have been shown to present with different behavioural patterns (more avoidance rather than aggression), and less recovery from renewed stress, compared with adult rats. Exposure to stress in

adolescence (compared with adulthood) was also associated with less neuronal activation in areas of the prefrontal cortex, cingulate and thalamus (Ver Hoeve, Kelly, Luz, Ghanshani, & Bhatnagar, 2013). This study did not include juveniles, limiting the conclusions that may be drawn regarding sensitive periods.

Adolescence may also be a vulnerable period for recovery from the experience of social stress (Pattwell et al., 2012). Fear extinction learning is key for a healthy response to stress (Maroun et al., 2013). For psychiatric conditions such as post-traumatic stress disorder, stress persists even though the stressor is no longer present. Fear extinction learning was found to be attenuated in adolescents (12 - 17 years), as compared to children (5 - 11 years) and adults (18 - 28 years) (Figure 1.7). The rodent data in the study indicated that a lack of synaptic plasticity in the ventromedial prefrontal cortex during adolescence is associated with decreased fear extinction. This implies that desensitization treatments, which are based on the principles of fear extinction learning, may be less effective in adolescence, and highlights the need for the development of alternative treatment approaches for this age group (Pattwell et al., 2012). The particular strength of this study lies in that fact that it included child, adolescent and adult age groups, as well as providing neural evidence in rodents. The results suggest that adolescence may be a sensitive, or vulnerable, period for recovery from stress.

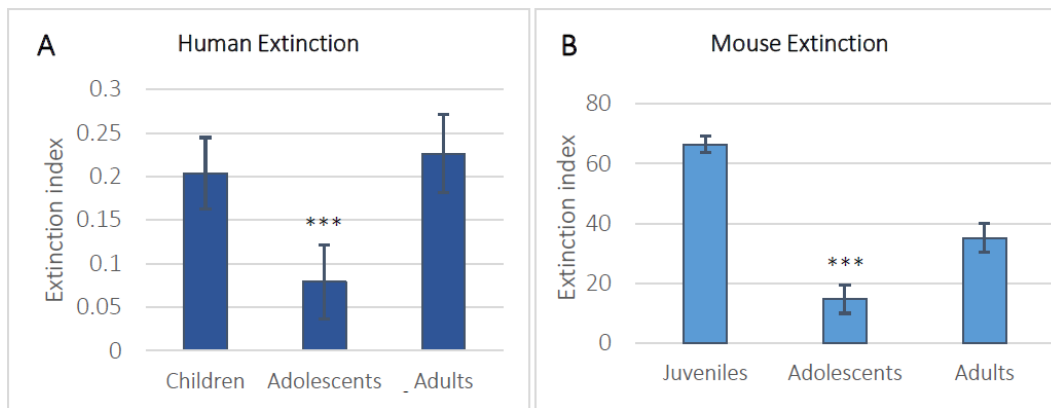


Figure 1.7. Fear Extinction Learning in Mice and Humans across Development. Mean indices for fear extinction learning with standard error bars in humans (A) and mice (B) for children or juveniles, adolescents, and adults. *** $p < .001$ for an attenuation in fear extinction compared with other age groups. Adapted from Pattwell et al. (2012) with permission from PNAS.

The absence of any social stimulation can have deleterious effects as well. Social isolation in male and female rats has been shown to have irreversible effects on some aspects of exploratory behaviour, but only if the isolation occurred between PND 25 - 45, and not before or after (Einon & Morgan, 1977). Therefore, this appears to be a vulnerable period for social deprivation in rats. This paradigm cannot be directly applied to humans, but as discussed in the section on peer influence, there is some evidence that human adolescents show greater levels of anxiety in response to social exclusion than adults (Sebastian et al., 2011; Sebastian et al., 2010). Chapter 4 systematically investigates the effects of social stress across adolescence so as to help develop and time mental health interventions aimed at strengthening resilience to social exclusion.

1.3.4. Adolescence as a sensitive period for the effects of drug use

Adolescence is a time of heightened experimentation with drugs (Eaton et al., 2012; Steinberg, 2008), with cannabis being one of the most widely used recreational drugs among adolescents and adults in the US and UK (Johnston, O'Malley, Bachman, & Schulenberg, 2013; The NHS Information Centre, 2011). It has been estimated that 15.2% of Europeans aged 15 - 24 have used cannabis in the last year and 8% in the last month (European Monitoring Centre for Drugs and Drug Addiction, 2011). Cannabinoid exposure during adolescence (before age 17) as compared to exposure after 17 years of age has been shown to result in lasting changes in brain structure and cognitive deficits, possibly making adolescence a vulnerable period for its effects (Ehrenreich et al., 1999; Pope et al., 2003).

Recreational cannabis use before the age of 18 (but not in adulthood), or heavy use at any age, has been linked to grey matter atrophy in the adult temporal pole, parahippocampal gyrus and insula (Battistella et al., 2014). Longitudinal data has indicated that self-reported, persistent cannabis use between 13 - 15 years is associated with a significant decline in IQ (Meier et al., 2012). The longer the period of cannabis consumption, the greater the decline in IQ (Meier et al., 2012). This decline in IQ was found to be more pronounced in participants who used cannabis before the age of 18 as compared to those who started to use cannabis after 18.

These findings suggest that the adolescent brain may be particularly sensitive to the adverse consequences of cannabis use. It should be noted, however, that alternative explanations, such as pre-existing mood or anxiety disorders mediating both cannabis-use and cognitive problems, cannot be ruled out in this study

(Blakemore, 2013). These studies also did not include younger age groups, and it is possible that the developing brain during childhood would show a similar or even greater sensitivity to cannabis than in adolescence. Even if that were the case, however, such sensitivities would not be commonly observed in humans as adolescent or adulthood will usually be the first point of contact with recreational drugs.

Molecular and cellular data on the effects of cannabis in adolescence is sparse but there is some indirect evidence for heightened sensitivity during adolescence. It has been shown that cannabis affects the endocannabinoid system, which, along with other neurotransmitter systems (e.g. the glutamatergic and dopaminergic systems), undergoes extensive restructuring during adolescence (Malone, Hill, & Rubino, 2010). While the two key cannabinoid receptors CB1 and CB2 are already present in the rodent embryo (gestational day 11 - 14 (Berrendero, Sepe, Ramos, Di Marzo, & Fernandez-Ruiz, 1999), neuroanatomical distribution and number of receptors change during development. CB1 receptor expression in several brain regions was found to peak with the onset of puberty in female and male rodents (Rodriguez de Fonseca, Ramos, Bonnin, & Fernandez-Ruiz, 1993). Any disturbance caused by cannabis exposure during the adolescent period may have lasting effects on the endocannabinoid system, which affects neurodevelopmental processes like neuronal genesis, neural specification, neuronal migration, axonal elongation and glia formation (Berghuis et al., 2007; Harkany, Keimpema, Barabás, & Mulder, 2008; Oudin et al., 2011). For instance, exposure to D9-tetrahydrocannabinol, the main psychoactive ingredient in cannabis, during puberty in female rats (PND 35 - 45) resulted in a decrease in CB1 receptor

density and functionality in several brain regions (Ellgren, Spano, & Hurd, 2007). However, comparative data from other age groups is lacking.

Strong evidence for an adolescent sensitive period for drug-use comes from a set of studies investigating chronic cannabinoid exposure in male rodents. Cannabinoid exposure in adolescence (PND 40 - 65) predicted long-term cognitive deficits in adulthood (object recognition memory), whereas similar exposure in prepubescent (PND 15 - 40) and young adult rodents (PND 70 and over) was not linked to such persistent deficits (Schneider, Drews, & Koch, 2005; Schneider & Koch, 2003). It is not clear, however, if this evidence directly translates to humans.

It should also be noted that only a subset of human adolescents experiment with drugs such as cannabis and that drug-use may be mediated by peer-influence. Adolescents whose friends regularly consume tobacco, alcohol and cannabis are more likely to use drugs themselves (Branstetter, Low, & Furman, 2011). Future studies are needed to investigate individual differences, particularly in relation to peer influence and risk-taking, to understand when and for whom adolescence may be a vulnerable period for drug use.

To summarize, evidence for plasticity in terms of memory and the effects of social stress and drug-use, is consistent with the proposal that adolescence is a sensitive period for these areas of development. The strongest evidence for sensitive periods to date comes from rodent studies showing a heightened vulnerability to the disruptive effects of social isolation and cannabis use, as well as reduced fear-extinction learning. There is little conclusive evidence for human adolescence, however. Studies are needed, which focus on the effects of training or stress

across adolescence and other age groups. This will have particular relevance for education, and may help answer the question of what to teach when, as well as identify the most effective time for providing school-based educational and mental-health interventions.

1.4. Learning, Plasticity and Education

As discussed in this chapter, developmental research on learning and plasticity has traditionally focused on children, with a particular emphasis being placed on the first few years of life. It is undisputable that plasticity for many important cognitive and motor functions is heightened during early childhood, and that development during this period of life can have bottleneck effects for later life (Eluvathingal et al., 2006; Howard-Jones, Washbrook, & Meadows, 2012). However, much of the evidence for early sensitive periods stems from animal models and studies of severe deprivation, which may not necessarily generalize to human development under typical environmental conditions. Yet, this evidence has been used to draw the rather extreme and perhaps premature conclusion that ages 0 - 3 years is *the* critical period for learning, after which developmental trajectories are more or less fixed (Howard-Jones et al., 2012; Thomas, 2012).

As highlighted in this chapter, many cognitive and social skills relevant to education normally develop beyond childhood. Adolescence, in particular, is characterized by protracted changes in brain structure and the maturation of cognitive skills requiring high levels of self-control. There is good evidence that plasticity for many higher level cognitive skills, like working memory or

reasoning, is maintained after childhood and some evidence suggestive of heightened plasticity in adolescence (Blakemore & Mills, 2014; Selemon, 2013; Spear, 2013; Steinberg, 2014), but conclusive evidence has been lacking.

To investigate learning and plasticity in adolescence, this thesis highlights areas of adolescent development that are particularly relevant to education. Chapter 2 discusses the methodology and design of the studies presented in this thesis. Chapter 3 - 6, the experimental chapters, address four main research questions:

(I) Are there age-related differences in face cognition between adolescence and adulthood? A better understanding of social cognition and social development is critical to fostering social competency inside and outside the classroom (Blakemore, 2010). Chapter 3 investigates development of face cognition as a model for broader changes in social cognition beyond childhood. Specifically, this chapter examines age-related differences in face cognition between early adolescence and adulthood. It probes whether age effects are specific to face cognition, or rather due to general increases in cognitive ability, and investigates developmental differences between face cognition sub-domains and genders.

(II) Do the effects of social exclusion on cognitive performance differ between age groups? Chapter 4 examines the effects of social exclusion on cognitive performance in adolescence and adulthood. It probes whether some age groups are more affected by social exclusion than others. This may help inform the timing of bullying interventions in schools.

(III) Do some age groups benefit more from cognitive training than others? Chapter 5 investigates whether complex cognitive skills like numerosity

discrimination, relational reasoning and face perception can be trained in adolescence and adulthood and probes whether some age groups benefit more from training than others. It also examines whether relational reasoning and face perception training generalizes to related cognitive skills by investigating transfer from relational reasoning to working memory, and from face perception to face memory.

(IV) What are the neurocognitive correlates of diligence? Using a different theoretical framework than the preceding chapters, namely, the dual systems hypothesis, Chapter 6 investigates individual differences in self-control and their ramifications for education. It probes whether the interplay between frontal control and striatal reward systems is related to academic diligence, the ability to regulate behaviour in the service of educational goals. Using behavioural, structural MRI, functional MRI and connectivity data, it assesses the neurocognitive correlates of diligence in adolescence. This may foster a new understanding of the mechanisms of academic diligence and ultimately inform the design of educational interventions aimed at strengthening adolescent self-control.

2. Chapter 2: Design Issues in Developmental Studies

In this chapter, design issues, that are particularly relevant to the experimental studies presented in Chapters 3 - 6, are discussed. First, different types of developmental designs are explored to highlight potential benefits and limitations of each. Then, validity questions, that have been especially pertinent to the design of the studies in this thesis, are discussed, namely instrumentation effects and missing data. Even though only one of the experimental chapters covers a training study, an entire section is devoted to the design of training studies, because there are many specific issues requiring consideration, such as temporal design and the choice of transfer tasks. This chapter ends with a discussion of statistical designs and analyses used in developmental studies. A case is made for Generalized Linear Mixed Models as a particularly useful tool for the interpretation of complex developmental data sets, and particular attention is devoted to discussing how age can be modelled within this framework.

2.1. Types of Developmental Designs

There are two main types of design that can be used to study development - cross-sectional and longitudinal designs - as well as designs that combine aspects of both (Baltes, 1968; Little, 2013). The choice of design will depend on the researchers' hypotheses as well as practical considerations and will affect inferences that can be drawn from it.

2.1.1. Cross-sectional designs

Cross-sectional designs typically compare participants of different ages on an outcome measure. Cross-sectional designs are relatively time- and cost-efficient, in that multiple variables can be collected at a single point in time. This efficiency was the principal reason why we chose cross-sectional designs for most of our studies (see Chapters 3, 4 and 6).

However, cross-sectional studies are limited in their ability to describe developmental change. Inter-individual variability makes it difficult to rule out the possibility that differences in age groups reflect accidental differences between groups rather than developmental differences (Mills & Tamnes, 2014). This problem can be ameliorated with large sample sizes, which increase power and protect against false positives (Button et al., 2013). Ultimately, however, cross-sectional data on age differences should be considered as preliminary for studies on developmental change with longitudinal aspects (Kraemer, Yesavage, Taylor, & Kupfer, 2000).

2.1.2. Longitudinal designs

A pure longitudinal study takes a single age-cohort and takes repeated measures on an outcome variable over time (Baltes, 1968). For example, a cohort of eight-year-olds may be recruited for the first wave of the study and then tested again five and ten years later to measure development on a reasoning task. Longitudinal designs are more powerful than cross-sectional designs and allow examining change over time (McArdle, 2008).

However, longitudinal studies take longer to complete than cross-sectional studies and are limited by the fact that age is perfectly confounded with time of measurement: If the same set of participants score higher on a reasoning test at age 18 than at age 8, it might be because of developmental changes, or it might be because participants have now taken the test more than once (Little, 2013). Moreover, changes in the testing protocol over time (e.g. changes in the testing personnel, imaging sequences, tasks used) can confound developmental effects.

2.1.3. Combination designs

There are a variety of mixed designs that attempt to combine the benefits of cross-sectional and longitudinal studies. One example of such designs is the cross-sequential design. It begins with a cross-sectional sample and follows this sample over time (Little, 2013). The advantage of cross-sequential designs for the study of developmental change is that data can be collected more speedily than in conventional longitudinal designs. This design also allows the researchers to model intra-individual change (Sabol, Chase-Lansdale, & Brooks-Gunn, 2015). Training studies with multiple age groups (see Chapter 5) are an example of cross-sequential designs.

Another type of design that has become very popular over recent years is the cohort-sequential or accelerated longitudinal design (Little, 2013). Accelerated longitudinal studies are like a single longitudinal study starting over and over again. Initially, researchers may recruit 10 -, 11 -, 12 - and 13 - year olds for their study. Two years later, they may test 12 -, 13 -, 14 - and 15 - year olds in the

second wave of the study. Participants, who were 10 and 11 in the first wave, are 12 and 13 in the second wave and can be tested again. Therefore, participants are followed longitudinally, but a broader age range can be covered in a shorter amount of time than in a traditional longitudinal study (Galbraith, Bowden, & Mander, 2017). Such designs are particularly suited to separating age effects from confounds such as learning effects, changes in experimenters or equipment (Prinzle & Onghena, 2014; Roe & Korn, 1993). Accelerated longitudinal structural imaging studies have provided rich insights into the protracted development of the human brain during adolescence and have highlighted an extraordinary amount of inter-individual variability in developmental trajectories (Giedd et al., 1999; Gogtay et al., 2004; Mills et al., 2014; Mills & Tamnes, 2014; Tamnes et al., 2017).

2.2. Validity Issues

Validity issues that are particularly important to the developmental studies presented in this this thesis include (I) instrumentation effects and (II) missing data.

2.2.1. Instrumentation effects

Instrumentation effects refer to violations of internal validity due to the measures or instruments used (Little, 2013). For developmental studies, floor effects (score limitation at the bottom of a scale) and ceiling effects (score limitation at the top of a scale) are a particular concern (Little, 2013; Uttl, 2005). For instance, a

researcher might not find a difference in reasoning scores between 10 - and 18 - year-olds because the particular test used was so simple that all participants performed nearly perfectly. If the test was harder, age group differences might have emerged. Instrumentation effects were a potential concern in the development of the studies presented in Chapters 3, 4, and 5, for which we recruited broad age ranges and participants in their early teens were compared to participants in their thirties.

Instrumentation effects can be limited to some extent with careful piloting and calibration of task difficulty. Particularly when the age range studied is broad, however, it can be difficult to find an appropriate measure that produces no floor or ceiling effects. Researchers might then be tempted to choose different tasks measuring the same construct in different age groups. Such heterotypic measures can limit the validity of age group comparisons: age group differences become confounded with task differences (Little, 2013). Adaptive and testing-the-limit designs (see section 2.3.2. below and Chapter 5 for examples) may be better suited to preventing floor and ceiling effect in developmental studies (Alloway et al., 2009; Baltes, Lindenberger, & Staudinger, 2006; Wechsler, 1999).

Statistically testing for floor or ceiling effects is a complex issue. One possibility is to compare the test scores of different age groups to the maximum or minimum possible scores. It is important to consider the possibility of asymptotes here - a score of 100% on a particular task may be theoretically possible but may never be achieved in practice (L. Wang, Zhang, McArdle, & Salthouse, 2009). Researchers therefore need to test their data against a realistic minimum or maximum value. See L. Wang et al. (2009) for an in-depth discussion of different analysis methods.

2.2.2. Missing data and selective attrition

In studies with longitudinal components, like the training study presented in Chapter 5, attrition (and therefore missing data) is almost unavoidable, but missing data can also be a problem in cross-sectional data. The source of attrition needs to be carefully considered, because it can be a serious limitation to statistical inference (Koutoumanou & Wade, 2015a; Little, 2013; Newsom, 2015).

Researchers need to report and analyse missing data patterns that are most pertinent to their hypotheses. The first step is to identify whether missing data is predicted by variables of interests or auxiliary variables (Little, 2013). In Chapter 5, for instance, we analysed whether drop-out from training differed between training and age groups.

Conceptually, and rather confusingly termed, data can be either *not missing at random (NMAR)*, *missing at random (MAR)*, or *missing completely at random (MCAR)*. NMAR means that missingness is related to the variable with missing scores, and cannot be explained by other variables in the dataset (Donders, van der Heijden, Stijnen, & Moons, 2006). NMAR is the most challenging scenario because even after including auxiliary variables, systematic differences in missingness remain (Koutoumanou & Wade, 2015a). For instance, participants may be more likely to drop out of a training study if they have low scores on the trained task. If this cannot be controlled for by any of the auxiliary variables collected, missingness would be NMAR.

If data is MAR, it means that missingness is not related to the variable that has missing data. It can be related to other variables in the dataset, however (Donders

et al., 2006). Any differences between missing values and observed values can be explained by variables in the dataset (Koutoumanou & Wade, 2015a). For instance, participants from private schools may be less likely to drop out of a training study than participants from state schools, because the teachers of the former might have more time to support their training. Taking school-type into account, would completely explain missingness in this example.

MCAR means that missingness is truly random and not related to any of the variables in the dataset. This means that there are no systematic differences between missing and observed values (Koutoumanou & Wade, 2015a). MCAR is the most desirable pattern of missingness because it poses the least difficulty in terms of analysing and interpreting the data. However, real data is rarely MCAR (Little, 2013).

The exact type of missingness is mostly impossible to establish (Newsom, 2015). Most missingness will combine aspects of MCAR, MAR and NMAR (Koutoumanou & Wade, 2015a; Little, 2013). This is problematic, particularly when choosing a method with which to address missingness because the optimal method depends on the type of missingness and results can vary depending on the method used (Newsom, 2015).

The most common approaches to addressing missingness are *list-wise deletion*, *pairwise-deletion* and *imputation*. In *list-wise deletion*, all data for a participant who has at least one instance of missingness is removed. This method is used by default when using t-tests or ANOVAs. It is the most common but usually least effective method of dealing with missing data (Koutoumanou & Wade, 2015a). If

data is MAR or NMAR, list-wise deletion can induce bias into statistical inference. If data is MCAR, list-wise deletion does not induce bias but still reduces effective sample size (T. A. Myers, 2011).

Pairwise deletion uses the entire available dataset but does not model the instances of missingness. Generalized Linear Mixed Models use pairwise deletion by default (see section 2.4.1. below) and was used for the studies presented in this thesis. Pairwise deletion has the advantage that it does not reduce or inflate sample size but it can still induce bias if data is NMAR (Koutoumanou & Wade, 2015a). In Chapter 5, our study with longitudinal aspects, we therefore carried out several supplementary analyses to check missing data patterns.

Imputation methods replace missing values with values predicted from the available data. Simple imputation methods include carrying the last observation forward or using the mean of the missing variable. These methods can both induce bias and falsely inflate the sample size (Koutoumanou & Wade, 2015a; Newsom, 2015). More modern approaches like multiple random imputation are much more effective at dealing with MCAR and MAR data (Newsom, 2015) but can be difficult to implement when using methods like Generalized Linear Mixed Models (but see Jagdhuber, 2016, for recent advances).

In summary, it can be difficult to identify missing data types and address missingness optimally in analyses. Nonetheless, it is important to explore how inferences may be influenced by missingness (Koutoumanou & Wade, 2015a; Newsom, 2015).

2.3. Design Issues in Training Studies

This section considers five particularly pertinent design issues in training studies like the one presented in Chapter 5: (I) temporal design; (II) adaptive designs; (III) control groups; (IV) randomization; and (V) transfer tasks.

2.3.1. Temporal design

Gollob and Reichardt (Gollob & Reichardt, 1987) stated the intuitive maxim that causes take time to exert effects and that the magnitude of effect depends on the measurement intervals. This is particularly relevant to training studies where researchers need to decide upon a multitude of temporal issues, including the duration of the individual training session, the total number of sessions, spacing between training sessions as well as the spacing between baseline testing and any follow-up testing. All of these decisions could affect the size of the training effects measured. Despite the obvious centrality of temporal issues in study designs, there is surprisingly little guidance and evidence available.

The most relevant line of research compares the effects of spaced and massed learning on subsequent recall (Ebbinghaus, 1885). This research generally indicates that the same amount of material is better recalled when encoded on more rather than fewer occasions (Sisti, Glass, & Shors, 2007). It is relatively difficult to derive specific recommendations from this literature however, because there are only a handful of studies that have directly tested the effect of spacing of training on training gains and these studies usually find that the maximum interval tested was most effective (Penner et al., 2012; Z. Wang, Zhou, & Shah,

2014). For our training study, we therefore had to refer to the design of earlier training studies for guidance. It is to be hoped that in the future temporal design will become a more active area of investigation so as to provide reliable, quantitative evidence for the design of interventions.

2.3.2. Adaptive designs

Adaptive training means that task difficulty is scaled according to performance. In an up-down design, for instance, participants are given a more difficult item after getting an item right. After getting an item wrong, they are given an easier item next. Adaptive designs therefore ensure that participants are continually challenged. They may prevent ceiling effects and the formation of strategies, as well as promote transfer (Schwaighofer, Fischer, & Bühner, 2015). For these reasons, we made training adaptive in the study presented in Chapter 5.

2.3.3. Active vs. passive control groups

Training studies should always include a control group. A passive control group receives no treatment and is the equivalent of a wait-list control group in pharmacological studies. An active control group receives an alternative treatment, which makes it similar to placebo groups in pharmacological trials (Klingberg, 2010). Passive control groups control only for the effects of repeated testing, while active control groups also control for expectancy or placebo effects as well as generic effects of adhering to a schedule and so on. The most common type of active control condition used in training studies is a non-adaptive version

of the treatment condition (Schwaighofer et al., 2015). In Chapter 5, we opted to use a different design. We used three different, but equally challenging cognitive tasks that served as control tasks for one another (relational reasoning, face perception and numerosity discrimination). This design was chosen because it was more efficient than using three different tasks with a non-adaptive control condition each. It also ensured that task difficulty did not differ between training and control groups.

2.3.4. Randomized-control studies

Randomized-control study designs have been adapted for education from the gold-standard for medical studies, randomized-control trials (RCTs). They feature random allocation to a training or control group and usually a double-blind design where experimenters and participants are unaware as to which group participants were allocated to (Sullivan, 2011). Some features of RCTs may not be realistic in behavioural interventions, however. It may be possible to blind experimenters to participants' training group, for example using automated testing. However, blinding participants to their treatment may be impossible when using cognitive tasks (as was the case in Chapter 5). Moreover, random allocation to training and control groups can be challenging if the intervention is delivered in the classroom. In Chapter 5, we therefore opted to ask participants to in their own time, instead of supervising training. The trade-off was, that we expected and got lower adherence to the training schedule than in supervised interventions.

2.3.5. Transfer tasks

Transfer tasks are used to evaluate how much training generalizes. This is key, particularly for interventions that are trialled for use in the classroom (Klingberg, 2010).

Whether training a certain cognitive skills can improve performance on non-trained tasks is still debated. Intervention studies have reported transfer mainly between skills that share similar cognitive processes such as different working memory tasks (*near-transfer*) (Klingberg, 2010; Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009). A small number of studies in children and adults have found evidence for *far-transfer*, that is, between less related skills. For instance, working memory training has been found to transfer to fluid intelligence (Bergman-Nutley & Klingberg, 2014; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008; Klingberg et al., 2005), arithmetic performance (Bergman-Nutley & Klingberg, 2014) and cognitive control (Klingberg et al., 2005), and reasoning training has been found to transfer to fluid intelligence (Bergman-Nutley & Klingberg, 2014; Klingberg et al., 2005; Mackey, Hill, Stone, & Bunge, 2011). However, large-scale studies and meta-analyses have failed to provide evidence for such far-transfer (Melby-Lervåg & Hulme, 2013; Owen et al., 2010; Schwaighofer et al., 2015). Nonetheless, the inclusion of far-transfer tasks is important when evaluating whether training generalizes to other skills, and was investigated in Chapter 5.

2.4. Statistical Design and Analysis

Developmental studies usually produce complex data sets. They often feature mixed designs, which contain both fixed and random effects. Moreover, the effect of interest is usually an interaction - for instance, the interaction of age and training group. They also often include nested effects, meaning that effects are not independent. There are now several methods available to deal with these kinds of complex datasets efficiently, of which Generalized Linear Mixed Models are rapidly gaining in accessibility and popularity (Bolker et al., 2009).

2.4.1. Generalized Linear Mixed Models (GLMMs)

GLMMs are based on simple linear regression (Judd, McClelland, & Ryan, 2009). Simple linear regression finds a linear function that predicts the dependent variable y from an independent variable x as:

$$y = \alpha + \beta x$$

Alpha (α) is the intercept of this model or the value of y for which $x = 0$. For the regression model depicted in Figure 2.1 $\alpha = 0.50$. Beta (β) is the slope of the regression line. It describes the change in y over the change in x . In Figure 2.1 $\beta = 0.85$. This means that when x increases by 1 unit, y increases by 0.85.

Each data point is predicted with a certain error, ϵ . Figure 2.1 shows ϵ_6 , the error for the 6th data point. Simple regression is guided by the least-squares approach, which means the models seek to minimize the sum of squared errors. The squared errors are highlighted in blue in Figure 2.1.

Simple linear regression is the basis for more traditional approaches like t-tests and ANOVAs as well as GLMMs. GLMMs have several advantages over ANOVAs, however. The explicit regression approach makes GLMMs more flexible and easily allows for extensions. Such extensions include the inclusion of mixed effects (General Linear Mixed Models), the prediction of outcome measures that do not follow the normal distribution (Generalized Linear Models) or a combination of both (Generalized Linear Mixed Models) (Judd et al., 2009).

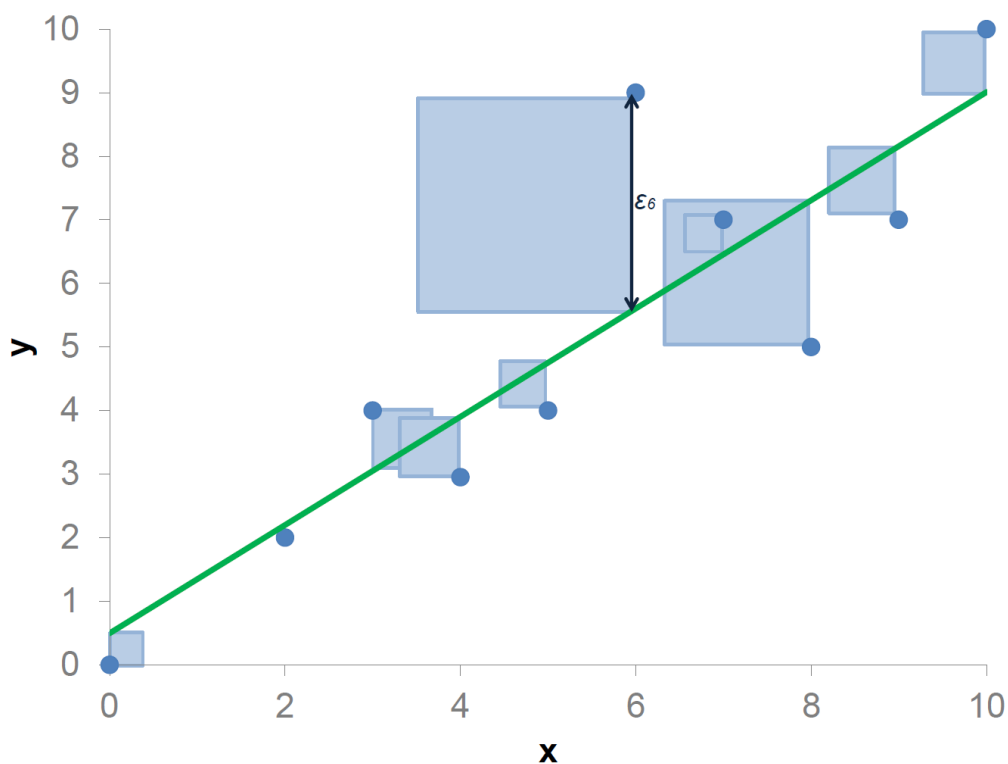


Figure 2.1. *The Regression Approach*. An example dataset with variables x and y , which have been fitted the linear regression line $y = 0.50 + 0.85x$ (shown in green).

Mixed effects models use both random and fixed effects (Bolker et al., 2009). The levels of random effects form only a subset of the population of levels (Venables & Ripley, 2002). Participant is a common random effect. Researchers usually only take a sample of the population they are investigating (adolescents, adults, etc.)

Therefore, the participants included in an experiment do not exhaust the population. Random effects are also often nested and therefore correlated. For instance, each participant is part of a class, which is part of a school. GLMMs can model and account for this correlation by estimating the variance of the random effects (Judd et al., 2009). Fixed effects, in contrast, are variables whose levels exhaust the population of levels (Venables & Ripley, 2002). For instance, the experimental condition may be a variable with two levels: treatment and control. The researcher designed only these two levels, therefore the two levels included in the analysis constitute all possible levels.

A feature of GLMMs that is particularly useful for cognitive researchers is the ability to predict outcome measures whose residuals cannot be modelled with a normal distribution. Most cognitive tasks produce accuracy data, which is dichotomous (1 = correct, 0 = incorrect). This kind of data cannot be modelled with a continuous normal distribution, which is a critical assumption of most linear regression approaches. Traditional approaches like ANOVA therefore require calculating an average score for each participant. GLMMs, however, allow modelling dichotomous outcome variables directly and predicting performance on each trial using a link function (usually the logit function). This function transforms dichotomous data into the unbounded, continuous probability of being correct, using the natural log of the odds. The link function therefore allows the predictors to vary linearly with the outcome variable, even though the outcome variable itself is not linear (Koutoumanou & Wade, 2015b). Using GLMMs instead of calculating averages has the advantage that it will not only increase power to detect effects, but also ensure that data is weighted correctly if not all

participants have completed the same number of trials (Koutoumanou & Wade, 2015b).

GLMMs can now be implemented in most statistical packages. R (R Core Team, 2015) is well supported for GLMMs with packages like lme4 (Bates, Maechler, & Bolker, 2013). R is also open source and uses scripts, which can aid replication efforts.

2.4.2. Modelling age

Age is naturally continuous, but it can be modelled either as a continuous variable or a categorical variable. The most common way of turning continuous variables into categorical ones is using median splits (e.g., as a dummy-coded predictor with 1 = greater than median, 0 = less than or equal to the median). Whether or not it is valid to do so is controversial (Fitzsimons, 2008; Iacobucci, Posavac, Kardes, Schneider, & Popovich, 2015a, 2015b; Irwin & McClelland, 2003; McClelland, Lynch, Irwin, Spiller, & Fitzsimons, 2015; Rucker, McShane, & Preacher, 2015). The arguments on both sides are briefly outlined and some suggestions are provided for how researchers can model age as a continuous or categorical variable in practice.

The main arguments against dichotomizing, or otherwise splitting variables, are that continuous variables are more realistic and increase power. Age, for instance, is naturally continuous. There is no real qualitative difference between being aged 14.9 or 15.0 (Rucker et al., 2015). Using median splits for grouping will also make results sample-dependent, which can hamper generalizability. Finally, splitting

variables results in a loss of power. This can increase both the risk of Type I and Type II errors, i.e. it makes it both harder to find effects and more likely that any effects detected are false-positives (McClelland et al., 2015).

The arguments for splitting continuous variables are mostly pragmatic. While splitting a variable does reduce power somewhat, this is often outweighed by the fact that categorical variables can be easier to interpret (Iacobucci et al., 2015a, 2015b). Iacobucci and colleagues also showed that the risk of false positives is actually not increased by the use of median splits as long as there is no multicollinearity, or high degree of correlation between the factors of the model (Iacobucci et al., 2015a). Multicollinearity can be reduced, for instance, by using orthogonal contrasts coding schemes for categorical predictors (Judd et al., 2009), as was done in this thesis.

If researchers decide to model age as a continuous variable they first need to determine which function age takes: linear, quadratic, cubic, logistic, or something else completely. Plotting data at this stage can help. Model comparisons can also be useful when deciding between nested models. For example, linear models ($y = \alpha + \beta_1 x_1$) are nested within quadratic models ($y = \alpha + \beta_1 x_1 + \beta_2 x_1^2$), which are nested within cubic models ($y = \alpha + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_1^3$). Model comparisons weighing the loss of power by inclusion of more factors with the increase of information explained, can be implemented easily in most packages and help decide between nested models (Judd et al., 2009).

If the function of age is not linear, the use of Generalized Additive Models (GAMs) may be an alternative to GLMMs. GAMs can be used to identify and characterise

non-linear regressions (Hastie & Tibshirani, 2006). However, GAMs are currently not easily specified or interpreted when models contain a mix of continuous and categorical factors and interactions thereof.

Generally speaking, the interpretation of models using age as a categorical rather than continuous variable is often easier if age is a part of an interaction term, as is the case in much developmental research. This is particularly true when age interacts with categorical variables with more than two levels or if the researcher is interested in 3-way interactions or more. For categorical variables, a range of tools are available that allow inspecting specific contrasts and answering questions like ‘Does one age group improve more after training than another?’, ‘Does gender moderate age group differences in motivation?’. Tools available in R include inspecting predefined contrasts within interactions using `lsmeans` (Lenth, 2016) or designing custom contrasts with `multcomp` (Hothorn et al., 2016). See Appendix 2.1 for an example script using these functionalities.

To summarise, this methods chapter discussed advantages and limitations of cross-sectional and longitudinal designs, validity issues such as instrumentation effects, missing data, the design of training studies and the analysis of mixed designs. These issues are relevant to the experimental studies of this thesis, which are presented in the following four chapters.

3. Chapter 3: Perception and Recognition of Faces in Adolescence ⁴

Most studies on the development of face cognition abilities have focussed on childhood, with early maturation accounts contending that face cognition abilities are mature by 3 - 5 years. Late maturation accounts, in contrast, propose that some aspects of face cognition are not mature until at least 10 years. Here, we measured face memory and face perception, two core face cognition abilities, in 661 participants (395 females) in four age groups (younger adolescents (11.27 - 13.38 years); mid-adolescents (13.39 - 15.89 years); older adolescents (15.90 - 18.00 years); and adults (18.01-33.15 years) while controlling for differences in general cognitive ability. We showed that both face cognition abilities mature relatively late, at around 16 years, with a female advantage in face memory, but not in face perception, both in adolescence and adulthood. Late maturation in the face perception task was driven mainly by protracted development in identity perception, while gaze perception abilities were already comparatively mature in early adolescence. These improvements in the ability to memorize, recognize and perceive faces during adolescence may be related to an increasing exploratory behaviour and exposure to novel faces during this period of life.

⁴ The study presented in this chapter has been previously published as:

Fuhrmann, D., Knoll, L.J., Sakhardande, A., Speekenbrink, M., Cohen Kadosh, K. & Blakemore, S-J. (2016). Perception and recognition of faces in adolescence. *Scientific Reports*, 6(33497), doi:10.1038/srep33497

3.1. Introduction

Faces are core to human sociality and changes in face cognition during adolescence may model broader changes in social cognition during this period of life (Adolphs, 2003; Scherf et al., 2012). Face cognition is known to improve between childhood and adulthood (Cohen Kadosh, 2011; Gur et al., 2012; Song et al., 2015). However, there is still controversy as to whether these improvements are domain-specific and environmentally-driven (*late maturation account*; Maurer & Mondloch, 2011) or whether they merely reflect changes in general cognitive abilities, while face cognition abilities themselves are actually mature by 3 - 5 years of age (*early maturation account*; Crookes & McKone, 2009; McKone et al., 2012). See pp. 28 - 30.

Models of face cognition distinguish between two sub-components: *face memory*, the ability to learn and recognize known faces, and *face perception*, the ability to discriminate facial features and configurations (Dolzycka et al., 2014). For face perception, the disagreement between early and late maturation accounts may in part be attributable to the fact that the identity, expression or gaze perception develop at different rates (Cohen Kadosh, 2011). These face aspects are thought to rely on different processing strategies. Previous studies have shown that configural processing strategies are key to recognizing facial identities (Mondloch, Dobson, Parsons, & Maurer, 2004), while featural processing strategies are more often used to determine the direction of gaze (Cohen Kadosh, 2011) and a mix of both is recruited for expression perception (Bombardi et al., 2013). Featural face processing matures early in development (Mondloch et al., 2002). In contrast, configural face processing is more complex, requires much training, and develops

well into the second decade of life (Mondloch et al., 2004). However, when in adolescence the ability to perceive identity, expression and gaze matures, remains uncertain.

It is also not yet clear, whether there are gender differences in face cognition during adolescence. For face memory, adult females generally perform better than males (Heisz, Pottruff, & Shore, 2013; Herlitz & Lovén, 2013; Sommer, Hildebrandt, Kunina-Habenicht, Schacht, & Wilhelm, 2013) and some studies have found a female advantage in face memory across the adolescent age range as well (Gur et al., 2012). However, little is known about gender differences in face perception during adolescence and adulthood.

The aim of the present study was to investigate whether there are age- and gender-related differences in face cognition between adolescence and adulthood, and whether these differences are independent of general cognitive ability.

To this end, we examined face cognition in 661 adolescents and adults aged 11 - 33 (395 females). Participants were split into four age groups: younger adolescents (11.27 -13.38 years); mid-adolescents (13.39 - 15.89 years); older adolescents (15.90 - 18.00 years); and adults (18.01 - 33.15 years). We investigated age-related changes in the two core face cognition abilities, face perception and face memory (Dolzycka et al., 2014; Wilhelm et al., 2007), between adolescence and adulthood. Participants completed the Cambridge Face Memory Test (Duchaine & Nakayama, 2006), which measures the ability to memorise and recognise faces, and the Face Same-Different face perception task (Cohen Kadosh, 2011), which measures the ability to recognise changes in identity, expression or gaze between faces. The task is designed to prevent

participants from using a simple strategy, such as always focussing on the eyes by changing only one of these face aspects at a time, and never a mix of aspects.

We investigated three hypotheses: *(I) developmental differences*: Based on previous developmental studies (Gur et al., 2012; Song et al., 2015), we predicted that face cognition abilities would improve from early adolescence to adulthood and investigated whether these developmental patterns differed between face memory and perception. We also examined *(II) gender differences* in face cognition: Based mainly studies in adults (Gur et al., 2012; Heisz et al., 2013; Herlitz & Lovén, 2013; Sommer et al., 2013), we predicted that females would perform better than males. Finally, we probed developmental differences in the perception of different *(III) face aspect differences* (identity, expression and gaze). Based on previous studies on face aspect cognition (Bombari et al., 2013; Cohen Kadosh, 2011; Cohen Kadosh, Johnson, Dick, Cohen Kadosh, & Blakemore, 2013; Cohen Kadosh, Johnson, Henson, et al., 2013) and configural and featural processing (Mondloch et al., 2004; Mondloch et al., 2002), we predicted that the ability to perceive changes gaze would mature earlier than changes in identity.

3.2. Methods

3.2.1. Participants

We recruited 821 participants from 16 local schools in the London area (adolescents) and through University College London (UCL) participant pools and posters (adults). The present study analysed cognitive performance at baseline of a larger training study (see Chapter 5). Of this sample, 661 participants were

included in the current analysis ($M_{\text{age}} = 16.21$ years, $SD_{\text{age}} = 4.12$, age range = 11.27 - 33.15 years, 397 females). Exclusion criteria were: missing baseline data ($N = 3$); missing parental consent for adolescents ($N = 123$); report of developmental conditions including ADHD and dyslexia ($N = 34$). Adolescents were split into three age groups of equal size and adults were included as a fourth age group (Table 3.1). We chose three age groups as a compromise between increased sensitivity with increasing numbers of groups and the loss of power this engenders. Adults were qualitatively different from the other groups (e.g. not tested in schools) and were therefore included as a fourth group.

Table 3.1. Demographic Information

Age group	Age range	Gender	<i>N</i>	<i>M_{RR}</i>	<i>SD_{RR}</i>
younger adolescents	11.27-13.38	female	118	0.63	0.16
		male	67	0.57	0.18
mid-adolescents	13.39-15.89	female	89	0.68	0.15
		male	96	0.68	0.17
older adolescents	15.90-18.00	female	109	0.73	0.13
		male	77	0.72	0.15
adults	18.01-33.15	female	81	0.81	0.12
		male	24	0.79	0.17

Note. M_{RR} = mean relational reasoning accuracy at T1; SD_{RR} = standard deviation of relational reasoning accuracy at T1. The relational reasoning task was based on Raven's matrices, a standard measure of IQ and general cognitive ability (Knoll et al., 2016; Raven, 2009). Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

General cognitive ability differed between age groups as indexed by accuracy in a non-verbal matrix reasoning task similar to those included in IQ tests (Knoll et al., 2016; Raven, 2009). Reasoning scores increased with age and there were significant differences between all age groups ($F(3,654) = 42.28$, $p < .001$; see Table 3.1 and Table 3.2). Reasoning scores were therefore included as a covariate

in all analyses to control for age-group differences in general cognitive ability (see section 3.2.6).

Table 3.2. Comparisons of Relational Reasoning Accuracy between Age Groups

Contrast	Estimate	SE	t	df	p	
young adolescents vs. mid-adolescents	-0.07	0.02	-4.52	654	< .001	***
young adolescents vs. old adolescents	-0.12	0.02	-7.52	654	< .001	***
young adolescents vs. adults	-0.20	0.02	-10.68	654	< .001	***
mid-adolescents vs. old adolescents	-0.05	0.02	-2.96	654	.019	*
mid-adolescents vs. adults	-0.13	0.02	-6.73	654	< .001	***
old adolescents vs. adults	-0.08	0.02	-4.30	654	< .001	***

Note. * $p < .001$, ** $p < .01$, *** $p < .05$. Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

The study was carried out in accordance with UCL Research Ethics Guidelines and was approved by the UCL Research Ethics Committee. Informed consent was obtained from all participants aged 18 and over, and parents of participants under the age of 18. Assent was obtained from all participants under 18.

3.2.2. Experimental design

A 4 x 2 x 2 design with age group (younger adolescents, mid-adolescents, older adolescents, adults) and gender as between-subjects factors, and task (face perception, face memory) as a within-subjects factor was employed. For the face perception task, face aspect was investigated as an additional within-subjects factor with three levels (identity, expression, gaze).

3.2.3. Testing procedure

Participants were tested on a battery of tasks including two face cognition tasks: a face perception and a face memory task (see section 3.2.4 and 3.2.5). Testing was carried out using an online platform developed by the research team and a software company (www.cauldron.sc).

Participants completed the test session in groups of 3 - 48 in school (adolescents) or in a university computer room (adults), using laptops, tablets or desktop computers. Responses were made using a mouse or touchscreen. Task order was counterbalanced using a Latin-square design between testing groups of participants and across test sessions.

An experimenter gave instructions before each task. Participants then completed practice trials until three were completed correctly. Participants were given visual feedback on their performance in the practice trials only. For the face memory task, 22 participants completed more than 3 practice trials. On average, these participants needed 5.2 practice trials to proceed to the task and never more than 7. For the face perception task, 173 participants completed more than 3 practice trials. On average, these participants needed 4.4 practice trials to proceed to the task and never more than 8. All participants completed three practice trials successfully during the test session and proceed to the task. Participants were given visual feedback on their performance in the practice trials only.

3.2.4. Face memory task

An adaptation of the Cambridge Face Memory Task (CFMT) was used to assess the ability to learn and recognise unknown faces using a 3-alternative forced choice (3-AFC) test (Duchaine & Nakayama, 2006). Participants were asked to memorise target faces and then find a target face in a panel of three faces. There was only ever one target face in the panel of three, the other two were distractor faces that had not been memorised (Figure 3.1).

The task took 9 min or 54 trials to complete, whichever came first. The task was shortened from the original 72 trials (Duchaine & Nakayama, 2006) due to time restrictions in schools. The three blocks of the original CFMT were preserved but block two and three were shortened to match the number of trials in block one. Adults' accuracy in our adaptation of the CFMT was 82.02% ($SD = 7.63$) and similar to adults' performance in the original CFMT, in which accuracy was 80.4% ($SD = 11.0$). The third block was repeated if participants finished all trials before the time limit but data from these repeated trials were not included in the analysis.

A set of 126 face stimuli matching the specifications of the original CFMT was created for the purpose of the present study and the larger training study, comprising three test sessions in total. Photographs of 42 Caucasian males from three angles (frontal/left quarter profile/right quarter profile) were obtained from the Facial Recognition Technology database (Phillips, Moon, Rizvi, & Rauss, 2000). Black and white images were cropped to exclude external features of the face (hair etc.) using the GNU Image Manipulation Program (GIMP Team, 2013). The size of each face was standardized to 180 x 245 pixels and luminance of the image was set to a value of 110 using GIMP's Levels function.

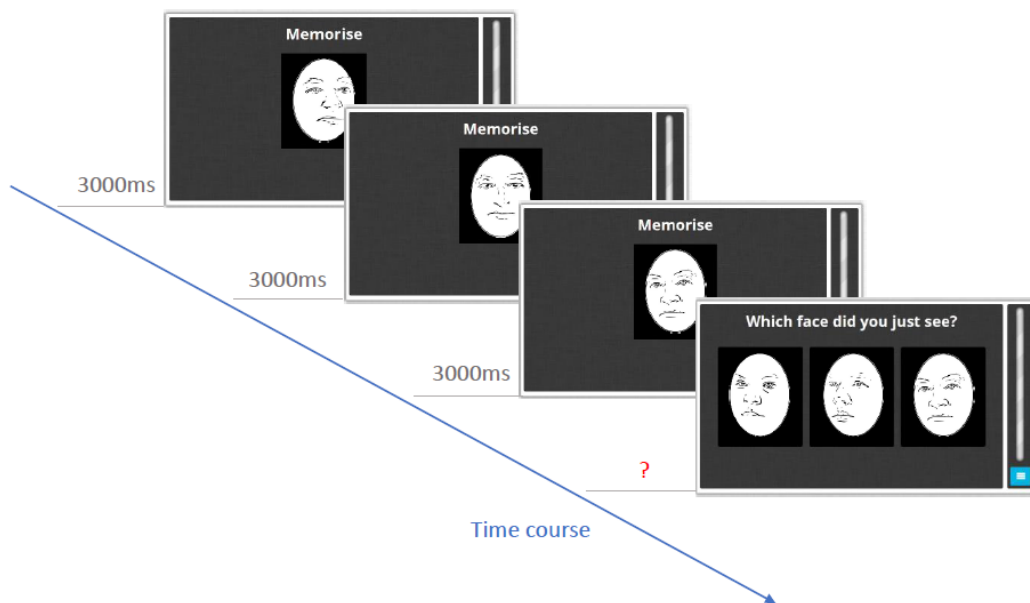


Figure 3.1. Cambridge Face Memory Task. The CFMT (Duchaine & Nakayama, 2006) consisted of three blocks. In the first block (shown here), a target face was shown at three different angles, for 3 s each, and this was followed by three 3-AFC trials. This procedure was repeated for five more target faces. In the second block, frontal views of the same six target faces were presented simultaneously for 20 s, and this was followed by eighteen 3-AFC trials. In the third block, frontal views of the same six target faces were presented simultaneously for 20 s, but a 50% Gaussian noise mask was added to the faces in the eighteen 3-AFC trials that followed. Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

3.2.5. Face perception task

The face perception task measured the ability to process featural and configural changes in faces (Cohen Kadosh, 2011). Participants were asked to decide whether two consecutively presented faces presented were the same or different

(Figure 3.2). Faces were considered to be different with regard to changes in any of the following face aspects: gaze direction (left/right), expression (happy/sad) or identity (person A/person B). Participants were informed that faces should be classified as the same only if they are exactly the same across all three face properties.

A test session took 7.5 min or 48 trials to complete, whichever came first. 16 out of 658 participants completed fewer than 48 trials. They still completed 89.7% of trials on average and never less than 62.5%. If participants finished the 48 trials within the 7.5min time limit, the set of faces was presented again, but these data were not included in the analysis. In half the trials, faces were the same; in the other half, faces differed (24 trials). In the trials in which faces differed, one third showed changes in expression, one third in identity and one third in gaze (8 trials per aspect). Trial difficulty was varied by adding noise masks of increasing strength (25 - 81% in 8% steps), except the first two trials, which had a noise mask of 25%.

Photos of two female, Caucasian faces were taken under standardised lighting conditions for the purpose of this experiment. Four photos were obtained for each face: happy expression-gaze left, happy expression-gaze right, sad expression-gaze left, and sad expression-gaze right. Using the GNU Image Manipulation Program (GIMP Team, 2013), coloured photos were scaled to a uniform size and cropped to exclude external features of the face (hair etc.) Using GIMP's Levels function, the lightness of the image and mean RGB values were standardized (Luminance: 105, R: 105, G: 75, B: 70). Task difficulty was increased

by presenting the images with a Gaussian noise mask of varying strength (25%, 33%, 41%, 49%, 57%, 65%, 73% or 81% noise).

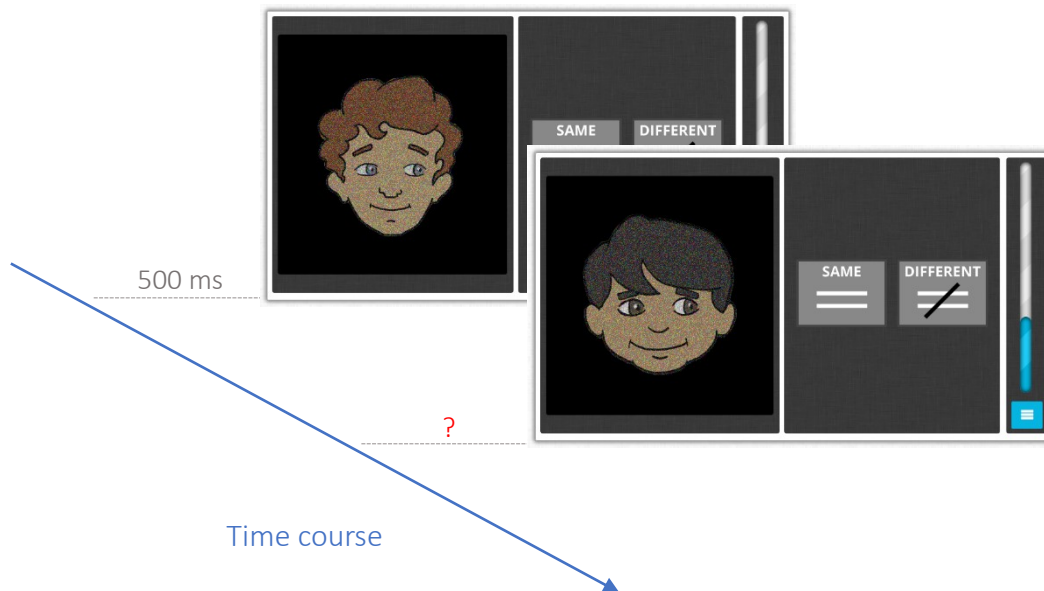


Figure 3.2. Face Perception Task. Procedure for the face perception task (Cohen Kadosh, 2011). Screen shots show stimuli from practice trials. Each trial started with a fixation cross presented for 800 ms, followed by the first face presented for 500 ms, then another fixation cross for 800 ms, and then the second face until a response was logged or 5000 ms passed, whichever came first. The two possible response options (same/different) were shown simultaneously with the presentation of the two faces. The next trial started immediately after the participant had entered their response. The trial displayed here shows an identity change. Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

3.2.6. Statistical analysis

GLMMs were implemented in R (R Core Team, 2015), lme4 (Bates et al., 2013) and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2016) to assess differences in task performance between age groups. Trials with a response time under 250 ms were excluded from the analysis. To assess age group differences in face cognition, a logistic model predicting correct/incorrect responses (accuracy) and a general linear model predicting response times in correct trials were computed. Orthogonal, Helmert-coded fixed effects included were age group, task and gender as well as all possible 2-way interactions and the 3-way interaction. Z-scored performance in the relational reasoning task for each participant was included as covariate to control for differences in general cognitive ability between age groups (see Table 3.1 and Table 3.2). Nested random intercepts for participant ID and school/university were used to reflect the repeated-measure design and clustered nature of participants tested. Planned comparisons of age group differences were carried out using lsmeans (Lenth, 2016) and Bonferroni-adjusted for multiple comparisons. To investigate the effect of face aspect on performance in the face perception task, two separate models predicting accuracy and response times in correct trials were computed. Age group and face aspect as well as their interaction were included as fixed effects. The covariate and random effects were computed as described above. Custom contrasts were computed using package multcomp (Hothorn et al., 2016) to investigate differences in performance dependent on face aspect and comparing differences between age groups within face aspects. These contrasts were Bonferroni-corrected for multiple comparisons as well. Finally, one model for accuracy and one for response times in face perception were computed, which were identical

to the models described above except for the fact that face aspect was dummy-coded with gaze perception as the reference group. This allowed inspection of contrasts of the interaction of face aspect with age group using the summary() function.

3.3. Results

3.3.1. The development of face memory and face perception

To assess whether developmental trajectories differed between face memory and face perception, we pooled the data for both tasks, generating an overall index of face cognition ability. We then investigated age effects in overall face cognition before determining whether these age effects were moderated by task (face perception versus face memory).

There were significant differences in overall face cognition accuracy between age groups ($\chi^2(3) = 24.40, p < .001$). Younger and mid-adolescents performed significantly worse than the two older age groups in face cognition (Figure 3.3A). There were also significant differences in response times on correct trials between age groups ($\chi^2(3) = 18.46, p < .001$). Response times were significantly slower in younger adolescents than in all older age groups. However, the contrast between mid-adolescents and the older age groups did not reach significance for response times (Figure 3.3B). The age effects did not differ between tasks (accuracy: $\chi^2(3) = 4.47, p = .215$; response times: $\chi^2(3) = 4.41, p = .220$) indicating that face memory and perception followed similar developmental trajectories.

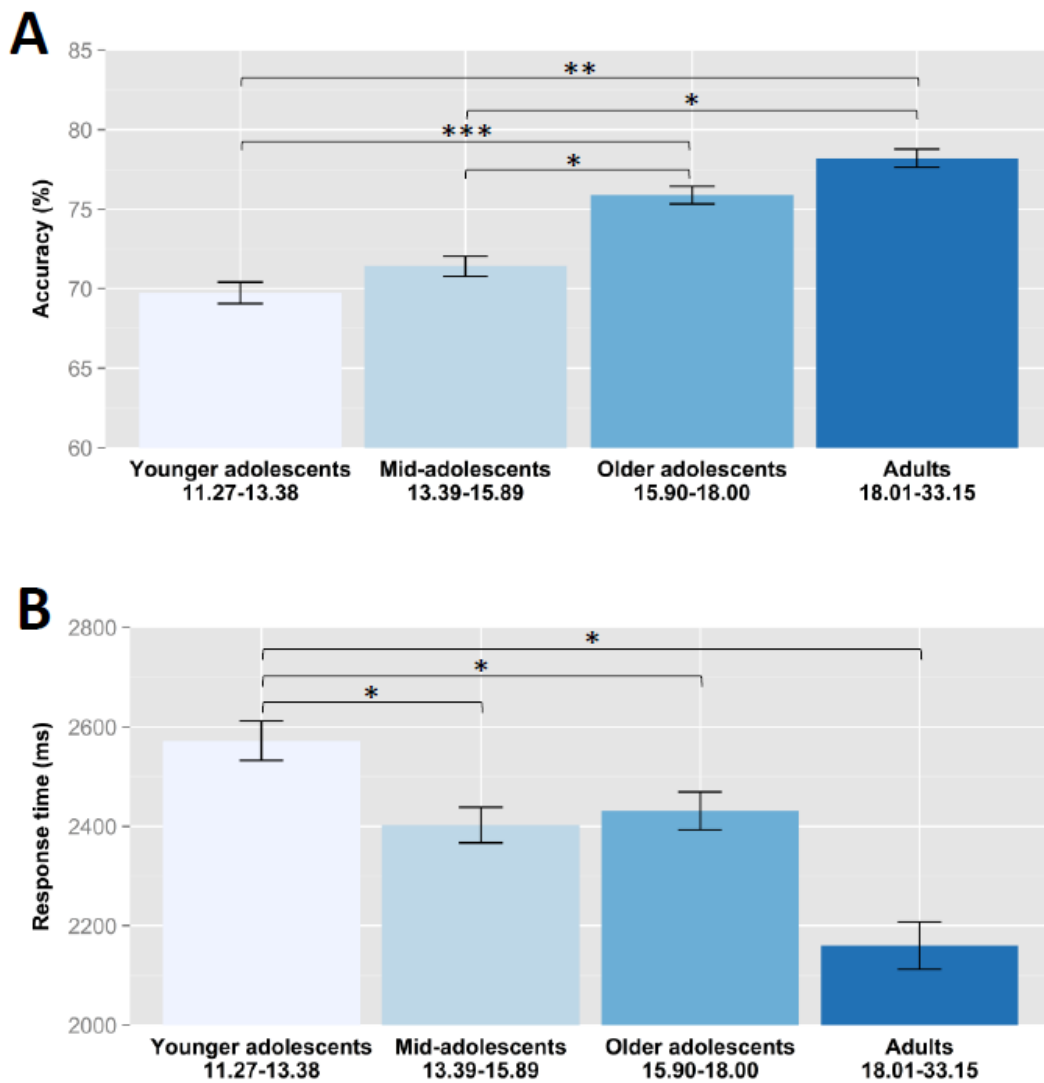


Figure 3.3. Face Cognition Performance. Panel A shows accuracy and panel B response times in four age groups: younger adolescents (11.27 - 13.38 years), mid-adolescents (13.39 - 15.89 years), older adolescents (15.90 - 18.00 years) and adults (18.01 - 33.15 years). Results are shown averaged over the face processing and face memory task with standard error bars (* $p < .05$, ** $p < .01$, *** $p < .001$). Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

3.3.2. Gender differences in face memory and face perception

There were no differences in reasoning scores between genders ($F(1, 654) = 2.52$, $p = .113$), indicating that general cognitive ability was matched. There was a main effect of gender on overall face cognition accuracy ($\chi^2(1) = 13.48$, $p < .001$) but not response times ($\chi^2(1) = 0.17$, $p = .682$). For both dependent measures, there was an interaction between gender and task (accuracy: $\chi^2(1) = 8.00$, $p = .005$; response times: $\chi^2(1) = 9.01$, $p = .003$) indicating that the effect of gender differed between face memory and face perception. For accuracy, females outperformed males in face memory but not face perception. Response times followed the same pattern (Table 3.3). There was no significant interaction between age group and gender (accuracy: $\chi^2(3) = 1.10$, $p = .776$; response times: $\chi^2(3) = 2.47$, $p = .481$), indicating that developmental trajectories did not differ between genders.

Table 3.3. Face Cognition Performance Overall

Measure	Task	Female	Male	Comparison
Accuracy	Face memory	0.79 (0.11)	0.77 (0.13)	$z = 4.54$, $p < .001$
	Face perception	0.70 (0.09)	0.68 (0.09)	$z = 1.83$, $p = .068$
Response times (ms)	Face memory	3144.53 (880.56)	3337.60 (935.46)	$t(1276.96) = -2.29$, $p = .022$
	Face perception	1626.21 (294.6)	1585.31 (299.01)	$t(1277.21) = 1.67$, $p = .095$

Note. Mean performance is shown. SD is shown in brackets. Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

3.3.3. Development of face aspect perception

The face perception task measures the ability to correctly detect changes in three face aspects: identity, expression and gaze. We investigated whether there were differences in the ability to process these face aspects and whether the developmental trajectories for identity, expression and gaze perception differed.

Accuracy

There was a main effect of face aspect on accuracy ($\chi^2(2) = 313.96, p < .001$). Participants performed significantly better in gaze perception ($M = 0.72, SD = 0.17$) than in identity perception ($M = 0.55, SD = 0.22; z = 16.59, p < .001$) and expression perception ($M = 0.58, SD = 0.20; z = 14.32, p < .001$). The difference between identity and expression perception did not survive correction for multiple comparisons ($z = 2.39, p = .051$).

The effect of face aspect was moderated by age group for accuracy ($\chi^2(6) = 14.84, p = .022$), indicating that the ability to correctly identify changes in identity, expression or gaze differed between age groups. There were developmental differences in identity perception such that younger adolescents were less accurate than older adolescents and adults, and mid-adolescents were less accurate than adults. Younger adolescents were also less accurate than older adolescents in expression perception while there were no developmental differences in gaze perception (Figure 3.4A).

To assess whether developmental differences in identity and expression perception were significantly greater than in gaze perception, we inspected contrasts. The difference between younger adolescents and the older age groups was significantly greater in identity than in gaze perception (Table 3.4), the difference between mid-adolescents and the older age groups was also significantly greater in identity than gaze perception. All other comparisons, including all comparisons between gaze and expression perception were not significant, indicating that developmental effects were mainly restricted to identity perception.

Table 3.4. Comparisons of Face Perception Accuracy between Age Groups and Face Aspects

Contrast	Estimate	SE	z	p
younger adolescents vs. older: gaze vs. expression	0.00	0.02	0.05	.963
mid-adolescents vs. older: gaze vs. expression	-0.06	0.04	-1.56	.118
older adolescents vs. older: gaze vs. expression	0.07	0.07	1.04	.299
younger adolescents vs. older: gaze vs. identity	-0.05	0.02	-2.23	.026 *
mid-adolescents vs. older: gaze vs. identity	-0.09	0.04	-2.58	.010 *
older adolescents vs. older: gaze vs. identity	0.01	0.07	0.16	.872

Note. * $p < .05$. Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

Response times

Face aspect also affected response times ($\chi^2(2) = 73.44, p < .001$). Following the same pattern as accuracy, participants responded significantly faster in gaze perception ($M = 1596.92, SD = 351.95$) than in identity perception ($M = 1761.81, SD = 483.19; t(1294.22) = -8.55, p < .001$) or expression perception ($M = 1691.43, SD = 428.05; t(1293.61) = -4.75, p < .001$). They were also quicker in expression than identity perception ($t(1294.64) = -3.08, p < .001$).

The effect of face aspect was not moderated by age group for response times ($\chi^2(6) = 7.58, p = .022$). There were similar developmental differences for all face aspects. Younger adolescents were significantly slower than all older age groups for all three face aspects, and mid-adolescents were significantly slower than adults in identity perception (Figure 3.4B). Developmental differences between younger adolescents and the older age groups were not stronger in identity or expression perception than gaze perception (Table 3.5).

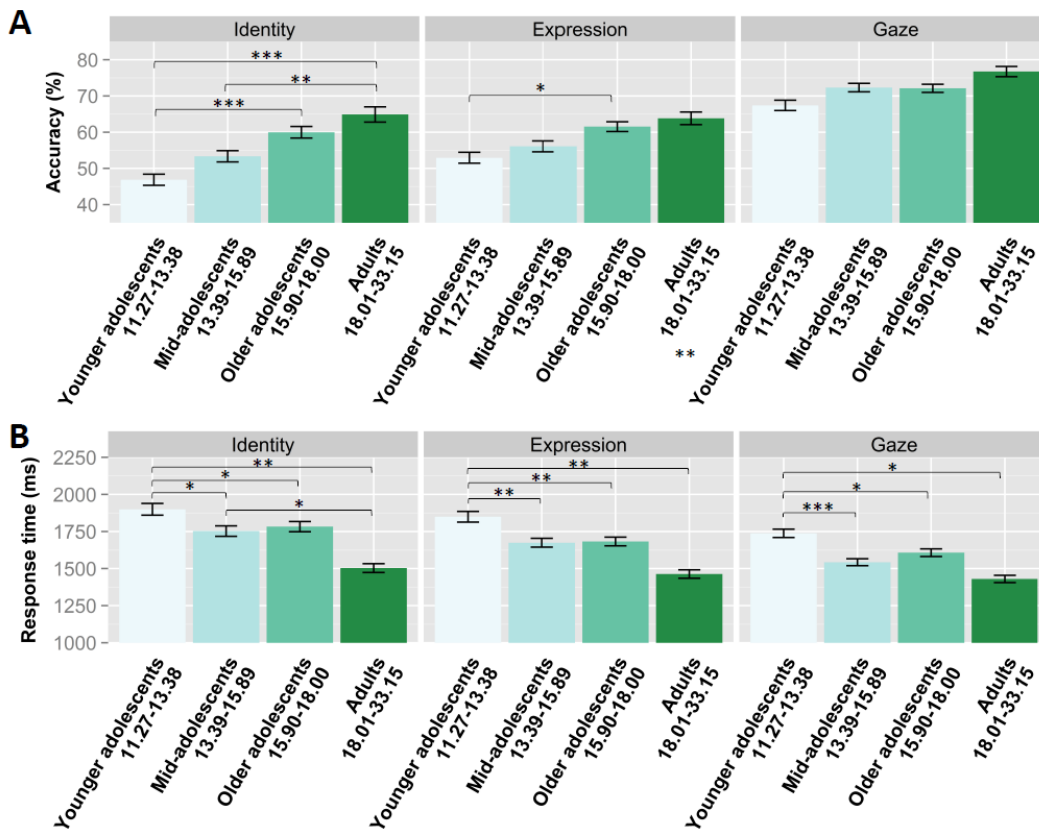


Figure 3.4. Face Perception Performance by Face Aspect. Panel A shows accuracy and panel B response times in four age groups: younger adolescents (11.27 - 13.38 years), mid-adolescents (13.39 - 15.89 years), older adolescents (15.90 - 18.00 years) and adults (18.01 - 33.15 years). Results are shown for identity, expression and gaze perception with standard error bars (* $p < .05$, ** $p < .01$, *** $p < .001$). Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

Table 3.5. Comparisons of Face Perception Speed between Age Groups and Face Aspects.

Contrast	Estimate	SE	df	t	p
younger adolescents vs. older: gaze vs. expression	8.45	10.04	1291.80	0.84	.400
mid-adolescents vs. older: gaze vs. expression	24.02	14.59	1293.40	1.65	.100
older adolescents vs. older: gaze vs. expression	20.94	27.75	1289.20	0.76	.451
younger adolescents vs. older: gaze vs. identity	2.77	10.06	1293.00	0.28	.783
mid-adolescents vs. older: gaze vs. identity	27.50	14.58	1292.90	1.89	.060
older adolescents vs. older: gaze vs. identity	52.98	27.79	1290.20	1.91	.057

Note. Adapted from Fuhrmann et al. (2016) with permission from Nature Publishing Group.

In summary, there were improvements with age in identity perception - both in accuracy and speed - while expression and gaze perception improved with age in speed only.

3.4. Discussion

The results from this large scale, cross-sectional study demonstrate that face cognition undergoes protracted development: older adolescents' and adults' face memory and face perception abilities were more proficient than those of younger adolescents and mid-adolescents. Most studies on the development of face cognition abilities have focussed on early and mid-childhood, with many studies suggesting that face cognition abilities are mature by 3 - 5 years of age (McKone et al., 2012). However, some previous studies have shown improvements from early adolescence to adulthood - a finding that we replicated here (Germine, Duchaine, & Nakayama, 2011; Gur et al., 2012; Song et al., 2015).

Our study extended previous evidence by showing that the main period of face cognition development in adolescence is roughly between 11 - 16 years. This was the case for both face memory and face perception, the two core components of face cognition. General cognitive ability (as measured by matrix reasoning) predicted face cognition scores but showed a different developmental trajectory with continuous improvements throughout adolescence and into adulthood. Developmental differences in face cognition also persisted after controlling for general cognitive ability. Our results thus indicate that the development of face cognition in adolescence is not solely due to improvements in general cognitive

ability (McKone et al., 2012). To further probe the domain-specificity of face cognition development, future studies will need to directly compare memory and perception of face- and non-face-objects.

The effects of gender differed between the two sub-components of face cognition, with a female advantage for face memory but not for face perception. This pattern did not differ between age groups. Previous studies have shown a female advantage in face memory in adults (Heisz et al., 2013; Herlitz & Lovén, 2013; Sommer et al., 2013) and in adolescents (Gur et al., 2012), which we replicated in the current study. Some studies have also found a female face perception advantage in adults, which was not replicated here (Sommer et al., 2013). The female advantage in face cognition is thought to be driven by the fact that female participants scan face stimuli more than males (Heisz et al., 2013). One explanation for why we found gender differences in face memory, but not in face perception, is, that increased face scanning by females may have led to gender differences in the task in which face stimuli were presented for a long period of time (up to 20s in the case of the face memory task), and precluded gender differences in the task where the stimuli were only presented for a short time (500 ms for the face perception task). Gender differences in face cognition in adults are thought to be partly explained by greater social interest and involvement in females compared to males (Sommer et al., 2013), but this remains to be tested in adolescents.

An inspection of the ability to recognise changes in the three face aspects manipulated in the face perception task – identity, expression and gaze – revealed that developmental effects in the face perception task were driven mainly by age

group differences in identity perception. Younger adolescents and mid-adolescents were less accurate than older adolescents and adults in identity perception, but not gaze perception. This supports the late maturation account of face cognition by showing not only quantitative differences between adolescent age groups but also qualitative differences, with identity perception maturing only in mid-adolescence. This finding also matches models of early maturation of featural versus late maturation of configural perception (Cohen Kadosh, 2011; Mondloch et al., 2002). Perceiving identity changes requires mainly configural perception, whereas perceiving a change in gaze direction recruits purely featural perception (Cohen Kadosh, 2011; Cohen Kadosh, Johnson, Henson, et al., 2013).

Mid-adolescence may be of particular importance, not only for the development of face cognition, but also for the development of social cognition in general. Developmental models (Scherf et al., 2012) and empirical evidence (McGivern et al., 2002) indicate a perturbation of face cognition with the onset of puberty. The ensuing period of rapid cognitive and neurological development may provide an ideal substrate for social learning (Blakemore & Mills, 2014; Fuhrmann, Knoll, & Blakemore, 2015). Exploratory behaviour in adolescence (Casey et al., 2008) may lead to more exposure to novel faces than earlier in life and new social roles in adolescence may increase the focus on facial information such as attractiveness and status (Scherf et al., 2012). This may then provide the environmental enrichment necessary for becoming a face expert. This interpretation fits with the perceptual expertise account of face perception (Bukach, Gauthier, & Tarr, 2006). Extensive experience with a specific category of objects, not just faces, is thought to lead to more efficient mental representations - perhaps through 'holistic'

encoding. The concept of holistic encoding is similar to configural perception and describes the ability to process a stimulus as a whole rather than the sum of its parts (Piepers & Robbins, 2012).

In conclusion, face memory and face perception abilities mature relatively late in development, between early and late adolescence. Improvements in face perception over adolescence are driven by increased identity perception abilities. These improvements over adolescence may be related to changes in adolescents' social environment and an increased exposure to novel faces during this period of life. The next chapter directly probes how the social environment affects adolescents by simulating social exclusion in the lab.

4. Chapter 4: Cognitive Performance after Social Exclusion in Adolescence

It has been suggested that adolescence is a sensitive period of social development. During this time of life, social exclusion may have a particularly detrimental effect on mood and psychological needs. However, little is known about how social exclusion affects cognitive performance in adolescence. The aim of this study was to test whether social exclusion reduces cognitive performance in adolescence more than it does in adulthood. To this end, we recruited 99 females in three age groups: young adolescents (N = 36, aged 10.1 - 14.0), mid-adolescents (N = 36, aged 14.3 - 17.9) and adults (N = 27, aged 18.3 - 38.1). Social exclusion was simulated using the Cyberball paradigm, a computerized ball-tossing game. Following inclusion and exclusion by virtual peers in the game, participants completed a mood questionnaire and two measures of cognitive performance: an n-back verbal and a dot-matrix visuo-spatial working memory task. All age groups showed reductions in mood after exclusion. Young adolescents also showed reduced accuracy in the n-back task and increased response times in the dot-matrix task following exclusion. In contrast, mid-adolescents' and adults' cognitive performance was not significantly affected by exclusion. These results suggest that young adolescent girls' cognitive performance is particularly susceptible to the adverse effects of a short, virtual social exclusion experience.

4.1. Introduction

Adolescence is traditionally thought of as a time of social reorientation during which peers become increasingly important (Crone & Dahl, 2012; Steinberg &

Monahan, 2007). It has been suggested that adolescence may even be a sensitive period, during which the brain is particularly susceptible to socio-cultural information (Blakemore & Mills, 2014). As such, negative social experiences, such as social exclusion, may be especially detrimental for adolescents (Buwalda, Geerdink, Vidal, & Koolhaas, 2011; Fuhrmann et al., 2015). See pp. 30 - 34 and section 1.3.3.

Social exclusion can be simulated in the lab using the Cyberball paradigm (Williams et al., 2000). Cyberball is an online ball-tossing game during which participants are seemingly either included or excluded by two peers. In adults, the exclusion condition consistently lowers mood and induces a threat to four fundamental psychological needs: self-esteem, belonging, control and a sense of meaningful existence (Williams, 2007; Williams et al., 2000). Such effects have been found to be heightened in adolescence in some studies (Pharo et al., 2011; Sebastian et al., 2010).

Social exclusion may affect not only mood and need-threat but also cognitive performance (Baumeister et al., 2002). Studies using the Cyberball paradigm in adults have mostly found negative effects of exclusion on cognitive functioning in adults, particularly on executive functions such as inhibitory control and working memory (Jamieson et al., 2010; Themanson et al., 2014). Cyberball has also been shown to disrupt cognitive performance in children aged 8 to 12 (Hawes et al., 2012). To date, however, there is little experimental evidence on how social exclusion affects cognitive performance in adolescence.

The aim of this study was to investigate the effects of experimentally-induced social exclusion on cognitive performance during adolescence, so as to gain a

better understanding of the ramifications of social exclusion in schools. To this end, we compared the impact of social inclusion and exclusion in 99 female adolescents and adults. We recruited only one gender because of sex differences in pubertal development during adolescence that may cause differences in cognitive development (Sisk & Zehr, 2005). We chose to recruit females because adolescent girls have been found to spend more time with peers than boys (Larson & Richards, 1991), potentially making peer-rejection more relevant to them. In childhood, girls have also been shown to be more sensitive to social exclusion than boys (Hawes et al., 2012).

Participants were divided into three age groups: young adolescents ($N = 36$, aged 10.1 - 14.0), mid-adolescents ($N = 36$, aged 14.3 - 17.9) and adults ($N = 27$, aged 18.3 - 38.1). Adolescent participants were divided into two, even-sized age groups because previous research suggested that younger adolescents may respond differently to Cyberball than mid-adolescents (Sebastian et al., 2010).

Participants experienced the inclusion and exclusion condition in the Cyberball game. After each Cyberball condition, participants completed a mood questionnaire, as well as verbal and visuo-spatial working memory tasks. We chose working memory tasks as indicators of cognitive performance because working memory is educationally relevant and predicts fluid intelligence (Engle et al., 1999; Kane et al., 2004) and academic performance (Alloway et al., 2009; Gathercole et al., 2003). Using working memory tasks also allowed us to compare our results to previous Cyberball studies using working memory and other executive function tasks in children and adults (Hawes et al., 2012; Jamieson et al., 2010; Themanson et al., 2014). We assessed both verbal (n-back) and visuo-

spatial (dot-matrix) working memory as these are proposed to rely on separable systems (Baddeley, 2003; Baddeley & Hitch, 1974). Assessing both components can therefore give a more complete picture of working memory performance.

We hypothesized that social exclusion would reduce n-back and dot-matrix task performance across age groups, and that this effect would be higher in the adolescent groups compared with adults. In line with previous studies, we also expected that social exclusion would be associated with lower mood in all age groups, and that effects would be stronger in both adolescent groups than in adults.

4.2. Methods

4.2.1. Participants

A total of 113 female participants aged 10 - 38 years were recruited for the purpose of this study. Adolescent participants were recruited from seven secondary schools (two state, four private and one grammar) in Greater London, Peterborough and Oxfordshire and tested individually in schools. Adult participants were recruited from UCL participant pools and tested in the lab. A researcher tested all participants individually in a quiet room. Seven participants were excluded from all analyses because they reported psychiatric or developmental disorders, one because they scored below 70 IQ points, one because of technical difficulties during testing and five because they didn't believe the Cyberball manipulation (see section 4.2.2). The remaining 99 participants were allocated to one of three age groups: young adolescents, mid-adolescents

and adults (Table 4.1). School-aged participants were allocated to the two adolescent age groups via a median-split to ensure similar sample sizes in both groups. Participants over the age of 18 were allocated to the adult group.

Table 4.1. Participant Characteristics

Age group	N	Age		IQ	
		Min	Max	M	SE
Young adolescents	36	10.1	14.0	101.38	2.06
Mid-adolescents	36	14.3	17.9	97.90	1.94
Adults	27	18.3	38.1	108.02	2.25

We examined IQ differences between age groups as a potential confound. IQ, as measured by matrix reasoning tests (WASI; Wechsler, 1999), did not differ significantly between young adolescents and mid-adolescents ($t(92) = 1.23$, $p = .667$), or young adolescents and adults ($t(92) = -2.18$, $p = .096$). It did, however, differ significantly between mid-adolescents and adults ($t(92) = -3.41$, $p = .003$). Because of this difference, IQ was controlled for in all analyses (see section 4.2.4).

The study was carried out in accordance with UCL Research Ethics Guidelines and was approved by the UCL Research Ethics Committee. Informed consent was obtained parents of participants under 18 and assent from participants themselves. Adult participants consented to taking part in the study.

4.2.2. Materials

Cyberball

Social inclusion and exclusion were simulated using the freeware Cyberball 4.0 program (Williams et al., 2000). This program features two virtual players who played an online ball-tossing game lasting ~2 min with the participant (Figure 4.1).

Whilst participants were told the other players were real and connected to them via the internet, the Cyberball players were in fact programmed to either include or exclude the participant from the game. Inclusion generated one third of the ball tosses to the participant. Exclusion generated only two tosses to the participant at the beginning of the game, after which the other players no longer threw the participant the ball.



Figure 4.1. Cyberball Game. Screenshot from Cyberball 4.0 (Williams, Yeager, Cheung, & Choi, 2012). The participant saw an icon for herself at the bottom of the screen ('You'). When she received the ball, she could click on one of the other two players ('Kate' or 'Emma') to throw the ball to them.

To check whether participants believed the other players to be authentic, we carried out a three-question probe during debrief (Will, van Lier, Crone, & Güroğlu, 2016):

1. What did you think of the Cyberball game?
2. How did you like being connected to other people though the internet?

3. What did you think the study was about?

We recorded whether or not participants verbalized suspicion about authenticity during this probe. Five participants (two mid-adolescents and three adults) voiced suspicion that the other players were not real, and were therefore excluded from all analyses.

Working memory measures

All participants completed two different measures of working memory: an n-back verbal working memory task and a dot-matrix visuo-spatial working memory task. The order of these tasks was counterbalanced between participants. The first 20 participants also completed a digit span task. This task was then cut from the procedure because of time constraints in schools, and data from this task was not analysed. All tasks were programmed in Cogent (Cogent 2000 Team, 2015) and MATLAB (The MathWorks, 2013) and accuracy (correct/incorrect) and response times for each task were recorded.

N-back task. In the n-back verbal working memory task (Gevins & Cutillo, 1993), numbers were flashed one-by-one on a screen for 500 ms with a variable delay in between (1000-3000 ms, mean delay: 2000 ms). The task required participants to indicate whether the current number on the screen was i) a zero (0-back task) or ii) the same as the number that appeared "two back" in the sequence (2-back task). Distractors were shown simultaneously with the number. Distractors consisted of photos of a house, a happy face or a fearful face. They appeared on both sides of the number and were included to vary the affective context of the task.

Participants were instructed to ignore them. Participants completed six blocks of 12 trials each. Half of these blocks were 0-back tasks, half were 2-back tasks. The order of blocks and response buttons was counterbalanced between participants.

Dot-matrix task. The dot-matrix task is a visuo-spatial working memory task (Alloway et al., 2009). Participants were shown a four-by-four white grid on a black background. Dots were flashed one-by-one for 300 ms and with a 600 ms delay in between. Dots were displayed in any of the 16 squares of the grid. After all dots in a particular sequence were shown, the grid turned orange for 1500 ms, and then turned white again. Participants were instructed to click on the fields of the grid where the dots had appeared; and in the order they had appeared. Sequence length increased from three to eight dots. Three sequences of each length were shown.

Questionnaire measures

Participants were administered a standard mood and need-threat questionnaire after each Cyberball condition (Williams et al., 2000). We analysed the mood questionnaire here in which participants rated how good/bad, happy/sad, friendly/unfriendly and tense/relaxed they were currently feeling, on a scale of 1 - 5. We calculated an average mood rating for each participant and each Cyberball condition.

4.2.3. Procedure

Participants practised the two working memory tasks at the beginning of the experiment. They were then introduced to Cyberball. All participants played Cyberball twice and experienced both inclusion and exclusion. The order of the Cyberball conditions was counterbalanced between participants. Participants completed the mood questionnaire and two working memory tests after each Cyberball condition. Participants were then fully debriefed. The experiment took ~60 min in total.

4.2.4. Design and analysis

We used a 2 x 3 mixed design with Cyberball condition (inclusion/exclusion) as the within subjects measure and age group as the between subjects measure (young adolescent/mid-adolescent/adult).

The data was analysed using GLMMs in R (R Core Team, 2015) and lme4 (Bates et al., 2013). For each of the working memory tasks, we specified one model for accuracy and one for response times. Accuracy was analysed as a binary dependent variable (correct/incorrect) and modelled using the binomial distribution. Response times and mood ratings were each averaged over Cyberball condition for each participant and analysed as continuous, dependent variables. In all models, Cyberball condition, age group and the interaction between the two were specified as orthogonal, Helmert-coded fixed effects. IQ was included as a z-scored covariate and participant number as a random intercept. We carried out planned comparisons of the fixed effects in these models using lsmeans (Lenth,

2016) and multcomp (Hothorn et al., 2016) and Bonferroni-corrected for three comparisons each.

In a supplementary analysis we specified four additional models predicting accuracy and response times for the n-back and dot-matrix task each. For the n-back task, models were specified as described above but additionally included task difficulty (0-back/2-back) and distractor type (happy face/fearful face/house) as orthogonal, Helmert-coded fixed effects. For the dot-matrix task, two models were specified as described above but task difficulty (low: 3 - 5 dots / high: 5 - 8 dots) was included as an additional factor. All models also included all possible interactions between the fixed effects.

4.3. Results

4.3.1. Working memory performance

We assessed working memory performance after inclusion and exclusion in the Cyberball game. Accuracy and response times in the n-back verbal working memory task and dot-matrix visuo-spatial working memory task were analysed using GLMMs.

N-back task

Accuracy. There was no main effect of Cyberball condition for n-back accuracy ($\chi^2(1) = 0.17, p = .678$), indicating that there was no overall difference in performance between inclusion and exclusion (Table 4.2).

Table 4.2. Overall Performance in the N-Back and Dot-Matrix Task

Cyberball condition	N-back				Dot-matrix			
	Accuracy (%)		RT (ms)		Accuracy (%)		RT (ms)	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Inclusion	93.32	0.08	792.30	160.28	67.89	0.12	3530.90	390.37
Exclusion	93.90	0.06	782.30	159.46	68.69	0.12	3423.95	369.49

Note. RT = response times. Model-predicted values are shown.

However, there was a significant interaction between Cyberball condition and age group ($\chi^2(2) = 7.97, p = .019$). Young adolescents were the only age group that showed reduced accuracy after exclusion compared to inclusion. Mid-adolescents and adults showed no significant difference between exclusion and inclusion (Figure 4.2; Table 4.3). The reduction in accuracy in young adolescents was significantly greater than differences between exclusion and inclusion in mid-adolescents. The difference between young adolescents and adults did not survive correction for multiple comparison (Figure 4.2; Table 4.4).

There was a significant 3-way interaction between Cyberball condition, age group and task difficulty ($\chi^2(2) = 7.94, p = .019$). This indicated that the age differences in response to exclusion were moderated by task difficulty (0-back or 2-back). Post-hoc tests showed that young adolescents demonstrated a greater reduction in performance in response to exclusion on the 0-back than the 2-back task ($z = -3.07, p = .007$) (Table 4.5). There was no difference between the 0- and 2-back task for any of the other age groups (mid-adolescents: $z = 1.10, p = .810$; adults: $z = 0.07, p = 1$). These age-group differences were unlikely to be due to ceiling effects as all age groups performed significantly below 100% (Table 4.5). There was no

significant interaction between Cyberball condition, age group and distractor type (happy face, fearful face, or house; $\chi^2(4) = 4.38, p = .358$).

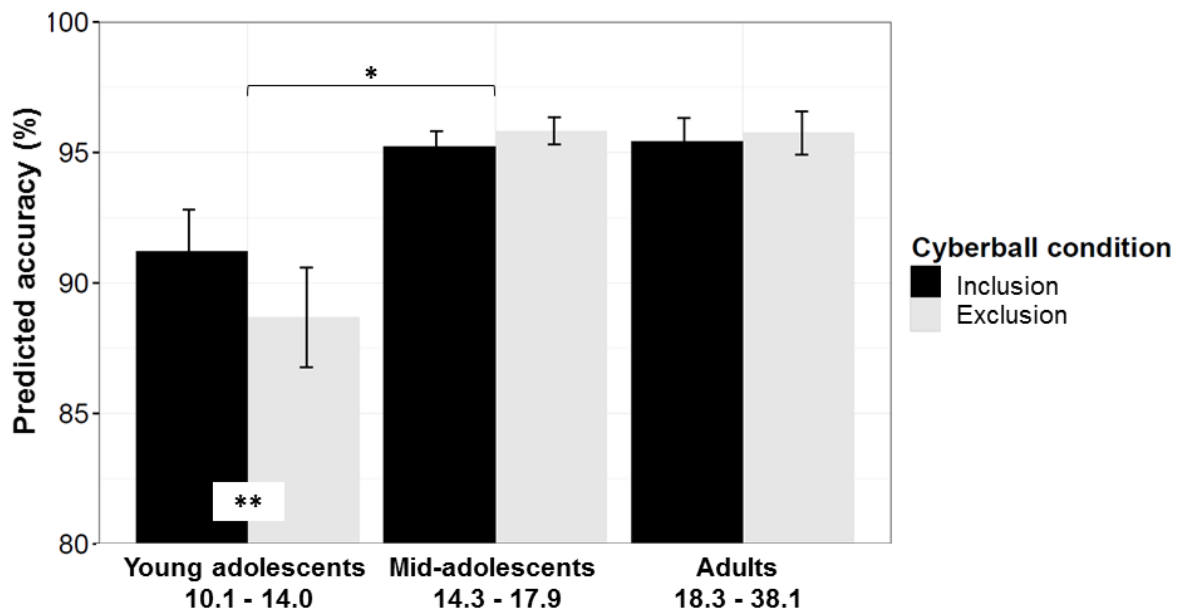


Figure 4.2. *N*-back Accuracy after Inclusion and Exclusion. Mean accuracy with standard error bars are shown for three age groups: young adolescents, mid-adolescents and adults. All values are model-predicted. Asterisks at the bottom of the bars in white boxes indicate significant differences between Cyberball conditions for a particular age group. Asterisks above the bars indicate that such effects differed between age groups (* $p < .05$).

Response times. There was no main effect of Cyberball condition for *n*-back response times ($\chi^2(1) = 0.19, p = .667$), indicating that overall, there was no difference in performance between exclusion and inclusion (Table 4.2). There was also no significant interaction between Cyberball condition and age group ($\chi^2(2) = 3.27, p = .195$) and planned comparisons showed that there were no significant differences between Cyberball exclusion and inclusion for any of the three age groups (Table 4.3; Table 4.4). There was also no significant interaction between Cyberball condition, age group and

task difficulty ($\chi^2(2) = 3.01, p = .222$) or distractor type ($\chi^2(4) = 2.51, p = .644$).

Table 4.3. Cognitive Performance and Mood Compared between Exclusion and Inclusion within Age Groups

N-back accuracy					
<i>Contrast</i>	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>	
Young adolescents	-0.31	0.11	-2.96	.003	**
Mid-adolescents	0.13	0.14	0.93	.353	
Adults	0.08	0.16	0.50	.616	
N-back response times					
<i>Contrast</i>	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>p</i>	
Young adolescents	-0.17	15.09	-0.01	.991	
Mid-adolescents	25.67	15.26	1.68	.096	
Adults	13.86	16.43	-0.84	.401	
Dot-matrix accuracy					
<i>Contrast</i>	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>	
Young adolescents	-0.01	0.13	-0.07	.943	
Mid-adolescents	-0.08	0.13	-0.64	.522	
Adults	-0.02	0.14	-0.14	.887	
Dot-matrix response times					
<i>Contrast</i>	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>p</i>	
Young adolescents	235.21	114.21	2.06	.042	*
Mid-adolescents	104.06	104.65	0.99	.323	
Adults	-48.41	122.62	-0.40	.694	
Mood ratings					
<i>Contrast</i>	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>p</i>	
Young adolescents	-1.57	0.17	-9.20	<.001	***
Mid-adolescents	-1.45	0.16	-9.02	<.001	***
Adults	-1.42	0.19	-7.49	<.001	***

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

Table 4.4. Cognitive Performance and Mood Compared between Exclusion and Inclusion between Age Groups

N-back accuracy				
Contrast	Estimate	SE	z	p
Young adolescents vs. mid-adolescents	-0.45	0.18	-2.50	.037 *
Young adolescents vs. adults	-0.39	0.19	-2.06	.119
Mid-adolescents vs. adults	0.05	0.22	0.25	1
N-back response times				
Contrast	Estimate	SE	z	p
Young adolescents vs. mid-adolescents	-25.84	21.45	-1.21	.685
Young adolescents vs. adults	13.69	22.30	0.61	1
Mid-adolescents vs. adults	39.53	22.42	1.76	.233
Dot-matrix accuracy				
Contrast	Estimate	SE	z	p
Young adolescents vs. mid-adolescents	0.07	0.18	0.40	1
Young adolescents vs. adults	0.01	0.19	0.06	1
Mid-adolescents vs. adults	-0.06	0.19	-0.33	1
Dot-matrix response times				
Contrast	Estimate	SE	z	p
Young adolescents vs. mid-adolescents	131.10	154.90	0.85	1
Young adolescents vs. adults	283.60	167.50	1.69	.271
Mid-adolescents vs. adults	152.50	161.20	0.95	1
Mood ratings				
Contrast	Estimate	SE	z	p
Young adolescents vs. mid-adolescents	-0.12	0.23	-0.51	1
Young adolescents vs. adults	-0.15	0.26	-0.58	1
Mid-adolescents vs. adults	-0.03	0.25	-0.12	1

Note. * $p < .01$

Table 4.5. Mean Accuracy in the 0-Back Task

Age group	Cyberball condition	Mean accuracy (%)	SE (%)	One-sample t-test (comparing mean to 100%)		
				t	df	p
Young adolescents	Exclusion	91.75	17.44	-2.84	35	.045 *
Young adolescents	Inclusion	95.21	7.29	-3.94	35	.002 **
Mid-adolescents	Exclusion	98.04	2.59	-4.35	32	< .001 ***
Mid-adolescents	Inclusion	97.19	2.94	-5.57	33	< .001 ***
Adults	Exclusion	97.36	3.99	-3.44	26	.012 *
Adults	Inclusion	97.06	3.56	-4.29	26	.001 **

Note. P-values smaller than .05 suggest that accuracy was significantly below 100% (* $p < .05$, ** $p < .01$, *** $p < .001$).

Dot-matrix task

Accuracy. There was no main effect of Cyberball condition for dot-matrix accuracy ($\chi^2(1) = 0.23, p = .629$), indicating that overall performance was matched between exclusion and inclusion (Table 4.2). The interaction between Cyberball condition and age group was not significant ($\chi^2(2) = 0.18, p = .912$). Accuracy did not differ significantly between inclusion and exclusion for any age group (Table 4.3; Table 4.4). There was also no significant 3-way interaction between Cyberball condition, age group and task difficulty (low: 3 - 5 dots / high: 6 - 8 dots) for dot-matrix accuracy ($\chi^2(2) = 2.37, p = .305$).

Response times. There was no main effect of Cyberball condition for dot-matrix response times ($\chi^2(1) = 2.17, p = .141$), indicating that overall performance did not differ between exclusion and inclusion (Table 4.2). The interaction between Cyberball condition and age group was not significant ($\chi^2(2) = 2.87, p = .239$). However, planned comparisons showed an increase in response times after exclusion compared to inclusion in young adolescents (Figure 4.3; Table 4.3; Table 4.4). There was no significant 3-way interaction between Cyberball condition, age group and task difficulty ($\chi^2(2) = 1.58, p = .453$).

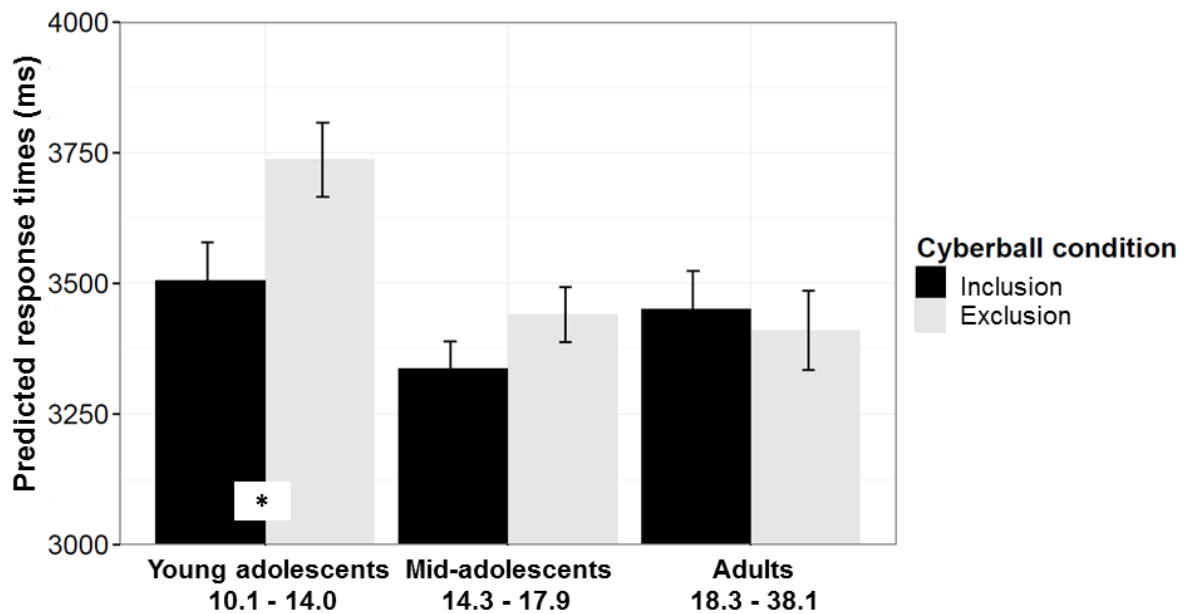


Figure 4.3. Dot-Matrix Response Times after Inclusion and Exclusion. Mean response times with standard error bars are shown for three age groups: young adolescents, mid-adolescents and adults. All values are model-predicted. Asterisks in white boxes at the bottom of the bars indicate significant differences between Cyberball conditions for a particular age group. None of the comparisons between age groups were significant (* $p < .05$).

4.3.2. Mood ratings

We analysed participants' mood ratings after inclusion and exclusion in the Cyberball game using GLMMs. There was a significant main effect of Cyberball condition on mood ($\chi^2(1) = 216.70, p < .001$). Mood was lower after exclusion ($M = 2.51, SE = 0.08$) than inclusion ($M = 3.99, SE = 0.08$) overall. This effect did not differ between age groups ($\chi^2(2) = 0.40, p = .818$), however (Figure 4.4; Table 4.3; Table 4.4).

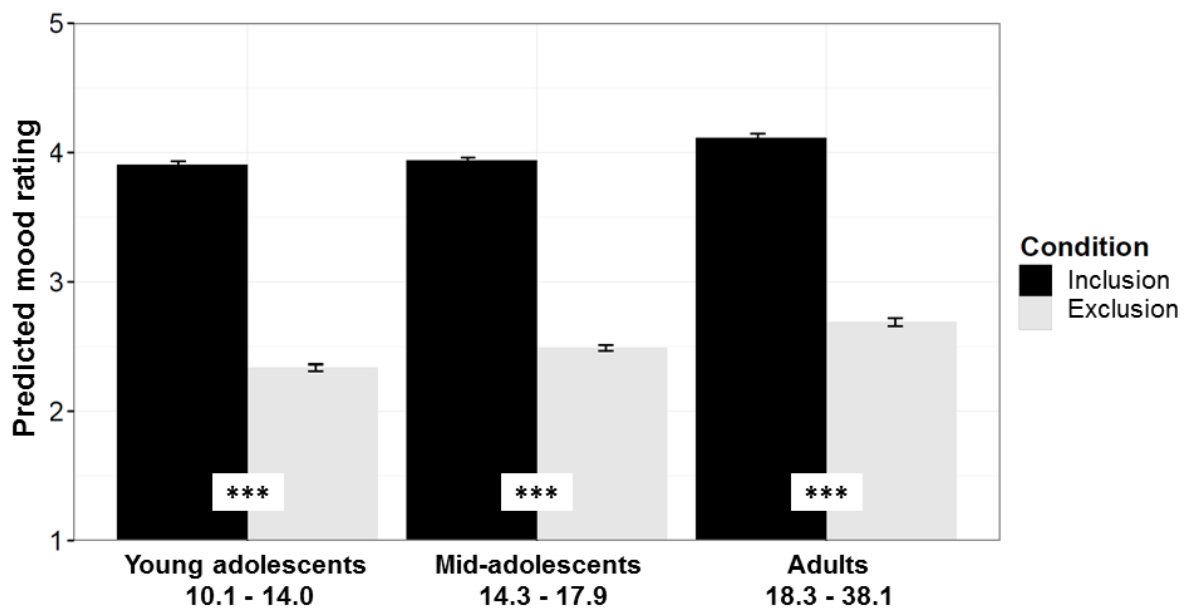


Figure 4.4. Mood Ratings after Inclusion and Exclusion. Mean ratings with standard error bars are shown for three age groups: young adolescents, mid-adolescents and adults. All values are model-predicted. Asterisks in white boxes at the bottom of the bars indicate significant differences between Cyberball conditions for a particular age group. None of the comparisons between age groups were significant (***) $p < .001$).

4.4. Discussion

Here we investigated the impact of social exclusion on cognitive performance and mood in three age groups: young adolescents (aged 10.1 - 14.0), mid-adolescents (aged 14.3 - 17.9) and adults (aged 18.3 - 38.1). While all age groups showed a similar and significant reduction in mood after social exclusion, the effect of exclusion on cognitive performance was age-dependent. Only young adolescents showed a reduction in performance on n-back and visuo-spatial working memory tasks after social exclusion; this was not the case for mid-adolescents or adults.

Previous research showed negative effects of Cyberball exclusion on executive functions in adults (Jamieson et al., 2010; Lustenberger & Jagacinski, 2010; Themanson et al., 2014) and working memory in children (Hawes et al., 2012). Therefore, we hypothesized that we would see reductions in cognitive performance after exclusion in all age groups, but expected the effects to be more pronounced in adolescence. While cognitive performance was reduced after exclusion in younger adolescents, we found no effect of social exclusion on cognitive performance in mid-adolescents or adults. It is possible that the effects of social exclusion in older populations depend on the specific executive function tasks used. Executive functions may decompose into rule-driven, explicit; and internalized, automatic processes (Crone & Steinbeis, 2017; Olsson & Ochsner, 2008). Effects in adults may be evident mostly when tasks require internalized, automatic executive function processes such as inhibition in the anti-saccade task (Jamieson et al., 2010) or the Flanker task (Themanson et al., 2014). Such automatic processes may be difficult to adjust to situational demands and might be easily disrupted by stressful situations. In contrast, our cognitive tasks were rule-driven and explicit and may be more adaptable under changing situational demands. This explanation is speculative at present, however, and remains to be tested in future research.

Contrary to our hypothesis, we found no significant age group differences in mood: all age groups showed similar significant reductions after social exclusion. While this finding is dissimilar to some previous studies on mood, anxiety and need-threat (Pharo et al., 2011; Sebastian et al., 2010), it is in line with a recent meta-analysis of 120 Cyberball studies. This meta-analysis showed that exclusion

generally has large ($d > |1.4|$) effects on intrapersonal outcome measures such as self-esteem and that these effects are mostly independent of age (Hartgerink, van Beest, Wicherts, & Williams, 2015).

This pattern of results makes it unlikely that the age-dependent effects of Cyberball on cognitive performance reported here were due to age differences in the emotional (mood) response to Cyberball. All three age groups tested showed similar mood reductions after exclusion and yet cognitive performance was affected in young adolescents only. This indicates that the age-dependent effects of Cyberball were relatively specific to cognitive performance. This finding is in line with previous studies in adults showing that the effects of social exclusion on cognitive performance are not mediated by mood (Baumeister et al., 2002; Buckley, Winkel, & Leary, 2004). Instead, self-regulatory processes such as suppression of ruminative thought, or active down-regulation of unwanted affect may be candidate mechanisms for the reduction in cognitive performance. These self-regulatory processes are thought to interact and compete with executive functions (Hofmann, Schmeichel, & Baddeley, 2012).

The effect of social exclusion on cognitive performance in young adolescents was evident across two measures of working memory, even though it manifested somewhat differently. For n-back performance, accuracy was reduced but response times did not increase after exclusion. For dot-matrix performance, the opposite pattern emerged: response times increased but accuracy was not reduced. One possible explanation for this pattern of results is that the n-back required quick responses, while responses were self-paced in the dot-matrix task. Participants may have increased their response times in the dot-matrix task to

maintain accuracy levels after exclusion while that was not possible in the n-back task. Overall, the effect was more robust for n-back performance: the increase in dot-matrix response times in young adolescents after exclusion was not significant at the interaction-level.

A supplementary analysis showed that the effect of exclusion was stronger for easy trials (0-back) than for difficult trials (2-back) in the n-back task. This finding is similar to the results of a previous study in which girls aged 8 - 12 also showed reduced cognitive functioning on easy but not hard working memory tasks (Hawes et al., 2012). One possible explanation for this pattern of results is that the easier 0-back trials may have allowed for more rumination than the demanding 2-back trials. Rumination, in turn, is known to increase the emotional impact of Cyberball exclusion (Wesselmann, Ren, Swim, & Williams, 2013) and also to disrupt cognitive performance (Curci, Lanciano, Soleti, & Rimé, 2013; Hofmann et al., 2012). There was no difference between easy and hard trials in the visuo-spatial working memory test. This task was self-paced, which may have allowed for some rumination on all types of trials.

There are some limitations of the sample included here, which need to be taken into account when interpreting the results. First, IQ differed between age groups. We think it unlikely, however, that IQ could explain stronger performance reductions after social exclusion in younger adolescents, as IQ was controlled for in all analyses. Furthermore, IQ did not differ significantly between young adolescents and the other two age groups, only between mid-adolescents and adults. It should be noted, however, that the difference between young adolescents and adults was approaching significance ($p = .080$). Second, our study

included female participants only and we do not know whether our results would generalize to males. Adolescent girls may spend more time with peers than boys (Larson & Richards, 1991) and may be particularly sensitive to social exclusion (Hawes et al., 2012). It is therefore important for future research to explore whether adolescent boys react differently to social exclusion than adolescent girls.

Overall, our results indicate high susceptibility of young adolescent girls to the effects of a short virtual social exclusion experience. This adds to previous research showing that children aged 8 - 12 were similarly affected by social exclusion (Hawes et al., 2012). These findings are relevant to understanding the effects of ostracism in schools. It highlights that experiencing social exclusion may place a particular burden on young girls. Exclusion reduces cognitive performance, which, in turn, may impact educational achievement (Nakamoto & Schwartz, 2010; Rigby, 2000; Sharp, 1995). This underlines the need to develop effective ostracism interventions in schools and to consider age differences in response to social exclusion in the design and timing of interventions.

This hypersensitivity to social exclusion in late childhood and early adolescence is in line with rodent studies, which have shown a sensitive period for social isolation during the late juvenile and early adolescent stage (Einon & Morgan, 1977). Future research could investigate effects of social exclusion in humans across a broader age-range to explore whether there is a peak of sensitivity to social exclusion in late childhood and early adolescence, as there is in rodents (Buwalda et al., 2011; Einon & Morgan, 1977), or whether sensitivity to exclusion simply decreases over development.

The study in this, and the preceding chapter, suggest that adolescence may be a period of high susceptibility to environmental input. If this is the case, we would not only expect increased vulnerabilities, however, but also predict opportunities for development. This line of inquiry is followed in the next chapter, which investigates whether some age groups benefit more from cognitive training than others.

5. Chapter 5: Age-Dependent Effects of Cognitive Training⁵

In the current study, we investigated windows for enhanced learning of cognitive skills during adolescence. Six hundred and sixty-three participants (11 - 33 years old) were divided into four age groups, and each participant was randomly allocated to one of three training groups. Each training group completed up to 20 days of online training in numerosity discrimination (i.e., discriminating small from large numbers of objects), relational reasoning (i.e., detecting abstract relationships between groups of items), or face perception (i.e., identifying differences in faces). Training yielded some improvement in performance on the numerosity discrimination task, but only in older adolescents or adults. In contrast, training in relational reasoning improved performance on that task in all age groups, but training benefits were greater for people in late adolescence and adulthood, than for people earlier in adolescence. Training did not increase performance on the face perception task for any age group. Our findings suggest that for certain cognitive skills, training during late adolescence and adulthood yields greater improvement than training earlier in adolescence, which highlights the relevance of this late developmental stage for education.

5.1. Introduction

Education policy tends to emphasise the importance of investing in early childhood intervention. This argument is partly based on well-established economics accounts of the added value of early childhood intervention (Heckman,

⁵ The study presented in this chapter has been previously published as:

Fuhrmann, D.*, Knoll, L.J.*, Sakhardande, A., Stamp, F., Speekenbrink, M. & Blakemore, S-J. (2016). A window of opportunity for cognitive training in adolescence. *Psychological Science*, 27(12), 1620-1631. doi: 10.1177/0956797616671327 *Joint first authors.

2000, 2006). However, there is a tension between the assumption that earlier is always better for learning, and studies showing that the human brain continues to develop throughout childhood, adolescence and into early adulthood.

Research has shown that several cortical regions in humans undergo protracted structural and functional development across adolescence (Cohen Kadosh, Johnson, Dick, Cohen Kadosh, & Blakemore, 2013; Giedd & et al., 1999; Giedd & Rapoport, 2010; Lebel, Walker, Leemans, Phillips, & Beaulieu, 2008; Tamnes et al., 2010). Regions that undergo particularly substantial development include the prefrontal and parietal cortices, which are involved in a variety of higher cognitive skills relevant to mathematics education, including reasoning and numerical skills (Blakemore & Robbins, 2012; Dumontheil, 2014; Houdé, Rossi, Lubin, & Joliot, 2010). There is evidence that protracted development of these cognitive skills occurs during adolescence (Crone, Wendelken, Donohue, van Leijenhorst, & Bunge, 2006; Dumontheil, Houlton, Christoff, & Blakemore, 2010; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Tamnes et al., 2013). See pp. 24 - 27 and section 1.3.2. However, little is known about when these skills are most efficiently learned.

Here, we trained performance of three cognitive skills: numerosity discrimination, relational reasoning and face perception. These cognitive skills were chosen because they involve brain regions that undergo development during adolescence (Cohen Kadosh, Johnson, Dick, et al., 2013; Dehaene, Piazza, Pinel, & Cohen, 2003; Dumontheil, Houlton, et al., 2010) and because they improve during adolescence (Dumontheil, Houlton, et al., 2010; Fuhrmann et al., 2016; Halberda et al., 2012). Therefore, these skills might be expected to be particularly trainable

in adolescence. In addition, both relational reasoning and numerosity discrimination are relevant to education. Numerosity discrimination is correlated with mathematics performance (Halberda et al., 2012), and relational reasoning is also related to fluid intelligence, a significant predictor of educational outcomes (Chuderski, 2014).

Face perception, the ability to identify changes in faces and facial features, was included as the control training task. Face perception also improves during adolescence and may be susceptible to training, but it relies on different cognitive processes and neural circuits than those involved in numerosity discrimination and relational reasoning (Cohen Kadosh, Johnson, Dick, et al., 2013; Cohen Kadosh, Johnson, Henson, et al., 2013). We thus reasoned that there would be no transfer from face perception training to numerosity discrimination and relational reasoning performance, or vice versa. Including a face perception training group also allowed us to control for non-specific aspects of participating in a training study such as adhering to a training schedule, online training over several days and so on (Klingberg, 2010).

Each of the three training tasks was tested before (test session one: T1) and immediately after training (test session 2: T2), and between three to nine months after training (test session three: T3) (Figure 5.1). In addition, we included two non-trained tasks in each test session: a working memory task (backward digit span) and a face memory task, in order to determine whether transfer effects were evident, and whether they differed between age groups.

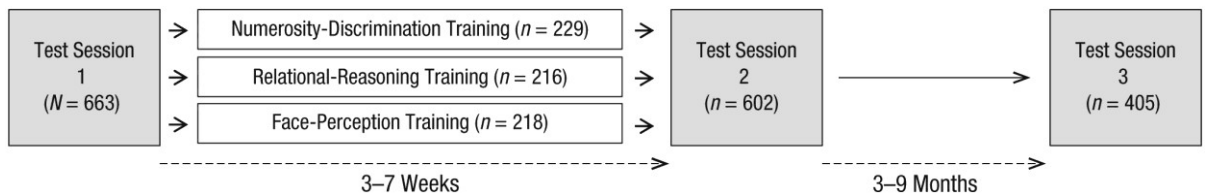


Figure 5.1. *Timeline of the Study.* ND = numerosity discrimination, RR = relational reasoning, FP = face perception, T1 = baseline testing, T2 = testing after training, T3 = follow-up testing. Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

We compared training effects between four age groups: 186 younger adolescents (11.27 - 13.38 years), 186 mid-adolescents (13.39 - 15.89 years), 186 older adolescents (15.90 - 18.00 years), and 105 adults (18.01 - 33.15 years). We investigated three central hypotheses: (I) *General training effects*: Training would improve performance on the trained task only; (II) *Age-dependent training effects*: Performance on the trained task would improve after training within some or all age groups and the strength of improvement would differ between age groups; (III) *Transfer effects*: Training effects might generalise to performance on a non-trained task that involves similar cognitive processes. Specifically, training in relational reasoning might lead to improvements in performance on an untrained working memory task (Klingberg, 2010), and training in face perception might lead to improvements in performance on an untrained face memory task (Dolzycka et al., 2014).

5.2. Methods

5.2.1. Participants

Participants recruited for the study described in Chapter 3, were tested on a number of tasks as part of a cognitive training study. As described in Chapter 3, Data from 821 participants was collected over a 16-month period. Adolescents were recruited from 16 schools in and around London. Adults were recruited through the UCL participant pools (which are databases that include individuals who are not students and have not previously studied at UCL) and through posters in central London, near the university. School-age participants were tested during lessons, and data were collected from all students present in the classroom. Data from 123 students was excluded because parental consent was not provided. Participants' data was also excluded if they reported a diagnosis of developmental conditions, including attention-deficit/hyperactivity disorder, autism, dyscalculia, dyslexia, and epilepsy ($N = 34$), or if they were not present during testing at Test Session 1 ($N = 1$). The final sample at Test Session 1 included 663 participants (398 females; $M_{age} = 16.50$ years, $SD_{age} = 4.42$, age range = 11.27 - 33.15 years) and was divided into four age groups: younger adolescents, mid-adolescents, older adolescents, and adults (Table 5.1). To create the three adolescent age groups, we sorted the 11 - to 18 - year - olds by age and then split them into three bins of equal size. We chose three age groups for adolescents as a compromise between the increased sensitivity that comes with increasing numbers of groups and the loss of power this produces. Adults were tested separately from adolescents and were assigned to their own age group.

Participants were randomly assigned to one of three training groups: numerosity discrimination, relational reasoning and face perception training (see Table 5.1. for group sizes and gender split). Experimenters were blind to participants' training group. We tested whether training groups and age groups differed in a number of potential confounds: the amount of training completed; days between training sessions; days between T1 and T2; days between T2 and T3; group size at testing, number of test sessions split over multiple days and missing data at T2 and T3. None of the training groups differed on any of these variables, but there were age group differences in all of them. We therefore carried out supplementary analyses to test whether these potential confounds with age influenced our main results (see section 5.3.4).

Table 5.1. Participant Numbers

Age group		Numerosity discrimination training group			Relational reasoning training group			Face perception training group		
		T1	T2	T3	T1	T2	T3	T1	T2	T3
Younger adolescents 11.27 - 13.38	N overall	62	57	37	61	56	38	63	58	43
	N female	41	39	26	32	30	26	45	42	31
Mid-adolescents 13.39 - 15.89	N overall	60	57	38	63	61	46	63	59	46
	N female	30	28	21	33	33	23	27	25	22
Older adolescents 15.90 - 18.00	N overall	71	60	42	57	49	33	58	43	25
	N female	41	37	30	33	30	21	35	26	14
Adults 18.01 - 33.15	N overall	36	36	17	35	34	22	34	32	18
	N female	28	28	14	25	24	15	28	27	17

Note. Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

5.2.2. Experimental design

Participants were tested at three time points: T2 occurred three to seven weeks after T1, and T3 occurred three to nine months after T2 (Figure 5.1). Between T1 and T2, participants were asked to complete 20 sessions of online training on one of three training tasks (numerosity discrimination, relational reasoning or face perception). Participants were tested on five tasks at each test session: numerosity discrimination, relational reasoning, face perception, face memory and backward digit span. The face memory and backward digit span tasks were included to investigate transfer effects between the trained tasks and related tasks. See below for details of each of the tasks.

5.2.3. Testing procedure

Testing and training were carried out using an online platform developed by the research team and Cauldron, a software company (<http://www.cauldron.sc>). Participants completed each of the three test sessions in groups; adolescents were tested in school and adults were tested in a university computer room (for average group sizes per age group, see Table 5.7). Participants used laptops, tablets, or desktop computers. Responses on all five tasks were made using a mouse, touchpad, or touchscreen. Before each task, an experimenter gave instructions, and participants completed practice trials until they correctly completed three trials on each of the five tasks. Participants were given visual feedback on their performance in the practice trials only. Task order was counterbalanced among training groups and across test sessions using a Latin-square design. Because of school scheduling constraints, Test Session 1 was split

over 2 or 3 days for four groups (see Table 5.6). All other sessions were completed in one sitting. To check whether this influenced the main results, we reran the analysis and excluded data from individuals whose test sessions were split over multiple days (see section 5.3.4).

5.2.4. Training procedure

Participants were asked to complete 20 days of training on any internet-enabled device except smart-phones. The training platform did not allow more than one training session to be started per day. Each training session lasted a maximum of 12 min or a set number of trials, whichever was reached first (see task descriptions). If a participant failed to respond for over 5 min, the training session timed out and was not included in the total number of training sessions. Task difficulty was adaptive according to performance and participants received feedback.

Training was designed to be motivating by providing positive feedback, such as flashing stars, after every correct response. Motivational phrases (e.g. 'awesome!', 'three in a row!') were shown as intermittent reinforcers (Ferster & Skinner, 1957). To incentivise training further, participants received virtual trophies. Before each training session, participants were asked to select a trophy chest (bronze, silver or gold). After the session, participants could open the chest to find a trophy in their online trophy cabinet. Participants were able to track the number of training sessions they had completed by viewing their trophy cabinet. Participants were reminded to train by automated daily e-mails and additional e-mail reminders sent by the research team. Additionally, teachers were asked to

remind adolescent participants to train. Volunteers also received monetary rewards at T2 if they had completed at least 15 training days. Adolescents received a £10 Amazon voucher; adults received £30 in cash. After the third test session adults received a further £10 in cash and adolescents received a certificate of participation. The training platform was designed to resemble school-based learning: testing was carried out in groups in the classroom and the training programme was comparable to homework in terms of duration and frequency.

5.2.5. Numerosity discrimination task

The numerosity discrimination task was used to measure the ability to rapidly approximate and compare the number of items within two different sets of coloured dots presented on a grey background. In this task, the total number of dots and dot proportions (i.e., the relative number of dots of each colour) in each array could be modified to vary difficulty level, such that a higher number of dots and a higher dot proportion represented a more difficult trial (Halberda et al., 2012).

Numerosity discrimination testing protocol

The dot proportions used were 0.30, 0.40, 0.42, 0.45, 0.47, and 0.49; the last four proportions, which were more difficult, appeared twice as often as the first two, easier proportions. Testing started with four easy trials (i.e., dot proportion = 0.30), but the proportion used in all subsequent trials was randomized. Only trials

with black and white dots were included in the testing. Individual dot positions for each array were selected pseudo-randomly: Their position was restricted such that none of the dots overlapped or touched and each dot was within the borders of the stimulus display.

Each trial started with a fixation cross presented for 250 ms, followed by a dot array presented for 200 ms. Participants were asked to select the colour of the more numerous dots. The two possible response options were displayed at the same time as the dot array and stayed on the screen until a response was given. The position (i.e., left or right) of the response buttons (i.e., 'black' or 'white') on the screen was counterbalanced between participants. There was no time limit on the response in each trial. After participants provided a response, the next trial started immediately. The numerosity discrimination task took 7 min to complete.

Numerosity discrimination training protocol

Each training session took 12 min or 64 trials to complete, whichever was reached first. All possible dot proportions were used. The first training session started with an initial dot proportion of 0.3. After each correct trial, difficulty increased one level (i.e., dot proportion came closer to 0.5); after each incorrect trial, it decreased two levels. The initial difficulty of each subsequent training session was two levels lower than the peak difficulty encountered in the previous training session. In training, randomly selected pairs of coloured dot sets were used (black and white, blue and yellow, blue and orange, violet and yellow, and violet and orange).

5.2.6. Relational reasoning task

A modified version of Raven's Progressive Matrices (Raven, 2009) was used to examine the ability to detect abstract relationships between groups of items. In this version of the relational reasoning task, puzzles consisted of a 3 × 3 matrix; eight of the cells contained shapes, but there was no shape in the bottom right cell. To select the correct response option, the participant had to deduce the pattern of change within the matrix. The items in a matrix could vary by colour, size, shape, and position across the matrix.

Relational reasoning testing protocol

Each trial started with a 500 ms fixation cross, followed by a 100 ms blank screen. In each trial, a puzzle was presented on the left side of the screen, and four possible response options were shown on the right side of the screen. Each puzzle was presented for 30 s. After 25 s, a clock appeared above the response options, indicating that 5 s remained until the next trial. The next trial started after participants responded or after 30 s had elapsed. The task took 8 min to complete. There were three test sessions; a different set of 80 puzzles using abstract shapes was created for each session. The order of the 80 puzzles within each set was the same for all participants, starting with five easy trials. The order of the three sets was counterbalanced across participants. If a participant completed all 80 puzzles within the 8 min time limit, the same set was presented again, but data from these additional puzzles were not included in the analysis.

Relational reasoning training protocol

Each training session took 12 min or 40 trials to complete, whichever was reached first. For each session, abstract and iconic puzzle shapes were selected. The first training session started with an easy puzzle. Training was adapted to performance such that the number of changing dimensions increased by one after each correct response and decreased by one after each incorrect response. The initial difficulty of each subsequent training session was two levels lower than that in the previous training session.

5.2.7. Face perception task

The design of the face perception task was as described in Chapter 3.2.5. The face perception task measured the ability to process featural and configural changes in faces (Cohen Kadosh, 2011). Participants were asked to decide whether two faces presented consecutively were the same or different. Faces were considered to be different when there were changes in any of the following face properties: gaze direction (left or right), expression (happy or sad), or identity (Person A or Person B). Participants were informed that faces should be classified as the same only if all three face properties were exactly the same.

Face perception testing protocol

Photos of 26 faces (16 white, 10 Asian; 16 female, 10 male), were taken under standardized lighting conditions for the purpose of this experiment. Four colour photos were obtained for each face: two with a happy expression (one with

leftward gaze and one with rightward gaze) and two with a sad expression (one with leftward gaze and one with rightward gaze). Photos were scaled to a uniform size and cropped to exclude external features of the face (e.g., hair) using the GNU Image Manipulation Program (GIMP Team, 2013).

Each trial started with a fixation cross presented for 800 ms, followed by the first face for 500 ms, and then another fixation cross for 800 ms, and then the second face for 500 ms. In the response display, the two possible response options ('same' or 'different') were shown simultaneously with the presentation of the two faces. The next trial started immediately after participants responded. One test took 7.5 min to complete.

Each test session contained a different set of stimuli, and each set comprised 48 different trials in which the faces of White women were shown. The order of the three sets of stimuli was counterbalanced across participants. If participants finished the 48 trials within the 7.5 min time limit, the trials were presented again, but the data were not included in the analysis. On the first 2 trials, the images had a noise mask of 25%, and difficulty in the remaining trials was increased by adding noise masks of increasing strength (from 25% to 81% in steps of 8 percentage points).

Face perception training protocol

Each training session lasted for 12 min or 48 trials, whichever was reached first. Twenty different sets of faces (five sets showed Asian women, five sets showed Asian men, five sets showed white women, and five sets showed white men) were generated for training. Training task difficulty was adapted to performance. In the

first training session, a 25% noise mask was applied to the first images. After a correct trial, noise strength was increased by 8 percentage points. After an incorrect trial, noise strength was decreased by 16 percentage points or kept at 25% - the lowest level. Each subsequent training session started with an initial difficulty level that was 16 percentage points lower than the peak difficulty encountered in the previous training session.

5.2.8. Backward digit span task

The backward digit span task was used to measure verbal working memory. Participants were asked to remember a sequence of digits in a certain order and to recall them in the reverse order. Minimum sequence length was two digits, sequences neither started nor ended with a 0, and no digit appeared twice or more in a row. Each trial started with a 500 ms fixation cross, followed by a 250 ms blank display. Digits were presented at a rate of one per second with an interstimulus interval of 250 ms. At the end of each sequence, participants were presented with a number of dashes equal to the length of the digit sequence they had just seen and were asked to input the digit sequence in reverse order, using the on-screen keyboard. Participants were not permitted to correct a response after a digit had been entered. There was no time limit on the response. After the response was given, the next trial started immediately. The task took 6 min to complete. The sequence length started at five digits, and trial difficulty was adapted to performance such that after correct trials, the difficulty level increased by one level (i.e., the sequence length increased by one), and after incorrect trials, the difficulty level decreased by 1 level (i.e., the sequence length decreased by 1).

5.2.9. Face memory task

The design of the face memory task was as described in Chapter 3.2.4. An adaptation of the Cambridge face memory task (CFMT) was used to assess the ability to learn and recognise unknown faces using a 3-Alternative-Forced Choice (3-AFC) test. Participants were asked to memorise 6 target faces and then locate one of the targets from a panel of three faces comprising the target plus two distractor faces that had not been memorised. A set of 198 face stimuli matching the specifications of the original CFMT was created for the purpose of the experiment. Black and white photographs of 66 Caucasian males from three angles (frontal/left quarter profile/right quarter profile) were obtained from the Facial Recognition Technology database (Phillips, Moon, Rizvi, & Rauss, 2000). Photos were cropped to exclude external features of the face (hair etc.) using the GNU Image Manipulation Program (GIMP Team, 2013). The task consisted of three blocks. In the first block, a target face was shown at three different angles, for 3 s each, and this was followed by three 3-AFC trials. This procedure was repeated for five more target faces. In the second block, frontal views of the same six target faces were presented simultaneously for 20 s, and this was followed by eighteen 3-AFC trials. In the third block, frontal views of the same six target faces were presented simultaneously for 20 s, but a 50% Gaussian noise mask was added to the faces in the eighteen 3-AFC trials that followed.

There was no time limit on the response in any of the blocks. After participants responded, the next trial started immediately. The task took 9 min or 54 trials to complete, whichever came first. Three sets of stimuli were created, one for each of the three test sessions. The order of presentation of these sets was

counterbalanced across participants. Each testing set contained 6 unique target faces and 6 unique distractor faces, as well as a set of 30 distractor faces that was used in all three test sessions. These common distractors were used to increase the difficulty of the task and prevent ceiling effects.

5.2.10. Statistical analysis

GLMMs implemented in the lme4 package (Bates et al., 2013) in R (R Core Team, 2015) were used to investigate the degree to which participants improved their task performance after training and whether the effect of training differed between age and training groups. Trials in any of the tasks with a response time under 250 ms were excluded from the analysis. For the numerosity discrimination, relational reasoning, face perception, and face memory tasks, the sums of correct and incorrect responses across trials were used as dependent variables. The models predicted each participant's task accuracy on the basis of four independent variables: training group, age group, test session, and number of completed training sessions (to control for differences in motivation). The models included fixed main effects of all four variables and fixed interaction effects between test session, training group, and age group as well as an interaction between training group and number of days trained. Orthogonal Helmert coding was used for all categorical fixed effects. Training days were standardized to z-scores. To account for individual differences, attrition, and the repeated measures for each participant, the model included a participant-specific random intercept (nested in school or university).

A LMM was used to investigate training effects on performance in the backward digit span task. This model incorporated each participant's maximal digit span as the dependent variable and the same random and fixed effects that were used in the GLMMs. The effects of the predictors on the dependent variables were investigated using an omnibus Type III Wald χ^2 test. Planned comparisons were performed to inspect differences across test sessions, age groups, and training groups using the multcomp package (Hothorn et al., 2016). For each of the five tasks, we inspected 26 comparisons of performance changes between T1 and T2 and between T1 and T3. To investigate general training effects, we analysed changes in performance in the trained tasks between test sessions within training groups (2 tests) and compared these effects between training groups (4 tests). Age-dependent training effects were investigated by looking at changes in performance in each age group on their trained task (8 tests). Between-age-group comparisons of age-dependent training effects were made by looking at changes in accuracy between age groups on their trained task (12 tests). All reported results were Bonferroni-corrected for these 26 comparisons. For additional analysis, which investigated potential confounds, see section 5.3.4.

5.3. Results

5.3.1. General training effects

Training on the numerosity discrimination, relational reasoning and face perception task improved performance on these respective tasks (Figure 5.2). Changes in performance differed between training groups, as indicated by

significant interactions between time point and training group for the numerosity discrimination task ($\chi^2(4) = 34.61, p < .001$), relational reasoning task ($\chi^2(4) = 328.48, p < .001$) and face perception task ($\chi^2(4) = 12.57, p = .014$).

Planned comparisons showed that participants who were trained in numerosity discrimination showed significantly improved performance in numerosity discrimination at T2 but those gains were not sustained to T3 (Table 5.2). Compared with participants who received training in one of the other two tasks, participants in the numerosity discrimination training group showed significantly higher gains in numerosity discrimination at T2. These effects were due mainly to the adult age group. When the adults' data were excluded, some of the effects of numerosity discrimination training became non-significant after Bonferroni correction (see section 5.3.4).

Participants who were trained in relational reasoning showed significantly improved performance in relational reasoning at T2 and T3 (Table 5.3). These gains were higher than those in participants trained in one of the other tasks (Figure 5.2).

Participants who were trained in face perception showed significantly improved performance in face perception at T2 but not T3 (Table 5.4). The gains at T2 were higher than those in participants trained in numerosity discrimination (Figure 5.2). However, these effects were not stable in supplementary analyses: The training effects in face perception lost significance when confounds like variation in group size were controlled for (see section 5.3.4).

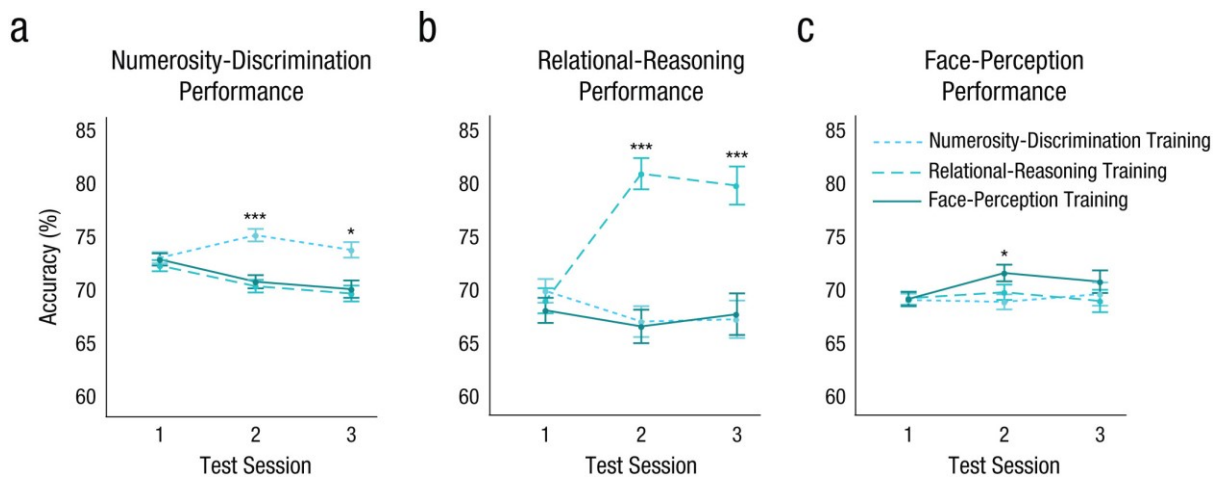


Figure 5.2. Performance by Training Group. Percentage accuracy with standard error bars in (a) the numerosity discrimination task, (b) the relational reasoning task, (c) the face perception task at the three test sessions of the. Asterisks in (a) and (b) indicate significantly greater gains at T2 and T3 in the group trained in the indicated task than in the other two groups (* $p < .05$, *** $p < .001$). The asterisk in (c) indicates a significant difference in gain at T2 between the group trained in face perception and the group trained on numerosity discrimination (* $p < .05$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

Table 5.2. Numerosity Discrimination Performance Change after Training

Contrast	Estimate	SE	z	p	
T1 vs. T2, ND	0.39	0.11	3.38	.019	*
T1 vs. T3, ND	0.18	0.13	1.39	1	
T1 vs. T2, ND vs. FP	0.75	0.16	4.65	< .001	***
T1 vs. T2, ND vs. RR	0.85	0.16	5.24	< .001	***
T1 vs. T3, ND vs. FP	0.58	0.18	3.15	.042	*
T1 vs. T3, ND vs. RR	0.63	0.18	3.50	.012	*
T1 vs. T2, ND, younger adol.	-0.07	0.06	-1.11	1	
T1 vs. T2, ND, mid-adol.	0.06	0.06	0.95	1	
T1 vs. T2, ND, older adol.	0.20	0.06	3.49	.013	*
T1 vs. T2, ND, adults	0.20	0.05	3.80	.004	**
T1 vs. T3, ND, younger adol.	-0.05	0.07	-0.74	1	
T1 vs. T3, ND, mid-adol.	-0.07	0.07	-1.08	1	
T1 vs. T3, ND, older adol.	0.06	0.06	0.99	1	
T1 vs. T3, ND, adults	0.24	0.07	3.52	.011	*
T1 vs. T2, ND, younger adol. vs. mid-adol.	-0.12	0.08	-1.45	1	
T1 vs. T2, ND, mid-adol. vs. older adol.	-0.14	0.08	-1.74	1	
T1 vs. T2, ND, older adol. vs. adults	-0.00	0.08	-0.06	1	
T1 vs. T2, ND, younger adol. vs. adults	-0.27	0.08	-3.34	.022	*
T1 vs. T2, ND, mid-adol. vs. adults	-0.15	0.08	-1.85	1	
T1 vs. T2, ND, younger adol. vs. older adol.	-0.26	0.08	-3.20	.036	*
T1 vs. T3, ND, younger adol. vs. mid-adol.	0.02	0.09	0.24	1	
T1 vs. T3, ND, mid-adol. vs. older adol.	-0.13	0.09	-1.47	1	
T1 vs. T3, ND, older adol. vs. adults	-0.18	0.09	-2.03	1	
T1 vs. T3, ND, younger adol. vs. adults	-0.29	0.10	-3.06	.059	
T1 vs. T3, ND, mid-adol. vs. adults	-0.31	0.09	-3.30	.025	*
T1 vs. T3, ND, younger adol. vs. older adol.	-0.11	0.09	-1.21	1	

Note. T1 = test session one; T2 = test session two; T3 = test session three; ND = numerosity discrimination training; RR = relational reasoning training; FP = face perception training; younger adol. = younger adolescents; mid-adol. = mid-adolescents; older adol. = older adolescents (* $p < .05$, ** $p < .01$, *** $p < .001$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

Table 5.3. Relational Reasoning Performance Change after Training

Contrast	Estimate	SE	z	p	
T1 vs. T2, RR	3.17	0.19	16.52	<.001	***
T1 vs. T3, RR	2.59	0.21	12.1	<.001	***
T1 vs. T2, RR vs. FP	3.76	0.25	14.89	<.001	***
T1 vs. T2, RR vs. ND	3.90	0.25	15.55	<.001	***
T1 vs. T3, RR vs. FP	3.11	0.29	10.8	<.001	***
T1 vs. T3, RR vs. ND	3.29	0.29	11.33	<.001	***
T1 vs. T2, RR, younger adol.	0.46	0.08	6.11	<.001	***
T1 vs. T2, RR, mid-adol.	0.36	0.07	5.17	<.001	***
T1 vs. T2, RR, older adol.	1.03	0.09	11.53	<.001	***
T1 vs. T2, RR, adults	1.32	0.13	9.76	<.001	***
T1 vs. T3, RR, younger adol.	0.46	0.08	5.57	<.001	***
T1 vs. T3, RR, mid-adol.	0.28	0.08	3.68	.006	**
T1 vs. T3, RR, older adol.	1.06	0.11	9.36	<.001	***
T1 vs. T3, RR, adults	0.79	0.14	5.54	<.001	***
T1 vs. T2, RR, younger adol. vs. mid-adol.	0.10	0.10	0.97	1	
T1 vs. T2, RR, mid-adol. vs. older adol.	-0.67	0.11	-5.87	<.001	***
T1 vs. T2, RR, older adol. vs. adults	-0.28	0.16	-1.75	1	
T1 vs. T2, RR, younger adol. vs. adults	-0.85	0.15	-5.51	<.001	***
T1 vs. T2, RR, mid-adol. vs. adults	-0.95	0.15	-6.27	<.001	***
T1 vs. T2, RR, younger adol. vs. older adol.	-0.57	0.12	-4.84	<.001	***
T1 vs. T3, RR, younger adol. vs. mid-adol.	0.18	0.11	1.63	1	
T1 vs. T3, RR, mid-adol. vs. older adol.	-0.78	0.14	-5.74	<.001	***
T1 vs. T3, RR, older adol. vs. adults	0.27	0.18	1.49	1	
T1 vs. T3, RR, younger adol. vs. adults	-0.33	0.16	-2.00	1	
T1 vs. T3, RR, mid-adol. vs. adults	-0.51	0.16	-3.17	.040	*
T1 vs. T3, RR, younger adol. vs. older adol.	-0.60	0.14	-4.28	<.001	***

Note. T1 = test session one; T2 = test session two; T3 = test session three; ND = numerosity discrimination training; RR = relational reasoning training; FP = face perception training; younger adol. = younger adolescents; mid-adol. = mid-adolescents; older adol. = older adolescents (* $p < .05$, ** $p < .01$, *** $p < .001$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

Table 5.4. Face Perception Performance Change after Training

Contrast	Estimate	SE	z	p	p
T1 vs. T2, FP	0.54	0.14	3.92	<.001	.002
T1 vs. T3, FP	0.46	0.16	2.79	.005	.137
T1 vs. T2, FP vs. RR	0.46	0.19	2.43	.015	.393
T1 vs. T2, FP vs. ND	0.59	0.19	3.16	.002	.042
T1 vs. T3, FP vs. RR	0.48	0.22	2.12	.034	.882
T1 vs. T3, FP vs. ND	0.34	0.23	1.50	.133	1
T1 vs. T2, FP, younger adol.	0.04	0.06	0.65	.515	1
T1 vs. T2, FP, mid-adol.	0.12	0.06	2.08	.038	.985
T1 vs. T2, FP, older adol.	0.13	0.07	1.91	.056	1
T1 vs. T2, FP, adults	0.25	0.09	2.89	.004	.100
T1 vs. T3, FP, younger adol.	0.09	0.06	1.39	.164	1
T1 vs. T3, FP, mid-adol.	0.13	0.06	1.99	.046	1
T1 vs. T3, FP, older adol.	0.08	0.09	0.85	.393	1
T1 vs. T3, FP, adults	0.17	0.10	1.59	.111	1
T1 vs. T2, FP, younger adol. vs. mid-adol.	-0.08	0.08	-1.01	.314	1
T1 vs. T2, FP, mid-adol. vs. older adol.	-0.01	0.09	-0.13	.899	1
T1 vs. T2, FP, older adol. vs. adults	-0.12	0.11	-1.06	.290	1
T1 vs. T2, FP, younger adol. vs. adults	-0.21	0.10	-2.03	.042	1
T1 vs. T2, FP, mid-adol. vs. adults	-0.13	0.10	-1.24	.217	1
T1 vs. T2, FP, younger adol. vs. older adol.	-0.09	0.09	-1.04	.297	1
T1 vs. T3, FP, younger adol. vs. mid-adol.	-0.04	0.09	-0.43	.666	1
T1 vs. T3, FP, mid-adol. vs. older adol.	0.05	0.11	0.48	.632	1
T1 vs. T3, FP, older adol. vs. adults	-0.09	0.14	-0.66	.511	1
T1 vs. T3, FP, younger adol. vs. adults	-0.08	0.12	-0.63	.531	1
T1 vs. T3, FP, mid-adol. vs. adults	-0.04	0.12	-0.31	.760	1
T1 vs. T3, FP, younger adol. vs. older adol.	0.01	0.11	0.12	.903	1

Note. T1 = test session one; T2 = test session two; T3 = test session three; ND = numerosity discrimination training; RR = relational reasoning training; FP = face perception training; younger adol. = younger adolescents; mid-adol. = mid-adolescents; older adol. = older adolescents (* $p < .05$, ** $p < .01$, *** $p < .001$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

5.3.2. Age-dependent training effects

General training effects were significantly moderated by age group for the numerosity discrimination task ($\chi^2(12) = 24.64, p = .017$), relational reasoning task ($\chi^2(12) = 80.13, p < .001$), but not for face perception task ($\chi^2(12) = 8.80, p = .720$) (Figure 5.3).

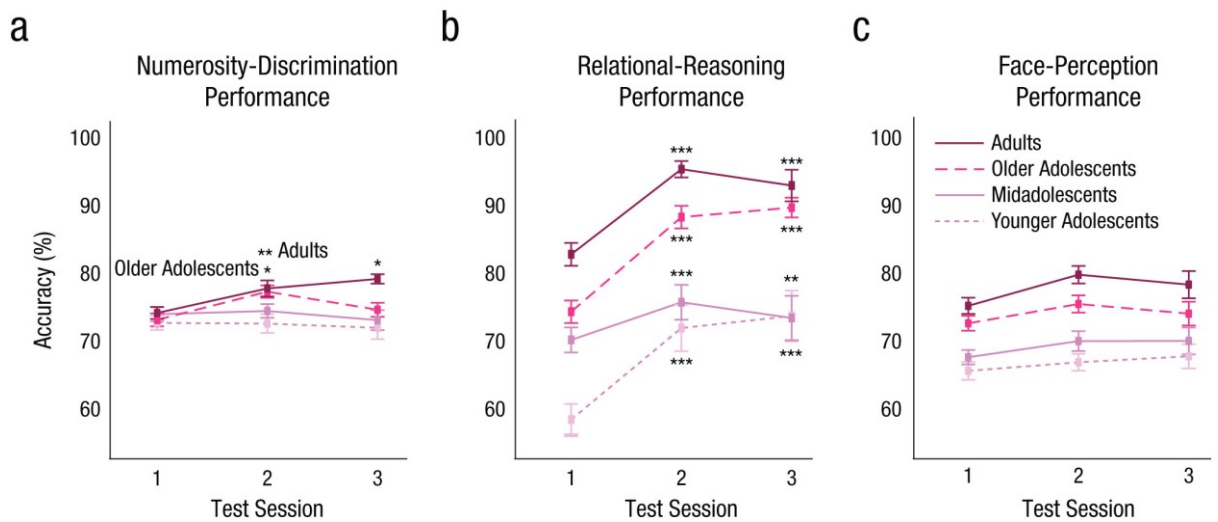


Figure 5.3. Performance by Age Group. Percentage accuracy on the (a) numerosity discrimination task, (b) relational reasoning task, and (c) face perception task is plotted as a function of test session. Asterisks indicate significant training gains at T2 or T3 within age groups (* $p < .05$, ** $p < .005$, *** $p < .001$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

The only age groups to improve their performance in numerosity discrimination at T2 were older adolescents and adults, who were trained in numerosity discrimination (Figure 5.3). These improvements were larger than the changes in performance in younger adolescents (Figure 5.4). Only adults showed a consolidation effect in numerosity discrimination at T3, and this effect was larger than that for mid-adolescents (Figure 5.3; Table 5.2). However, the training effects of numerosity discrimination did not remain statistically significant when we included covariates for differences in spacing of testing and group size at testing. This was particularly the case for the consolidation effects at T3 (see section 5.3.4).

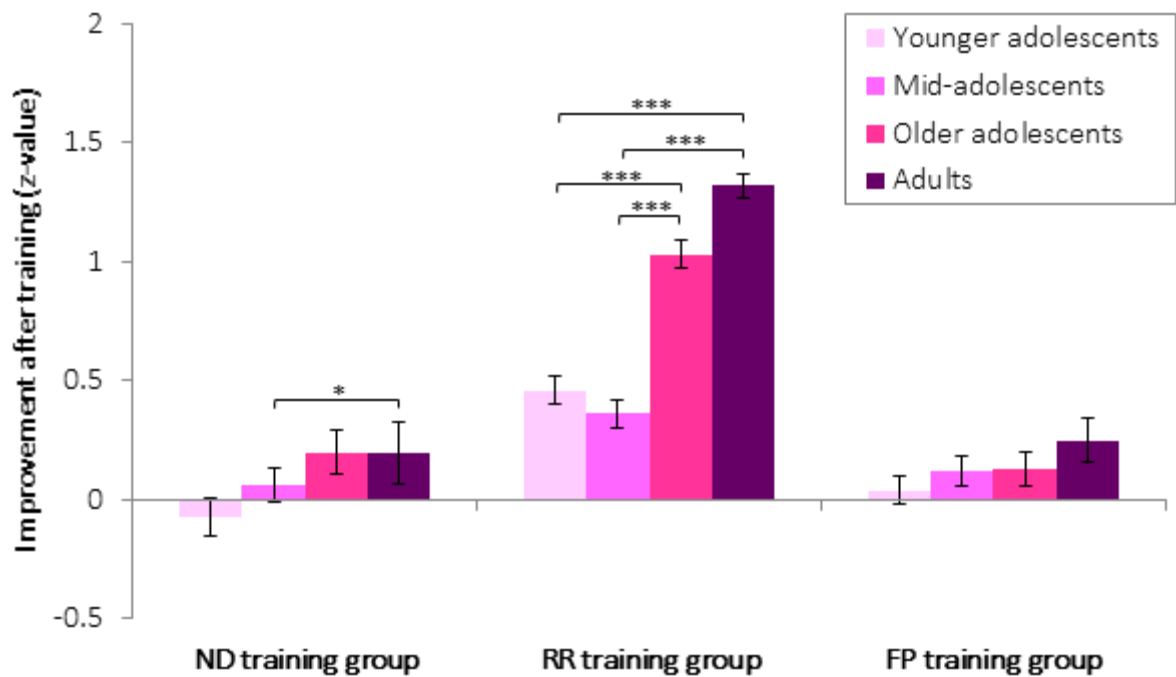


Figure 5.4. Training Gains Compared between Age Groups. Improvement after training (performance at T2 – performance at T1) for the numerosity discrimination (ND) task, relational reasoning (RR) task, and face perception (FP) task is plotted as a function of age group. Asterisks indicate significant differences in training gains at T2 between age groups (* $p < .05$, ** $p < .005$, *** $p < .001$).

All age groups trained in relational reasoning showed improved relational reasoning performance at T2 (Figure 5.3). Improvements were stronger in older adolescents and adults than in younger adolescents and mid-adolescents (Figure 5.4). Improvements were sustained at T3 in all age groups (Table 5.3), but were stronger in older adolescents and adults than the younger age groups (Table 5.3).

None of the contrasts for face perception training was significant (Figure 5.3; Figure 5.4; Table 5.4).

5.3.3. Transfer effects

There was no evidence of transfer from relational reasoning training to backward digit span or from face perception training to face memory.

The two-way interaction between time point and training group was not significant for the backward digit span task ($\chi^2(4) = 2.54, p = .637$), and no significant improvements at T2 or T3 were found in the relational reasoning group (Table 5.5). There was no effect of age group on transfer to digit span ($\chi^2(12) = 14.87, p = .249$) and none of the age groups trained in relational reasoning significantly increased their digit span (Table 5.5).

For the face memory task, the overall two-way interaction between time point and training group was significant ($\chi^2(4) = 12.31, p = .015$). However, no significant improvements in the face perception training group to T2 or T3 were found (Table 5.6). There was no effect of age group on transfer to face memory ($\chi^2(12) = 13.31, p = .347$) and none of the age groups trained in face perception improved significantly in face memory (Table 5.6).

Table 5.5. Change in Backward Digit Span Performance after Training

	Estimate	SE	z	p
T1 vs T2, RR	0.52	0.32	1.61	1
T1 vs T3, RR	0.84	0.38	2.23	.671
T1 vs T2, RR vs. FP	-0.56	0.46	-1.21	1
T1 vs T2, RR vs. ND	-0.14	0.45	-0.30	1
T1 vs T3, RR vs. FP	0.19	0.55	0.35	1
T1 vs T3, RR vs. ND	0.04	0.54	0.07	1
T1 vs T2, RR, younger adol.	0.10	0.16	0.63	1
T1 vs T2, RR, mid-adol.	0.13	0.14	0.92	1
T1 vs T2, RR, older adol.	0.45	0.16	2.91	.094
T1 vs T2, RR, adults	-0.16	0.19	-0.85	1
T1 vs T3, RR, younger adol.	0.18	0.18	0.99	1
T1 vs T3, RR, mid-adol.	-0.08	0.16	-0.46	1
T1 vs T3, RR, older adol.	0.48	0.18	2.65	.207
T1 vs T3, RR, adults	0.25	0.22	1.14	1
T1 vs T2, RR, younger adol. vs mid-adol.	-0.03	0.21	-0.14	1
T1 vs T2, RR, mid-adol. vs older adol.	-0.32	0.21	-1.54	1
T1 vs T2, RR, older adol. vs adults	0.61	0.25	2.51	.318
T1 vs T2, RR, younger adol. vs adults	0.26	0.25	1.06	1
T1 vs T2, RR, mid-adol. vs adults	0.29	0.24	1.23	1
T1 vs T2, RR, younger adol. vs older adol.	-0.35	0.22	-1.59	1
T1 vs T3, RR, younger adol. vs mid-adol.	0.26	0.25	1.04	1
T1 vs T3, RR, mid-adol. vs older adol.	-0.56	0.25	-2.29	.578
T1 vs T3, RR, older adol. vs adults	0.23	0.29	0.81	1
T1 vs T3, RR, younger adol. vs adults	-0.07	0.29	-0.25	1
T1 vs T3, RR, mid-adol. vs adults	-0.33	0.28	-1.19	1
T1 vs T3, RR, younger adol. vs older adol.	-0.30	0.26	-1.18	1

Note. ND = numerosity discrimination training; RR = relational reasoning training; FP = face perception training; younger adol. = younger adolescents; mid-adol. = mid-adolescents; older adol. = older adolescents. Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

Table 5.6. Change in Cambridge Face Memory Task Performance after Training

	Estimate	SE	z	p	
T1 vs T2, FP	-0.59	0.14	-4.20	<.001	***
T1 vs T3, FP	-0.78	0.17	-4.53	<.001	***
T1 vs T2, FP vs. RR	0.29	0.20	1.48	1	
T1 vs T2, FP vs. ND	0.50	0.19	2.57	.262	
T1 vs T3, FP vs. RR	-0.31	0.24	-1.31	1	
T1 vs T3, FP vs. ND	-0.17	0.24	-0.71	1	
T1 vs T2, FP, younger adol.	-0.29	0.06	-4.84	<.001	***
T1 vs T2, FP, mid-adol.	-0.21	0.06	-3.56	.010	*
T1 vs T2, FP, older adol.	0.01	0.07	0.08	1	
T1 vs T2, FP, adults	-0.11	0.09	-1.22	1	
T1 vs T3, FP, younger adol.	-0.16	0.07	-2.25	.640	
T1 vs T3, FP, mid-adol.	-0.29	0.06	-4.52	<.001	***
T1 vs T3, FP, older adol.	-0.14	0.09	-1.52	1	
T1 vs T3, FP, adults	-0.19	0.11	-1.74	1	
T1 vs T2, FP, younger adol. vs mid-adol.	-0.08	0.08	-0.94	1	
T1 vs T2, FP, mid-adol. vs older adol.	-0.21	0.09	-2.27	.607	
T1 vs T2, FP, older adol. vs adults	0.11	0.11	0.99	1	
T1 vs T2, FP, younger adol. vs adults	-0.18	0.11	-1.70	1	
T1 vs T2, FP, mid-adol. vs adults	-0.10	0.11	-0.96	1	
T1 vs T2, FP, younger adol. vs older adol.	-0.29	0.09	-3.09	.053	
T1 vs T3, FP, younger adol. vs mid-adol.	0.13	0.10	1.33	1	
T1 vs T3, FP, mid-adol. vs older adol.	-0.15	0.11	-1.34	1	
T1 vs T3, FP, older adol. vs adults	0.05	0.14	0.34	1	
T1 vs T3, FP, younger adol. vs adults	0.03	0.13	0.20	1	
T1 vs T3, FP, mid-adol. vs adults	-0.10	0.13	-0.81	1	
T1 vs T3, FP, younger adol. vs older adol.	-0.02	0.12	-0.19	1	

Note. ND = numerosity discrimination training; RR = relational reasoning training; FP = face perception training; younger adol. = younger adolescents; mid-adol. = mid-adolescents; older adol. = older adolescents (* $p < .05$, *** $p < .001$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

5.3.4. Potential confounds

We tested whether participants varied by age and training group in a number of potential confounds (Table 5.7). There were no significant differences between training groups but age groups varied in the number of days they trained, the spacing between training and test sessions, number of participants missing at T2 and T3 and their group size at testing. Missing data was addressed by our main

analyses (see section 5.2.10). The other confounds were addressed in three supplementary analyses.

Table 5.7. Confounds with Age

	Younger adolescents 11.27-13.38	Mid- adolescents 13.39-15.89	Older adolescents 15.90-18.00	Adults 18.01- 33.15	Test of difference
Days trained	14.90 (0.48)	14.74 (0.44)	14.62 (0.45)	18.83 (0.54)	$F(3,535)=15.53$, $p < .001$ ¹
Days between training	2.05 (0.15)	2.06 (0.14)	2.53 (0.15)	1.48 (0.17)	$F(3,535)=7.25$, $p < .001$ ²
Days between T1 and T2	36.72 (0.69)	32.90 (0.68)	39.69 (0.74)	31.93 (0.89)	$F(3,596)=22.36$, $p < .001$ ³
Days between T2 and T3	134.04 (2.71)	146.09 (2.60)	124.32 (3.01)	150.4 (3.88)	$F(3,390)=14.64$, $p < .001$ ⁴
Group size at testing	14.98 (1.98)	14.78 (1.95)	23.340 (2.00)	7.10 (8.00)	$\chi^2(3) = 12.74$, $p = .005$ ⁵
Split test sessions	34	4	0	0	$\chi^2(3) = 72.02$, $p < .001$ ⁶
Participants missing at T2	15	9	34	3	$\chi^2(3) = 132.72$, $p < .001$ ⁷
Participants missing at T3	68	56	86	48	$\chi^2(3) = 27.93$, $p < .001$ ⁸

Note. Mean with *SD* in brackets or total *N* are shown. Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

¹ Adults trained more than all other age groups ($p < .001$).

² Adults' training sessions were spaced more closely than older adolescents' ($p < .001$).

³ Only adults' and mid-adolescents' test sessions T1 and T2 were similarly spaced, all other comparisons differed at $p < .05$.

⁴ T2 and T3 for adults and mid-adolescents, and for older and younger adolescents were similarly spaced; all other comparisons differed at $p < 0.05$.

⁵ Group size at testing differed between older and younger adolescents, and between older and mid-adolescents at $p < .001$.

⁶ There were more younger adolescents with split test sessions than all other age groups ($p < .001$); all other comparisons were non-significant.

⁷ There were more older adolescents missing at T2 than all other age groups ($p < .05$); all other comparisons were non-significant.

⁸ There were more older adolescents missing at T3 than mid-adolescents ($p = .012$); all other comparisons were non-significant.

Differences between adults and the other age groups in the amount of training completed and the spacing between test sessions

Adults completed more training than all other age groups and completed training more quickly than older adolescents (Table 5.7). To check whether this affected our main findings, we re-ran the models for the three training tasks (numerosity discrimination, relational reasoning, face perception) and excluded adults' data.

The results were qualitatively similar, but there were some quantitative changes.

One interaction and five planned contrasts became non-significant (Table 5.8):

- *Numerosity discrimination*: Some of the training effects for numerosity discrimination training became non-significant, particularly at T3.
- *Relational reasoning*: No changes.
- *Face perception*: The interaction between time point and training group became non-significant for the face perception task in this analysis ($\chi^2(4) = 8.52, p = .074$, previously $p = .014$). The training effects in face perception lost significance.

All other reported effects remained the same.

Table 5.8. Contrasts that Became Non-Significant after Excluding Adults' Data

Original analysis					
Contrast	Estimate	SE	z-ratio	p	
T1 vs T2, ND	0.39	0.11	3.38	.019	*
T1 vs T3, ND vs. FP	0.58	0.18	3.15	.042	*
T1 vs T3, ND vs. RR	0.63	0.18	3.50	.012	*
T1 vs T2, FP	0.54	0.14	3.92	.002	**
T1 vs T2, FP vs. ND	0.59	0.19	3.16	.042	*
Supplementary analysis without adults' data					
Contrast	Estimate	SE	z-ratio	p	
T1 vs T2, ND	0.19	0.10	1.83	1	
T1 vs T3, ND vs. FP	0.30	0.16	1.90	1	
T1 vs T3, ND vs. RR	0.30	1.16	1.94	.950	
T1 vs T2, FP	0.29	0.11	2.71	.123	
T1 vs T2, FP vs. ND	0.36	0.15	2.46	.253	

Note. ND = numerosity discrimination training; RR = relational reasoning training; FP = face perception training (* $p < .05$, ** $p < .01$, *** $p < .001$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

Age-group differences in the spacing between test sessions and group sizes at testing

The spacing between test sessions as well as group sizes at testing differed because of schools' timetabling constraints (Table 5.7).

To check whether this influenced our main results, we re-ran the models for the three training tasks (numerosity discrimination, relational reasoning, face perception) and included covariates for spacing between test sessions and group sizes.

Our results were qualitatively similar: the overall interactions were still significant and the effects retained their directionality. However, there were some quantitative changes in that nine planned contrasts that were significant in our main analysis, became non-significant after Bonferroni-correction (three without

Bonferroni-correction; Table 5.9). This may partly be due to the loss of power incurred by including additional covariates.

- *Numerosity discrimination*: Some of the training effects for numerosity discrimination training became non-significant, particularly at T3.
- *Relational reasoning*: Younger and mid-adolescents did not show a training effect.
- *Face perception*: The overall training effect disappeared.

All other reported effects remained the same.

Table 5.9. Contrasts that Became Non-Significant after Including Covariates for Group Size and Spacing between Test Sessions

Original analysis					
<i>Contrast</i>	<i>Estimate</i>	<i>SE</i>	<i>z-ratio</i>	<i>p</i>	
T1 vs T2, ND	0.39	0.11	3.38	.019	*
T1 vs T3, ND vs. FP	0.58	0.18	3.15	.042	*
T1 vs T3, ND vs. RR	0.63	0.18	3.50	.012	*
T1 vs T2, ND, older adol.	0.20	0.06	3.49	.013	*
T1 vs T2, ND, adults	0.20	0.05	3.80	.004	**
T1 vs T3, ND, adults	0.24	0.07	3.52	.011	*
T1 vs T3, RR, younger adol.	0.46	0.08	5.57	<.001	***
T1 vs T3, RR, mid-adol.	0.28	0.08	3.68	.006	**
T1 vs T2, FP	0.54	0.14	3.92	.002	**
Supplementary analysis with covariates					
<i>Contrast</i>	<i>Estimate</i>	<i>SE</i>	<i>z-ratio</i>	<i>p</i>	
T1 vs T2, ND	0.26	0.14	1.85	1	
T1 vs T3, ND vs. FP	0.54	0.18	2.92	.091	
T1 vs T3, ND vs. RR	0.62	0.18	3.42	.416	
T1 vs T2, ND, older adol.	0.16	0.06	2.6	.244	
T1 vs T2, ND, adults	0.17	0.06	3.08	.053	
T1 vs T3, ND, adults	0.08	0.12	0.63	1	
T1 vs T3, RR, younger adol.	0.40	0.17	2.43	.398	
T1 vs T3, RR, mid-adol.	0.22	0.17	1.34	1	
T1 vs T2, FP	0.35	0.17	2.07	.996	

Note. ND = numerosity discrimination training; RR = relational reasoning training; FP = face perception training; younger adol. = younger adolescents; mid-adol. = mid-adolescents; older adol. = older adolescents (* $p < .05$, ** $p < .01$, *** $p < .001$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

Testing group sessions were sometimes split

Testing was sometimes split over 2 or 3 sessions due to participants' time-constraints (Table 5.7).

To check whether this influenced the main results, we re-ran the models for the three training tasks (numerosity discrimination, relational reasoning, face perception) and excluded data from individuals whose test sessions were split over several days.

There were minor quantitative changes. Two planned contrasts became non-significant after Bonferroni-correction only (Table 5.10):

- *Numerosity discrimination*: Older adolescents' training effect was not significantly stronger than younger adolescents' training effects.
- *Relational reasoning*: Mid-adolescents' sustained training effect at T3 was no longer significant.
- *Face perception*: No changes.

All other reported effects remained the same.

Table 5.10. Contrasts that Became Non-Significant after Excluding Data from Individuals whose Test Sessions Were Split over Several Days

Original analysis					
Contrast	Estimate	SE	z-ratio	p	
T1 vs T2, ND, younger adol. vs older adol.	-0.26	0.08	-3.2	.036	*
T1 vs T3, RR, mid-adol.	0.28	0.08	3.68	.006	**
Supplementary analysis without split test sessions					
Contrast	Estimate	SE	z-ratio	p	
T1 vs T2, ND, younger adol. vs older adol.	-0.26	0.09	-2.99	.073	
T1 vs T3, RR, mid-adol.	0.2	0.08	2.66	.201	

Note. ND = numerosity discrimination training; RR = relational reasoning training; younger adol. = younger adolescents; mid-adol. = mid-adolescents; older adol. = older adolescents (* $p < .05$, ** $p < .01$). Adapted from Knoll et al. (2016) with permission from SAGE Publishing.

5.4. Discussion

This training study aimed to investigate cognitive skills relevant to maths education and when during adolescence such skills may be best trained. Numerosity discrimination training yielded small improvements only in late adolescence and adulthood. Relational reasoning training was already effective in early adolescence but showed a linear increase in benefit from mid - to late adolescence, and then no further improvement into adulthood. Training on face perception did not result in different levels of improvement in the different age groups. The results suggest that the ability to learn numerosity discrimination and especially relational reasoning is greater in late than in early adolescence.

Overall, participants who were trained in numerosity discrimination improved their numerosity discrimination skills more than participants trained in the other tasks. However, these effects were age-dependent: Only older adolescents' and

adults' performance improved significantly after training. Previous studies have shown that numerosity discrimination, which is related to mathematics performance, peaks at around the age of 30 (Halberda et al., 2012), and that approximate number processing can be trained in adulthood (Cappelletti, Pikkat, Upstill, Speekenbrink, & Walsh, 2015; DeWind & Brannon, 2012; Park & Brannon, 2013). However, ours is the first study to compare the training effect between age groups.

Relational reasoning performance was more improved by relational reasoning training than by training in other tasks. This training effect was observed in all age groups: Relational reasoning training improved relational reasoning task performance throughout adolescence and adulthood. This finding supports results from previous research in children and adults (Mackey et al., 2011; Mackey et al., 2013; Mackey et al., 2012). The training effects survived a 6-month no-training period. Between age group comparisons showed that the benefit from relational reasoning training increased from mid- to late adolescence, after which no further benefit was found in adulthood. This age effect was similar to the pattern of results observed in the numerosity discrimination task. This finding provides further evidence that training during older adolescence results in greater improvements in performance than does training during early adolescence.

The fact that relational reasoning can be trained in all the age groups tested here, and that it is particularly amenable to training during late adolescence, does not support the notion that matrix reasoning gives an indication of some kind of innate, fixed ability. This has implications for education because matrix reasoning is commonly used in IQ tests and school entrance exams.

Participants who were trained in face perception showed improvements in identifying changes in faces and facial features compared with participants trained only in numerosity discrimination. There were no age-dependent training effects. Previous studies on face perception training in adults have also yielded inconsistent results. For example, face cognition speed training was found to be effective in adults, whereas face memory training was not (Dolzycka et al., 2014).

There was no evidence of far-transfer from relational reasoning training to working memory performance or from face perception training to face memory performance. A small number of studies have demonstrated transfer effects from a trained task to a non-trained task, particularly if they are closely related (Klingberg, 2010; Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009). Many others have not (Melby-Lervåg & Hulme, 2013; Owen et al., 2010; Schwaighofer et al., 2015). Future studies should investigate near- and far-transfer to a broader range of tasks and over a wide age range to evaluate the significance of age-dependent transfer effects for education.

There are several possible explanations for the increased effects of training in late adolescence and adulthood observed here for numerosity discrimination and relational reasoning. First, improvements in training with age might be related to neurocognitive development. The prefrontal cortex is particularly late developing (Tamnes et al., 2017) and may retain high levels of plasticity (Fuhrmann et al., 2015). Tasks such as relational reasoning, which rely heavily on this region (Dumontheil, Houlton, et al., 2010), may therefore be better trained later in development. Performance on executive function tasks undergoes gradual improvement throughout adolescence (Crone & Dahl, 2012; Zelazo & Carlson,

2012), and this might also contribute to improved learning with age. Until recently, most studies investigating plasticity have concentrated on early childhood and have suggested that the adaptive processes of the nervous system are heightened in early development (Kuhl, 2004; Lewis & Maurer, 2005). In contrast, studies focusing on sensitive periods in later development are rare. Our findings indicate that the acquisition of relational reasoning and numerosity discrimination is more efficient in late adolescence than earlier in the teenage years, suggesting that plasticity for certain cognitive skills is sustained or even heightened at this relatively late stage of development. However, our study did not include participants younger than 11 years old, and we therefore cannot exclude the possibility that training would be efficient in younger participants. Future studies will need to elucidate the neurocognitive mechanisms of cognitive training and include younger as well as older age groups to show the trajectory of plasticity before and after adolescence.

Second, improved learning in late adolescence might be due to better strategy use. Older adolescents and adults have greater general cognitive abilities than young and mid-adolescents (Gur et al., 2012), which might enable them to develop and deploy strategies that result in greater training improvements. Of the three trained tasks, relational reasoning might be most amenable to improvement through enhanced cognitive strategies (Goodwin & Johnson-Laird, 2005).

Third, the age-dependent training effects may be due to a number of confounding variables. The testing and training conditions and behaviour were similar for the three adolescent groups, but the adult group was unavoidably different from the adolescent groups, in that the adults were self-selected and were paid more for

taking part than were the adolescents. The adolescent groups were self-selected to a lesser degree in that entire school classes took part. Given these differences, as might be expected, adults trained more and completed their training more quickly than adolescents. There were also differences between age groups in the spacing between test sessions and group sizes at testing. We controlled for these possible confounds by including the number of training days as a covariate in our main statistical analyses. In addition, supplementary analysis showed that excluding the adult data, or including covariates for confounds named above, did not result in major differences in results: The interactions remained significant and effects were still in the same direction. However, although the training and age effects of relational reasoning were remarkably robust in all analyses, the effects of numerosity discrimination were weaker and should therefore be interpreted with caution.

In summary, we found that complex cognitive skills relevant to maths education, particularly relational reasoning, show larger training effects in late adolescence than earlier in adolescence. These findings highlight the importance of late adolescence for education and, in contrast to the common assumption that 'earlier is better' for learning, highlight the need to investigate late adolescence as a potential window of opportunity for educational interventions.

While the study presented in this, and the prior two, chapters discusses cognitive and socio-cognitive performance across broad age ranges, the final experimental chapter of this thesis investigates individual differences in a narrow age group to better understand adolescent self-control, its relationship to educational predictors, and its neural correlates.

6. Chapter 6: The Neurocognitive Correlates of Academic Diligence

The dual-systems hypothesis of adolescent development predicts reduced self-control and heightened reward sensitivity during adolescence. Here we tested whether the interplay between frontal control and striatal reward systems is related to academic diligence, a predictor of educational attainment. We combined behavioural, structural MRI, functional MRI and connectivity data to assess the neurocognitive correlates of diligence. We recruited adolescent girls (N = 40, 14 - 15 years) and obtained behavioural measures of diligence using the Academic Diligence Task, which models students' choices when doing school-work. We also collected structural imaging data for each participant, as well as functional imaging data during an emotional go-no-go self-control task. As predicted by the dual-systems hypothesis, we found that inferior frontal activation correlated with diligence. However, frontal and striatal connectivity and structure showed no clear relation with diligence. Instead, we found prominent activation of temporal areas during the go-no-go task. This highlights the need to investigate more extended brain networks in future studies.

6.1. Introduction

Adolescence is thought to be a time of protracted development of self-control and increased reward seeking (Casey et al., 2008; Steinberg, 2008; Steinberg et al., 2017). This 'imbalance' between control and reward-sensitivity during adolescence is proposed to originate in the relatively early maturation of the subcortical reward system, including the striatum, while frontal control systems

still undergo protracted development during adolescence (Casey et al., 2008; Hall, 1904; Steinberg, 2008) (see pp. 17 - 22).

This dual-systems hypothesis of frontal self-control and striatal reward systems has been one of the most influential models of adolescent development (Shulman et al., 2016; Steinberg et al., 2017). Somerville and colleagues (2011), for instance, used an emotional go-no-go self-control task to show that inferior frontal activation and connectivity correlates with response inhibition in children, adolescents and adults. Adolescents, compared to children and adults, also showed an increased activity in the ventral striatum, which was linked to nonlinear reductions in impulse control to rewarding cues (happy faces), while response inhibition to neutral cues (neutral faces) improved linearly with age. This was taken as evidence that adolescents find it harder than other age groups to resist responding to rewarding social cues (Somerville, Hare, & Casey, 2011).

While many studies find similarly increased average impulsivity and reduced self-control in adolescence, compared to other age groups (Braams et al., 2015; Casey et al., 2008; Steinberg, 2008), several studies and reviews have highlighted pervasive individual differences in the maturation of self-control and fronto-striatal systems during adolescence (Crone & Dahl, 2012; Mills et al., 2014). Such individual differences may affect students' academic diligence (Duckworth & Steinberg, 2015). Academic diligence is the ability to regulate behaviour in the service of goals and been shown to be related to educational attainment (Galla et al., 2014). It has been proposed that diligence is the product of conflicting psychological processes – the exercise of will and the drive to seek immediate gratification (Duckworth & Steinberg, 2015). This rationale is similar to that of the

dual-systems hypothesis. Therefore, diligence might be hypothesized to correlate with fronto-striatal structure and function.

Here, we investigated this proposal. We took an individual differences approach and correlated individual levels of diligence with fronto-striatal structure and function. We recruited 40 girls aged 14 - 15 years, as previous studies have highlighted that mid-adolescents may find self-control tasks particularly challenging (Braams et al., 2015). We chose to recruit a relatively narrow age range so as to not confound individual differences in self-control with the ongoing development of executive functions during adolescence (Baum et al., 2017; Crone & Steinbeis, 2017).

We obtained behavioural measures of diligence using the Academic Diligence Task, which is designed to model students' behaviour when doing school-work (Galla et al., 2014). In this task, participants can freely allocate their time between doing useful but boring maths exercises and playing fun video games. The task has been shown to have incremental predictive validity for educational outcomes such as Grade Point Averages and performance on standardized maths and reading tests (Galla et al., 2014). Questionnaire measures of grit and self-control also reliably predicted unique variance in task behaviour, whereas agreeableness, a personality trait encompassing compliance, did not, thus demonstrating discriminant validity of this task (Galla et al., 2014).

We investigated how behaviour on the Academic Diligence Task was related to structure, function and connectivity of the inferior frontal gyrus and the striatum. We collected functional imaging data during an emotional go-no-go task with

happy and neutral peer faces as cues (Somerville et al., 2011). We chose a go-no-go task as an established measure of self-control, which is consistently associated with activation in well-defined frontal (inferior frontal gyrus) and striatal (ventral and dorsal) regions of interest (Ahmed, Bittencourt-Hewitt, & Sebastian, 2015; Simmonds, Pekar, & Mostofsky, 2008; Somerville et al., 2011). Using this task allowed us to interpret our findings in relation to previous studies with adults and adolescents and to test whether neural activation in the go-no-go task is predictive of behaviour on more naturalistic self-control tasks like the Academic Diligence Task. We chose the emotional variant of the go-no-go task as adolescents have been shown to be particularly responsive to emotional face stimuli (Somerville et al., 2011) and in affective contexts in general (Kilford et al., 2016; Prencipe et al., 2011)

Based on the dual-systems hypothesis and previous go-no-go studies (Casey et al., 2008; Somerville et al., 2011; Steinberg, 2008), we predicted that increased functional activation of the inferior frontal gyrus and decreased activation of the ventral striatum in the go-no-go task would correlate positively with diligence. We further predicted that increased diligence would be associated with increased connectivity strength between the inferior frontal gyrus and dorsal striatum, as well as decreased grey matter volume in the inferior frontal gyrus and the striatum.

6.2. Methods

6.2.1. Participants

42 typically developing girls aged 14 - 15 years were recruited for the purpose of this study. We chose to recruit only girls because of differences in pubertal development between the sexes during adolescence (Sisk & Foster, 2004). Participants attended eight different schools in Greater London and Cambridgeshire, UK, and were recruited through advertisements in schools and on social media. 28 participants attended state schools and 14 participants attended private schools. 38 participants were tested over the summer holidays, the remaining four were tested after school. Two participants were excluded from all analyses because of excess motion in the scanner (see section 6.2.4) leaving a total of 40 participants in the sample (Table 6.1).

Table 6.1. Participant Characteristics

Age	<i>range</i>	14.10 - 15.90
	<i>M</i>	14.99
	<i>SE</i>	0.09
IQ	<i>range</i>	85.03 - 122.51
	<i>M</i>	105.68
	<i>SE</i>	1.30
SES	<i>range</i>	2 (GCSE) - 6 (postgraduate degree)
	<i>median</i>	5 (undergraduate degree)
	<i>IQR</i>	4 (A-levels) - 5.5 (university degree)

Note. SES = socio-economic status; IQR = interquartile range; IQ was measured by matrix reasoning tests (Wechsler, 1999); SES was measured by parental education, a robust indicator of SES (Dubow, Boxer, & Huesmann, 2009).

The study was carried out in accordance with UCL Research Ethics Guidelines and approved by the UCL Research Ethics Committee. Informed consent from parents and assent from all participants was obtained.

6.2.2. Behavioural task

The *Academic Diligence Task* (ADT; Galla et al., 2014) is designed to mirror real-world choices students face when completing school-work. The task is available for preview and download as freeware: <https://angeladuckworth.com/research/academic-diligence-task/> The ADT consists of a split-screen interface with the choice to complete simple single-digit arithmetic problems (Figure 6.1A; e.g. $6 + 1$, 5×2 etc.) or play games (Tetris™ etc.).

Students were first shown an introduction screen that highlighted the benefits of practising maths equations: “New scientific research shows that students who practiced maths by doing more subtraction, addition and multiplication problems went on to earn higher grades. Even doing simple and easy maths problems can make you a better problem solver, which can help you in all areas of your life.” Participants then practised arithmetic problems before being instructed to solve as many maths problems as quickly and accurately as possible in the main task. They were also told that they always had the option to take a break and play games: “Remember, you will be able to play games whenever you feel like it, but the more problems you do, the better you will become at problem solving”. Students then completed three blocks during which they could freely allocate

their time between maths and games. After each block, participants were asked to rate how bored they felt on a 5-point scale (1 = not at all bored to 5 = very bored). After the last block, participants also rated how tempting they found the games on a 5-point scale (1 = not at all tempting to 5 = very tempting).

We operationally defined *diligence* as the percentage of time participants spent doing maths. The ADT took 20 minutes in total.

6.2.3. fMRI task

We used an *emotional go-no-go* task (Somerville et al., 2011) to measure self-control (Figure 6.1B). Participants were presented with happy or neutral faces and were instructed to respond to one of them by clicking a button (go stimulus, e.g. neutral faces) and not respond to the other (no-go stimulus, e.g. happy faces). We used adolescent faces as stimuli to reflect the importance of peers in this age group (Crone & Dahl, 2012). The stimuli were 18 girls' faces (happy and neutral expression for each) obtained from the NIMH-ChEFS adolescent face stimulus set (Coffman et al., 2015).

Two-thirds of stimuli were go and one-third no-go stimuli. This weighting was used to make the frequent go stimuli the pre-potent response and to increase the difficulty of inhibiting responses on infrequent no-go stimuli (Simmonds et al., 2008). Participants completed go-no-go conditions in which happy faces were the frequent go and neutral faces the infrequent no-go stimuli, and conditions in which neutral faces were the frequent go and happy faces the infrequent no-go stimuli.

Go-no-go blocks were interspersed with never-go blocks during which participants passively viewed faces. These blocks were used to control for potential confounds (see section 6.2.5). Stimulus frequency in never-go blocks was weighted just as in go-no-go blocks, i.e. blocks contained either infrequent happy and frequent neutral, or frequent happy and infrequent neutral faces.

Participants completed one functional run in which two-thirds of stimuli were happy faces, and one run in which two-thirds of stimuli were neutral faces. The order of the runs was counterbalanced between participants. Each of these runs consisted of eight blocks, four of which were go-no-go blocks and four never-go blocks. Each block consisted of 12 trials. A fixation cross was presented during a jittered (2000 - 7000 ms, $M = 4500$ ms) inter-stimulus interval. Each functional run took 8 min in total. The task was presented and responses were acquired with Cogent 2000 (Cogent 2000 Team, 2015) and Matlab (The MathWorks, 2013).

A recent review highlighted the need to test whether happy stimuli are actually rewarding (Foulkes & Blakemore, 2016). We therefore asked participants to rate how much they liked looking at each of the stimuli on a scale of -2.5 to +2.5 after the scanning session.

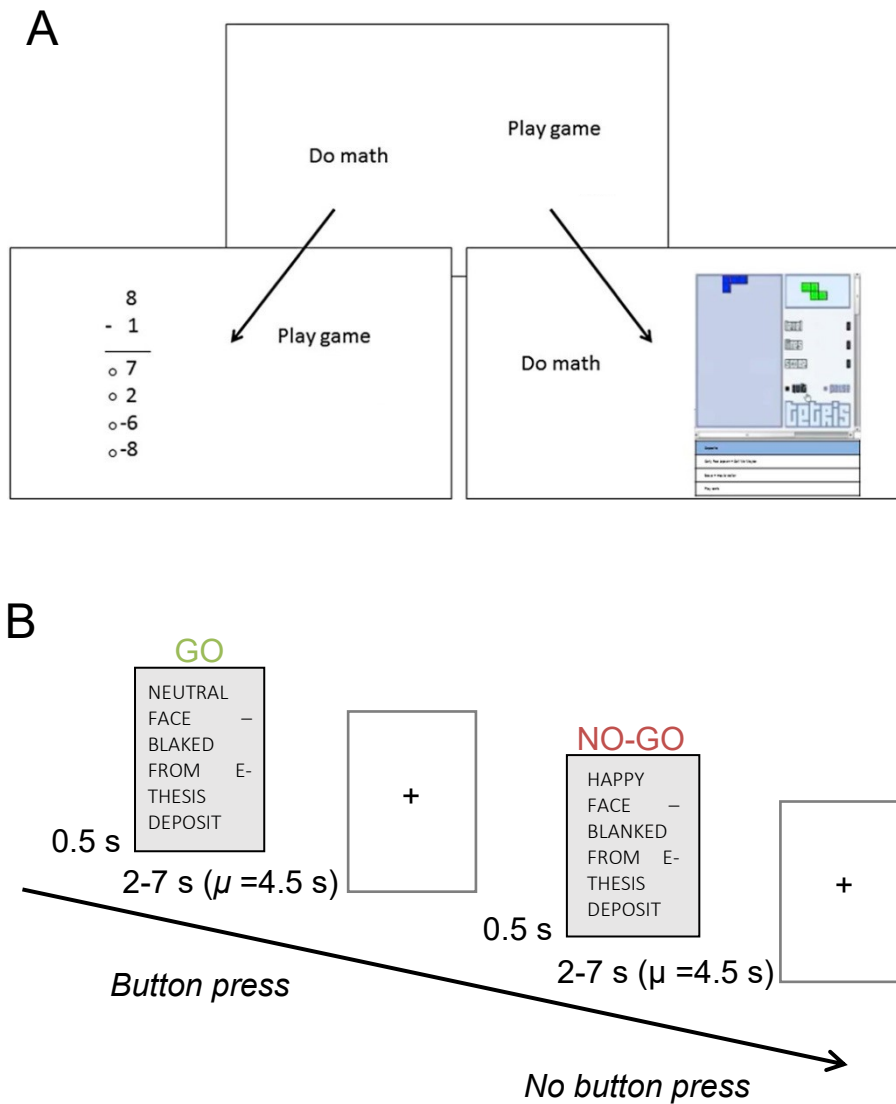


Figure 6.1. The Academic Diligence Task and Emotional Go-No-Go Task.

Panel A was adapted from Galla et al. (2014) and shows the Academic Diligence Task. Permission to reproduce this figure has been granted by Elsevier. Panel B shows the emotional go-no-go task (Somerville et al., 2011). Face stimuli were obtained from the NIMH-ChEFS adolescent face stimulus set (Coffman et al., 2015).

6.2.4. Imaging data acquisition and pre-processing

Imaging data were acquired using Siemens Avanto 1.5T MRI scanner. We ran a structural sequence (T1-weighted, 64 slices, TR = 1.17 s, TE = 0.01 s), two functional runs (T2-weighted, each run: 520 volumes, 44 slices, TR = 1 s, TE = 0.045 s) and a fieldmap in two sequences (each sequence: 64 slices, TR = 1.17 s, TE = 0.01 s). Each participant spent 30 min in the scanner.

Imaging data were pre-processed and analysed using SPM12 (Wellcome Trust Centre for Neuroimaging, 2014). To allow for T1 equilibration effects, the first eight volumes of each session were discarded. The EPI images were sinc interpolated in time for correction of slice-timing differences. Images were also realigned to the first scan by rigid body transformations to correct for head movements. The field map scans were pre-processed with the FieldMap toolbox (Andersson & Hutton, 2017) and used to correct for magnetic field distortions in functional scan.

Using a Gaussian kernel of full-width-half-maximum of 8 mm, EPI images (voxel size of $3\times 3\times 3$ mm³) and structural images (voxel size $1\times 1\times 1$ mm³) were co-registered and normalized to the T1 standard template in Montreal Neurological Institute space. Proportional scaling and high-pass temporal filtering with a cut-off of 128 s was applied to remove low-frequency drifts in signal.

Realignment estimates were used to calculate frame-wise displacement (FD) for each volume, which is a composite, scalar measure of head-motion across the six realignment estimates (Siegel et al., 2014). Volumes with FD > 0.9 mm were censored and excluded from further analysis by including a regressor of no interest for each censored volume in the general linear model (see section 6.2.5).

Scanning sessions with more than 5% of volumes censored or a root mean square movement over any run greater than 1.5 mm were excluded from the analysis. This applied to the two participants whose data were excluded from all analyses.

6.2.5. Functional analyses

Following pre-processing, statistical analyses were conducted on correct trials using a GLM. Activated voxels for inhibition (no-go > go trials) and emotion (happy > neutral trials) were identified using an epoch-related statistical model, convolved with a canonical haemodynamic response function and mean-corrected. The GLM included the main effects of inhibition and emotion, as well as their interaction.

To investigate the effects of inhibition, emotion, and the interaction between the two, we first conducted a whole-brain analysis (cluster-level $p < .05$ false discovery rate (FDR) corrected) with an exclusive mask for which we used contrasted infrequent never go trials with frequent go trials ($p = .001$). This mask was used to isolate activation due to inhibitory processes and to exclude activation due to the absence of a motor response or to viewing an infrequent stimulus.

To investigate the interaction between diligence, inhibition, and emotion we extracted activations of a-priori regions of interests (ROIs) using MarsBaR (Brett, Anton, Valabregue, & Poline, 2002). The ROIs were defined by Somerville et al. (2011) and consisted of 4 mm spheres in the right inferior frontal gyrus (IFG: $x =$

34, $y = 26$, $z = 0$), the dorsal striatum (DS: $x = 10$, $y = 16$, $z = 4$), and ventral striatum (VS: $x = -4$, $y = 15$, $z = -13$).

6.2.6. Connectivity analyses

We used psycho-physiological interaction (PPI) analysis to estimate task-related changes in connectivity between the IFG and other brain regions (Wellcome Trust Centre for Neuroimaging, 2014). The PPI analysis involved extracting the blood-oxygen-level dependent signal from the IFG ROI source region described above and forming the interaction term between the source signal and the eight conditions of our task. A second GLM analysis was then carried out that included the interaction term, the source region's extracted signal, the experimental factors and the movement regressors as effects of no interest. Participant-specific PPI models were run, and contrast images generated for each condition. These 'first level' contrast images were then entered into the full-factorial model to assess connectivity of the IFG during inhibition. We carried out a whole-brain analysis ($p < .05$ FDR-corrected) and extracted connectivity strength between the IFG and DS using MarsBaR (Brett et al., 2002).

6.2.7. Structural analysis

We analysed grey matter volumes within each of our ROIs using the CAT12 toolbox (Dahnke & Gaser, 2016). We estimated total intracranial volume (TIV) using a function provided by Ridgway (Ridgway, 2007). TIV was then added a covariate into the analysis to correct for differences in head size as recommended

by Peelle and colleagues (Peelle, Cusack, & Henson, 2012). Grey matter volume in the IFG, VS and DS ROI were extracted using MarsBaR (Brett et al., 2002).

6.2.8. Regression Models

To investigate the interaction between diligence and task-dependent activation, we implemented LMMs predicting structure, function and connectivity of the IFG, VS and DS using the lme4 (Bates et al., 2013) package in R (R Core Team, 2015). Significance tests were obtained using an omnibus Type III Wald χ^2 test. We built separate models for structure, function and connectivity because the structural analyses necessarily contained less fixed effects than the other two analyses (i.e. no contrasts for inhibition and emotion). For each ROI, we built one model predicting functional activation and one model predicting grey matter volumes. For connectivity, we built one model predicting connectivity between the IFG and DS. The functional models and connectivity model contained inhibition (no-go/go), emotion (happy/neutral) as orthogonal, Helmert-coded fixed effects. Diligence was included as a z-scored fixed effect in all models. We further included all possible interactions of the fixed effects. Participant ID and school were included as nested random intercepts. These random intercepts were used to reflect the repeated-measures design and the clustered nature of participants tested. Random slopes were not included in any model because their inclusion led to overfitting and non-convergence of models. The models predicting grey matter volumes included diligence as a z-scored fixed effect and school only as a random effect. Participant was not included as a random effect here because the structural models contained no repeated measures.

6.3. Results

6.3.1. Behaviour in the Academic Diligence Task

As expected, participants performed well on the simple arithmetic tasks (percentage accuracy: $M = 97.86\%$; $SE = 0.26\%$). Diligence scores, reflecting percentage of time spent doing maths, were high overall ($M = 84.14\%$, $SE = 2.70\%$) but individual scores ranged from 34.44% to 96.67%. Participants found maths moderately boring (ratings: $M = 2.71$ out of 5, $SE = 0.17$) and games moderately tempting (ratings: $M = 2.84$ out of 5, $SE = 0.20$). Diligence did not correlate significantly with IQ ($r(38) = -0.03$, $p = .841$) or SES ($r(31) = -0.06$, $p = .741$).

6.3.2. Behaviour in the go-no-go task

Participants gave happy faces in the emotional go-no-go task positive ratings, indicating that they found them rewarding to look at ($M = 0.58$, $SE = 0.04$). Neutral faces received negative ratings and were therefore not perceived as rewarding ($M = -0.22$, $SE = 0.04$). The difference between the two ratings was significant ($t(37) = 6.76$, $p < .001$). False alarm rates in the go-no-go task were low and did not significantly differ between happy ($M = 6.72\%$, $SE = 1.29\%$) and neutral faces ($M = 8.13\%$, $SE = 1.29\%$), ($\chi^2(1) = 0.96$, $p = .328$). There was, however, a difference in participants' reaction times ($t(39) = -3.38$, $p = .002$). Reaction times on happy trials ($M = 582.83$ ms, $SE = 22.71$ ms) were significantly faster than on neutral trials ($M = 622.21$ ms, $SE = 28.16$ ms), indicating that participants respond faster to rewarding social cues than to neutral ones.

In an exploratory analysis, we checked whether reaction times on happy go-no-go trials were associated with diligence. Although the direction of the association was negative, as would be expected, the correlation was not significant ($r(38) = -0.14, p = .406$).

6.3.3. fMRI results

Whole-brain results showed activation of mainly bilateral temporal clusters during inhibition (no-go > go; Table 6.2). No clusters survived cluster-level FDR-correction for emotion (happy > neutral) or for the interaction between inhibition and emotion.

Table 6.2. Results of the Cluster-Corrected Whole-Brain Analysis during No-Go Compared to Go Trials.

peak activation			z	cluster size	cluster level p_{FDR}	cluster location
x	y	z				
-51	5	-19	4.64	111	.023	L middle temporal gyrus and pole, posterior orbitofrontal cortex
51	-22	2	4.47	97	.023	R middle temporal gyrus
51	8	-25	4.35	87	.023	R inferior, middle temporal gyrus, middle temporal pole
-27	-97	2	4.17	125	.006	L inferior occipital gyrus

ROI analyses showed that there was no main effect of inhibition, or emotion for the IFG, VS or DS. There was an interaction between inhibition and emotion for the VS but not the IFG or DS (Table 6.3). VS activation was lower for happy no-go trials than for happy go trials while the reverse held for neutral trials (Figure 6.2). This indicates that the VS was activated more for trials where participants

responded to rewarding cues as compared to trials in which they withheld responses.

Diligence correlated significantly with IFG activation but not with activation in the DS or VS (Table 6.3). Participants with higher diligence showed higher activation of the IFG during the go-no-go task ($\beta = 0.20$; Figure 6.3). This effect was not moderated by inhibitory load or emotional valence of stimuli, however (Table 6.3).

Table 6.3. Results of Models Predicting Functional Activation of the IFG, DS and VS

IFG				
<i>Effect</i>	χ^2	<i>df</i>	<i>p</i>	
inhibition	1.87	1	.172	
emotion	0.17	1	.677	
diligence	4.03	1	.045	*
inhibition : emotion	0.62	1	.432	
inhibition : diligence	0.05	1	.833	
emotion : diligence	0.67	1	.412	
inhibition : emotion : diligence	0.30	1	.581	
DS				
<i>Effect</i>	χ^2	<i>df</i>	<i>p</i>	
inhibition	3.07	1	.080	
emotion	0.66	1	.418	
diligence	3.08	1	.079	
inhibition : emotion	1.57	1	.210	
inhibition : diligence	0.18	1	.670	
emotion : diligence	0.84	1	.360	
inhibition : emotion : diligence	1.21	1	.272	
VS				
<i>Effect</i>	χ^2	<i>df</i>	<i>p</i>	
inhibition	0.32	1	.571	
emotion	0.04	1	.845	
diligence	0.25	1	.617	
inhibition : emotion	4.58	1	.032	*
inhibition : diligence	0.11	1	.745	
emotion : diligence	0.28	1	.594	
inhibition : emotion : diligence	0.55	1	.460	

Note. * $p < .05$

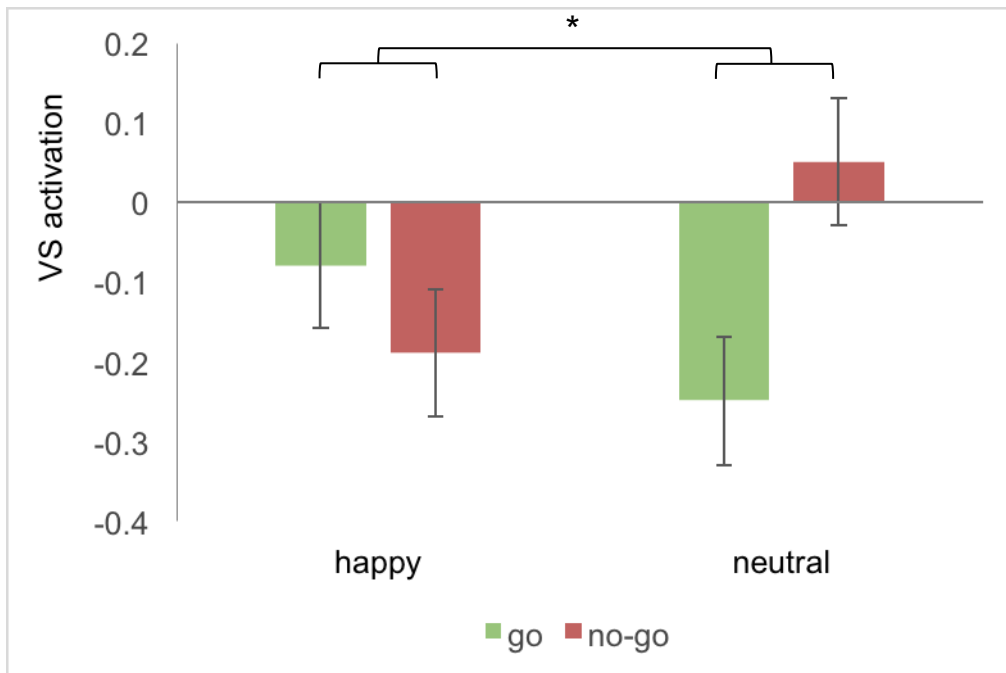


Figure 6.2. Ventral-Striatum Activation during the Emotional Go-No-Go Task. Mean predicted activation with standard error bars are shown for responding (go) and withholding responses (no-go) to happy and neutral faces. VS = ventral striatum (* $p < .05$).

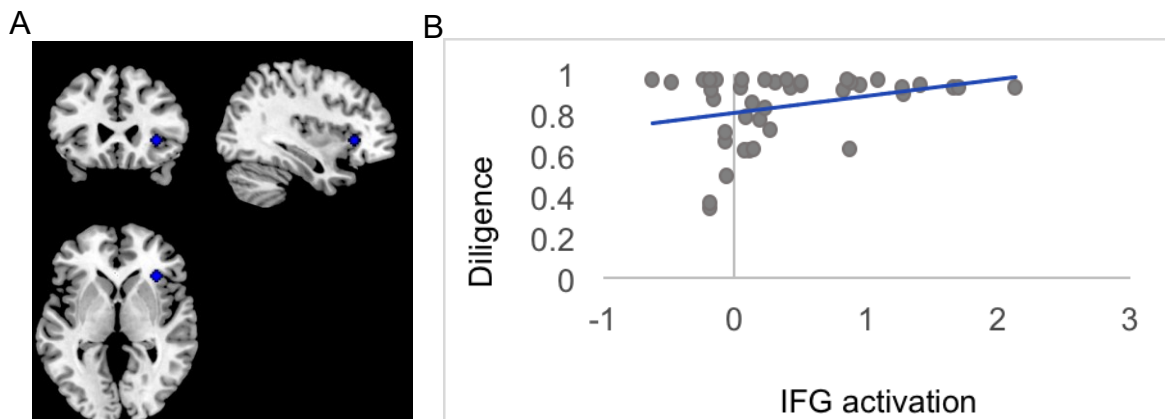


Figure 6.3. Inferior Frontal Gyrus ROI and Correlation with Diligence. Panel (A) shows the ROI of the inferior frontal gyrus (IFG). Panel (B) shows IFG activation during the emotional go-no-go task by diligence (proportion of time spent doing maths rather than playing games).

6.3.4. Connectivity results

Connectivity between the IFG and DS was not significantly associated with inhibition, emotion diligence, or any of their interactions (*Table 6.4*).

Table 6.4. Results of a Model Predicting Connectivity between the IFG and DS

Effect	χ^2	df	p
inhibition	3.08	1	.079
emotion	0.00	1	.949
diligence	0.76	1	.383
inhibition : emotion	0.01	1	.923
inhibition : diligence	0.21	1	.645
emotion : diligence	1.16	1	.282
inhibition : emotion : diligence	3.73	1	.053

6.3.5. Structural results

Grey matter volumes did not significantly correlate with diligence for any of our three ROIs (IFG: $\chi^2(1) = 0.01$, $p = .919$; DS: $\chi^2(1) = 1.15$, $p = .285$; VS: $\chi^2(1) = 1.15$, $p = .283$).

6.4. Discussion

The current study investigated the neurocognitive correlates of academic diligence, a predictor of educational attainment. We assessed whether individual differences in diligence during adolescence were related to the interplay between inferior frontal self-control and striatal reward systems, as predicted by the dual-systems hypothesis (Casey et al., 2008; Duckworth & Steinberg, 2015; Steinberg, 2008). The results were mostly inconsistent with the dual-systems hypothesis.

There was a link between inferior frontal activation and diligence. However, there was no clear association between diligence and striatal structure and function, or diligence and connectivity between frontal and striatal regions. Instead, we found widespread activation of temporal areas during the go-no-go task.

The functional ROI analysis provided some evidence that frontal activation was associated with diligence, in line with the dual-systems hypothesis (Duckworth & Steinberg, 2015). Activation of the inferior frontal gyrus during the emotional go-no-go task correlated positively with diligence, although this association was not dependent on inhibitory load or emotional valence. This finding is similar to resting-state studies linking prefrontal activation to self-control (Gianotti et al., 2009; Knoch, Gianotti, Baumgartner, & Fehr, 2010) and consistent with evidence from previous go-no-go studies showing a positive correlation between right inferior frontal activation and self-control (Simmonds et al., 2008; Somerville et al., 2011). It also complements lesion and correlational studies in adults showing that the personality trait conscientiousness, which is closely related to other measures of self-control (Credé et al., 2016), is associated with lateral frontal functioning (DeYoung et al., 2010; Forbes et al., 2014).

However, most predictions made by the dual-systems hypothesis were not supported by our data. There was no clear link between diligence and striatal functional activation and structure, or between diligence and inferior frontal gyrus structure or connectivity. We did find that the ventral striatum showed greater activation for trials in which participants responded to happy adolescent faces as compared to trials in which they withheld responses to these faces, while the pattern of activation was reversed for neutral faces. This indicates that the ventral

striatum responded preferentially to trials on which participants were looking out for subjectively rewarding social cues. This is in line with previous findings of heightened ventral striatum activation in response to rewards in adolescence (Braams et al., 2015; Haber, 2016; Somerville et al., 2011; van Leijenhorst et al., 2010). However, this pattern of striatal activation was not directly related to individual differences in diligence in our study.

A surprising finding was the prominent activation in the temporal cortex during the inhibition task: The whole-brain analysis showed that the emotional go-no-go recruited mainly temporal regions. While many previous inhibitory control studies have mostly focussed on frontal regions (Simmonds et al., 2008; Somerville et al., 2011), there are now several go-no-go studies in adolescents that have also shown prominent temporal activation. MEG (Vara, Pang, Doyle-Thomas, et al., 2014; Vara, Pang, Vidal, Anagnostou, & Taylor, 2014) and fMRI studies (Tamm, Menon, & Reiss, 2002) have found that adolescents recruit temporal regions, particularly the right temporal sulcus, more than adults during go-no-go tasks. This recruitment of temporal regions has been proposed to support frontal functioning during development (Vara, Pang, Doyle-Thomas, et al., 2014). Future studies may benefit from probing the interaction between more extended networks than just fronto-striatal systems to better understand the development of self-control during adolescence.

It is possible that some of our null findings are due to limitations of our sample or the tasks used. There was a range of individual diligence scores (34.44% to 96.67%), but diligence was high overall in our sample: participants and chose to do simple and boring maths over playing games 84.14% of the time on average. It

is possible, that larger sample sizes with even more variability in diligence are needed to detect a stronger correlation between diligence and brain structure and function. This possibility should be investigated in future studies - so far only a handful of studies have investigated neural correlates of diligence and related constructs and very few have probed striatal functioning (C. A. Myers et al., 2016; Nemmi et al., 2016).

Another limitation of this study is the limited amount of data available on the subjective experience of the emotional go-no-go and particularly the Academic Diligence Task. We directly probed how participants perceived our emotional go-no-go stimuli and found evidence that happy, but not neutral, faces were perceived as rewarding. The Academic Diligence Task also included questionnaire items probing participants' boredom during maths and temptation by games: our participants found maths relatively boring and games relatively tempting. Similarly, an earlier, larger validation study by Galla and colleagues (2014) used multilevel growth curve models to show that boredom during the Academic Diligence Task increased over time if participants chose maths, but not if they chose games. They also found that higher levels of boredom and temptation were linked to lower diligence. The Academic Diligence Task could be improved in future studies by including the same questionnaire items (temptation and boredom) for both maths and games to allow for a more direct comparison. Nonetheless, the evidence available to date suggests that the task is likely to capture every-day conflicts between the wish to pursue educational goals and the temptation to engage in more pleasurable distractions.

Despite these limitations, our study has some tentative implications for the study of adolescent self-control. It echoes previous theoretical work highlighting the limited ability of the dual-systems framework to explain the wide range of adolescent self-control observed in naturalistic settings (Crone & Dahl, 2012; Pfeifer & Allen, 2012). In conjunction with previous research, it also highlights that it may be useful to move away from the duality of fronto-striatal systems and instead explore more extended brain networks (Baum et al., 2017; Vara, Pang, Vidal, et al., 2014). More avenues for future research, as well as wider implication of this study, and the other studies in this thesis, is discussed in the next chapter, the overall discussion.

7. Chapter 7: Discussion ⁶

There is a tension between the wide-spread assumption that earlier is always better for learning and a growing body of research highlighting that the human brain and mind undergoes changes past childhood. Previous research showed that adolescence in particular is characterized by extensive changes in brain structure, paralleled by protracted development of cognitive functions relevant to education. These changes in brain structure have been linked to the protracted development of cognitive and socio-cognitive skills and have led to the suggestion that adolescence is a period of relatively high levels of plasticity, during which the environment has a heightened impact on brain development and behaviour. This thesis investigated this proposition in three behavioural studies and one neuroimaging study. The findings of each of these studies will be summarised in this chapter and synthesized in terms of their implications for policy and practise. Methodological limitations will be discussed, as well as directions for future research.

⁶ Parts of this chapter have been published as:

Fuhrmann, D., Knoll, L.J., Blakemore, S.-J. (2015). Adolescence as a sensitive period of brain development. *Trends in Cognitive Sciences*, 19 (10), 558-566. doi:10.1016/j.tics.2015.07.008

Fuhrmann, D.*, Knoll, L.J.*, Sakhardande, A., Stamp, F., Speekenbrink, M. & Blakemore, S-J. (2016). A window of opportunity for cognitive training in adolescence. *Psychological Science*, 27(12), 1620-1631. doi: 10.1177/0956797616671327 *Joint first authors.

Fuhrmann, D., Knoll, L.J., Sakhardande, A., Speekenbrink, M., Cohen Kadosh, K. & Blakemore, S-J. (2016). Perception and recognition of faces in adolescence. *Scientific Reports*, 6(33497), doi:10.1038/srep33497

Fuhrmann, D. Knoll, L.J., Sakhardande, A. & Blakemore, S-J. (2016). When to teach what: Are there sensitive periods for learning in adolescence? Nuffield Foundation, London, UK. URL: <http://www.nuffieldfoundation.org/sites/default/files/files/Blakemore%20%20Nuffield%20Main%20Report%2012%20Jan%202017%20Final.pdf>

7.1. Summary of Findings

The research presented in the current thesis investigated four main research questions: (I) Are there age-related differences in face cognition between adolescence and adulthood?; (II) Do the effects of social exclusion on cognitive performance differ between age groups?; (III) Do some age groups benefit more from cognitive training than others?; and (IV) What are the neurocognitive correlates of academic diligence? Each of these questions is addressed below.

7.1.1. Are there age-related differences in face cognition between adolescence and adulthood?

Face cognition is a fundamental building block of social cognition. Faces are a pre-eminent social signal involved in almost all aspects of social cognition and processed by a distributed neural network overlapping substantively with networks for other aspects of social cognition (Blakemore, 2012; Blakemore & Mills, 2014; Scherf et al., 2012). Thus, the development of face cognition between childhood and adulthood has been proposed to serve as a model for broader developments in sociality during adolescence (Scherf et al., 2012).

Despite the growing interest in adolescent face cognition, most studies on the development of face cognition abilities have still focussed on childhood, and it has been unclear whether there are still substantive changes in face cognition during adolescence. Early maturation accounts contend that face cognition abilities are mature by 3 - 5 years of age (Crookes & McKone, 2009; McKone et al., 2012). Late maturation accounts, in contrast, propose that at some aspects of face cognition

are not mature until at least 10 years (Cohen Kadosh, Johnson, Dick, et al., 2013; Mondloch et al., 2002).

In Chapter 3, we measured face memory and face perception, two core face cognition abilities to probe whether there are age-related differences in face cognition past childhood. We recruited 661 participants (395 females) and split them into four age groups: younger adolescents (11.27 - 13.38 years), mid-adolescents (13.39 - 15.89 years), older adolescents (15.90 - 18.00 years), and adults (18.01 - 33.15 years).

We showed that younger adolescents' and mid-adolescents' (11.27 to 15.89 years) face memory and face perception abilities were less proficient than those of older adolescents and adults (15.90 and over). General cognitive ability predicted face cognition scores but showed a different developmental trajectory with continuous improvements throughout adolescence and into adulthood. Thus, face cognition abilities mature relatively late, at around 16 years of age, and this protracted development is at least partly independent of general cognitive ability.

The improvements in the ability to memorize, recognize and perceive faces during adolescence may be related to broader changes in adolescents' social lives. Adolescents spend less time supervised by parents and caregivers and increasingly engage with peers (Lam et al., 2014). This may lead to more exposure to novel faces than earlier in life. New social roles in adolescence may also increase the focus on facial information such as attractiveness and status (Scherf et al., 2012). This may, in turn, provide the environmental enrichment necessary

for becoming a face expert (Bukach et al., 2006). Whether the maturation of face cognition is indeed driven by environmental inputs remains to be tested in future intervention studies examining age-related differences in learning face recognition, identity, expression and gaze perception. More work is also needed to establish to what extent developmental trends in face cognition model changes in social cognition in general. Such studies will help us better understand whether adolescence is a sensitive period for social development, during which time the brain is particularly susceptible to changes in socio-cultural information (Andersen & Teicher, 2008; Blakemore & Mills, 2014).

7.1.2. Do the effects of social exclusion on cognitive performance differ between age groups?

If adolescence is indeed a sensitive period of social development, we would expect social stress to have a particularly detrimental effect on adolescents (Buwalda et al., 2011; Fuhrmann et al., 2015). Most studies that have experimentally investigated social stress in adolescence have used animal models. Such studies have shown that adolescent rats show reduced recovery from defeat stress compared to adult rats, for instance (Ver Hoeve et al., 2013). Rodent studies have also highlighted that social isolation may have similarly detrimental effects as direct confrontation during adolescence. Social isolation during adolescence (but not earlier or later) was shown to have irreversible effects on exploratory behaviour in rats (Einon & Morgan, 1977). However, little experimental evidence is available for the effects of social isolation and exclusion human adolescence.

Chapter 4 investigated and compared the effects of cyber-ostracism on cognitive performance in 99 females in three age groups: young adolescents (10.1 - 14.0 years), mid-adolescents (14.3 - 17.9 years) and adults (18.3 - 38.1 years). Participants in all age groups completed verbal n-back and visuospatial dot-matrix working memory tasks, as well as a mood questionnaire, after social inclusion and social exclusion in the online ball-tossing game Cyberball.

All age groups showed a similar and significant drop in mood after Cyberball exclusion but the effect of social exclusion on cognitive performance was age-dependent. Only young adolescents (aged 10.1 - 14.0 years) showed a reduction in performance on n-back and visuo-spatial working memory tasks after social exclusion. In contrast, performance did not differ between social inclusion and exclusion for mid-adolescents and adults (aged 14.3 and over).

These results are relevant to understanding the effects of ostracism in schools. They suggest that experiencing social exclusion may place a particular burden on young adolescents, compared to older age groups. Exclusion reduces cognitive performance, which, in turn, may impact educational achievement (Nakamoto & Schwartz, 2010; Rigby, 2000; Sharp, 1995). This underlines the need to develop effective ostracism interventions in schools and to consider age differences in response to social exclusion in the design and timing of interventions. As the study did not include any younger age groups, however, we do not yet know whether sensitivity to social exclusion peaks in early adolescence or instead decreases continuously from early childhood. This question needs to be addressed by future cross-sectional and longitudinal studies.

7.1.3. Do some age groups benefit more from cognitive training than others?

If adolescence is a period of high levels of plasticity, as suggested by the previous two studies, we would not only expect stress to have detrimental effects during adolescence but also predict that enrichment should have positive effects during this time of life (Fuhrmann et al., 2015). However, education policies often focus on early education as a key window for interventions (Heckman, 2000, 2006). Much less is known about the effect of enrichment in older age groups. In Chapter 5, we used a cognitive training intervention to investigate the possibility that cognitive skills related to maths performance in schools are more efficiently acquired later in development.

A total of 663 adolescents and adults aged 11 - 33 years were trained in different cognitive tasks for 10 minutes a day, for up to 20 days. Participants were tested on a range of cognitive tasks before and after the training, and several months after training had ceased. They were assigned at random to one of three training groups. One training group was trained to discriminate small from large numerosities, an important skill, as we often have to compare and judge quantities in our everyday life (Halberda & Feigenson, 2008). The second group was trained in relational reasoning, which is the ability to detect abstract relationships between groups of items and is related to fluid intelligence (Mackey et al., 2011). Both numerosity discrimination and relational reasoning correlate with mathematics performance (Halberda et al., 2008; Knuth, Kalish, Ellis, Williams, & Felton, 2012). The third group was trained in face perception, the control task. As discussed in section 7.1.1, face perception also improves during adolescence and may be susceptible to training. However, it was chosen as a

control task as it relies on different cognitive processes and neural circuits than those involved in numerosity discrimination and relational reasoning (Cantlon, Brannon, Carter, & Pelphrey, 2006; Cohen Kadosh, Johnson, Henson, et al., 2013; Mackey, 2012).

Training in the numerosity discrimination task yielded some improvement in performance, but only in late adolescence and adulthood (aged 15.90 and over). All age groups improved when trained in relational reasoning, but older adolescents and adults showed the highest training benefits (again, aged 15.90 and over). Face perception showed some training effects overall but no age-related differences emerged. There was no evidence of far-transfer from relational reasoning training to working memory performance or from face perception training to face memory performance.

These findings suggest that skills related to mathematics are more efficiently learned in late adolescence and adulthood than earlier in adolescence. This highlights the relevance of this relatively late developmental stage for learning and, in contrast to the common assumption that 'earlier is better' for learning, underlines the need to investigate late adolescence as a potential window of opportunity for educational interventions. As the study did not include older adults, no conclusions can be drawn about the offset of sensitive periods, i.e. we do not know at what age training effects might start to reduce again. We also cannot be sure what the mechanisms of developmental increases in training effects are (e.g. plasticity, strategy learning or skill learning). These questions will need to be elucidated by future studies combining cognitive training with neuroimaging.

7.1.4. What are the neurocognitive correlates of academic diligence?

While the previous three studies compared average behaviour of adolescent and adult age groups, the final study of this thesis took a closer look at a relatively narrow age range (14 to 15 years) and sought to investigate individual differences in adolescent brain structure and function.

Adolescence is often described as a time of protracted development of self-control and increased reward seeking (Casey et al., 2008; Steinberg, 2008; Steinberg et al., 2017). The dual-systems hypothesis holds that this developmental 'imbalance' originates in the relative maturity of subcortical reward systems compared to late-developing frontal control systems (Casey et al., 2008; Hall, 1904; Steinberg, 2008). Most dual-systems studies have focussed on explaining increased risk-taking behaviours in adolescence (Casey et al., 2008; Steinberg, 2008; van den Bos et al., 2015). However, reward processing and self-control are also known to affect many other phenomena including educational outcomes (Duckworth & Steinberg, 2015; Steinbeis & Crone, 2016).

In Chapter 6, we investigated whether the interplay between inferior frontal control and striatal reward systems is related to individual differences in academic diligence, the ability to pursue long-term educational goals, and a predictor of educational attainment (Galla et al., 2014). We combined behavioural data, structural MRI and functional MRI during an emotional-go-no-go self-control task to assess neurocognitive correlates of diligence in 40 girls aged 14 - 15.

The results only partially supported the dual-systems hypothesis by showing a link between inferior frontal activation and diligence. However, we found no clear

association between diligence and inferior frontal structure or striatal function and structure. Instead, we found prominent activation of temporal areas during the go-no-go task. It is possible that some of our null-findings are due to limitations of our sample. There was a range of individual diligence scores (34.44% to 96.67%), but diligence was high overall in our sample: participants chose to do simple and boring maths over playing games 84.14% of the time on average. It is possible that larger sample sizes with even more variability in diligence are needed to detect a stronger correlation between diligence and brain structure and function. This possibility will have to be investigated by future studies - so far there have only been a handful of studies investigating neural correlates of diligence and related constructs and very few have probed striatal functioning (Myers et al., 2016; Nemmi et al., 2016).

Nonetheless, our findings echo previous studies and reviews highlighting the limited ability of the dual-systems framework to explain the wide range of adolescent self-control observed in naturalistic settings (Crone & Dahl, 2012; Pfeifer & Allen, 2012). In conjunction with previous research, it also highlights that it may be useful to move away from the focus on the duality of fronto-striatal systems and instead explore more extended brain networks, including temporal areas (Baum et al., 2017; Vara, Pang, Doyle-Thomas, et al., 2014; Vara, Pang, Vidal, et al., 2014).

7.2. Implications for Policy and Practise

The findings presented in this thesis highlight some facets of adolescent development that may inform educational policy and practise. This section discusses implications of this research in the context of previous theoretical and empirical work and highlights five areas of adolescent development particularly relevant to education: (I) protracted social development; (II) malleability of cognitive ability; (III) enhanced learning of complex cognitive skills; (IV) limited transfer from cognitive training; and (V) individual differences.

7.2.1. Adolescence as a key period for social development

Chapter 3 showed that even very fundamental aspects of social cognition such as face identity perception and recognition undergo protracted development during adolescence and do not mature until around 16 years of age. This finding mirrors a growing body of work from many labs around the world, identifying adolescence as a time of substantial development in almost all aspects of social cognition including mentalizing (Dumontheil, Apperly, et al., 2010), processing social emotions like embarrassment or guilt (Burnett, Bird, Moll, Frith, & Blakemore, 2009; Goddings et al., 2012), peer influence (Blakemore, 2012; Casey, 2013; Crone & Dahl, 2012; Kilford et al., 2016; Steinberg, 2008), and sharing (Meuwese, Crone, de Rooij, & Guroglu, 2015; Steinbeis & Singer, 2013).

These changes in social cognition highlight that adolescence is a key time for social development (Blakemore & Mills, 2014). As such, social stress and bullying during adolescence can have lasting effects on mental health and educational

attainment (Blakemore & Mills, 2014; Fuhrmann et al., 2015). Bullying is relatively common in childhood and adolescence: about 34% of schoolchildren are bullied. Victims are often perpetrators themselves and only about 42% of schoolchildren are neither bully nor are bullied themselves (Forero, McLellan, Rissel, & Bauman, 1999). Bullying is bi-directionally associated with severe and long-lasting effects on mental health (Arseneault, Bowes, & Shakoor, 2009). Children with depressive symptoms are more likely to be bullied (Arseneault et al., 2006) and chronically bullied children are more likely to have suicidal ideations and depression up to 17 years later - particularly among girls (Arseneault et al., 2009; Klomek et al., 2009). Correlational studies have linked bullying to reduced educational attainment, as well as mental health problems (Rigby, 2000; Sharp, 1995; Sigurdson, Undheim, Wallander, Lydersen, & Sund, 2015).

Chapter 4 built on previous research and provided experimental evidence that social exclusion, a common form of bullying (J. Wang, Iannotti, Luk, & Nansel, 2010), can have detrimental effects on cognitive performance in adolescence, particularly before the age of 14.

This work highlights the importance of adolescence for social development and underlines the need for adequate, age-appropriate social and mental health support in schools so that schools can foster an enriching, supportive learning environment for adolescents. School-based bullying interventions are usually relatively effective. A meta-analysis found that bullying decreased by 20 - 23% on average after such interventions (Ttofi & Farrington, 2011). Programs involving parents (e.g. parent training) were found to be particularly effective while interventions involving peers (e.g. peer mentoring) were found to be counter-

productive and led to an increase in victimization (but see Paluck, Shepherd, & Aronow, 2016).

7.2.2. Cognitive ability is not fixed

The study presented in Chapter 5, as well as previous work by others (Crone et al., 2009; Dumontheil, Houlton, et al., 2010; Li et al., 2004; Mackey et al., 2011; Mackey et al., 2013), showed that reasoning is not fixed after childhood. In a lifespan study, Li and colleagues showed that reasoning scores increase up until the mid-twenties and start to decline only in the mid-thirties (Li et al., 2004). Reasoning has also been found to be highly susceptible to training. Mackey and colleagues (2010) showed that children aged 7 - 9 from low socio-economic background increase their IQ by 10 points after 8 weeks of reasoning training. In 2013, the same group provided evidence that reasoning training induces plasticity by showing that 70 h of reasoning training in adults (mean age: 22) altered brain connectivity at rest and strengthened fronto-parietal and parietal-striatal connections (Mackey et al., 2013). The research presented in Chapter 5 added to this previous evidence by showing that reasoning training is effective throughout adolescence and early adulthood; and that the effects of reasoning training even increase over this age range.

This is relevant to education. Matrix reasoning tests, like those used in the training study in Chapter 5, were originally designed to measure aptitude and are often a part of IQ and fluid intelligence tests (Wechsler, 1999), as well as school-entrance exams (Spring, 2016). However, the findings from this area of research do not

support the notion that matrix reasoning gives an indication of some kind of innate, fixed ability (Jensen, 1969). Instead, they highlight that reasoning is a late-developing, highly malleable skill. This calls into question whether IQ tests can be ‘tutor-proof’ and thus suitable for school entry exams.

7.2.3. Earlier is not always better for learning

At present, education policy tends to emphasize the importance of investing in early-childhood intervention (Allen & Smith, 2009; Barnett, 2011; Heckman, 2000, 2006). Plasticity is undoubtedly high early in development, making the early environment of primary importance for the development of many aspects of cognition, including visual or language development (Kuhl, 2010; Maurer & Lewis, 2012). However, previous research has called into question the rather extreme conclusions that have sometimes been drawn from research into early sensitive periods, particularly the ‘myth of three’ - the idea that the first three years of life are *the* critical period for learning after which educational trajectories are more or less fixed (Howard-Jones et al., 2012; Wenger & Lövdén, 2016).

Numerous studies have shown that complex cognitive skills relevant to education, including reasoning, numeracy and cognitive control, continue to develop into adolescence and adulthood (Baltes et al., 2006; Baum et al., 2017; Halberda et al., 2012; Li et al., 2004; Murre et al., 2013). There is also good evidence that these skills can be trained into adulthood and even old age (Baltes et al., 2006; Mackey, 2012; Mackey et al., 2011; Melby-Lervåg & Hulme, 2013). Additionally, the study presented in Chapter 5 showed that both reasoning and numeracy might be

trained more effectively relatively late in development, from 15 onwards, than earlier in adolescence.

These findings may explain why some early childhood interventions have had limited success (Barnett, 2011; Howard-Jones et al., 2012; Leonard, 2000; Wenger & Lövdén, 2016). For instance, 'hot-housing', i.e. attempts to teach young children complex cognitive skills like reading and arithmetic, have not been very effective (Barnett, 2011; Wenger & Lövdén, 2016). It is likely that attention and working memory have to be sufficiently developed before complex cognitive skills relying heavily on these executive functions can be learned (Livesey & Dawson, 1981; Thomas, 2012). Therefore reading and complex arithmetic may simply not be within the 'zone of proximal development' (Vygotsky, 1978) of very young children. Educational efforts for young children may be more effective when focussed on foundational language, motor and social skills, which will help children acquire more complex cognitive skills later in development (Kuhl, 2004; Thomas, 2012).

Our findings support previous calls and efforts to invest in life-long learning (Knowland & Thomas, 2014; Thomas, 2008; Wenger & Lövdén, 2016). Early investment may be necessary to create educational potential, but without continued educational investment throughout childhood, adolescence and adulthood, learners may not be able to realize this potential (Knowland & Thomas, 2014; Livesey & Dawson, 1981). Programs facilitating the acquisition of cognitive skills like literacy past childhood may also be critical for improving quality of life for disadvantaged adolescents and adults, particularly in the developing world (Deshpande, Desrochers, Ksoll, & Shonchoy, 2017). In ageing

Western countries, lifelong education is becoming increasingly important for national economies as well (Knowland & Thomas, 2014; The Royal Society, 2011).

7.2.4. Transfer effects are limited

There was no evidence of transfer from cognitive training to untrained cognitive tasks in the training study presented in Chapter 5. Training on relational reasoning, for example, had no effect on the other skills tested, including working memory, even though these skills are known to be related (Kane et al., 2004). This mirrors much of the cognitive training literature, which shows that although transfer to closely-related skills (*near-transfer*, e.g. from one working memory task to another) is relatively robust (Schwaighofer et al., 2015), cognitive training often does not generalize to other, less closely related cognitive domains or classroom behaviour (*far-transfer*; Dunning, Holmes, & Gathercole, 2013; Melby-Lervåg & Hulme, 2013; Owen et al., 2010; Schwaighofer et al., 2015).

Because of this lack of robust evidence of generalizability, commercial cognitive training programs available may not be useful in the classroom yet (Goswami, 2006; Owen et al., 2010). When using cognitive training as a tool for education or interventions, far-transfer is important. The aim of practitioners is usually not only to improve performance on specific tasks, but to train generalizable skills (see Chapter 5). Future research will therefore need to invest in improving the efficacy of cognitive training interventions.

7.2.5. Individual differences matter

The extent of variability in brain and cognitive development between people of the same age is substantial (Baltes et al., 2006; Mills et al., 2014; Mills & Tamnes, 2014). This variability is sometimes as large, or larger than, variability between age groups (Figure 7.1).

Adolescence has been proposed to be a time of heightened inter-individual variability, particularly in terms of self-control (Crone & Dahl, 2012) and in the maturation of the brain structures sub-serving it (Mills et al., 2014). On average, adolescents may be hyper-sensitive to rewards (van Leijenhorst et al., 2010) and more impulsive than other age groups (Somerville et al., 2011). At the level of the individual, however, many adolescents show remarkable levels of self-control (Crone & Dahl, 2012; Galla et al., 2014). Chapter 6, for instance, highlighted that many adolescents are able to diligently complete a tedious academic task. This highlights the need to refine our understanding of the neurocognitive mechanisms of adolescent self-control and to develop alternatives to the dual systems hypotheses. Promising avenues include exploring whether self-control consist of separable sub-components, which may develop at different rates (Crone & Steinbeis, 2017). Future research should also develop candidate neurocognitive mechanisms for developmental increases in self-control other than fronto-striatal systems. Probing changes in modularity of extended neural networks may be a particularly promising line of research here (Baum et al., 2017).

More generally, this evidence highlights that most of the effects discussed in this thesis are likely moderated by individual differences. Training effects may differ

systematically between people, with motivation and diligence predicting training gains (Jaeggi, Buschkuhl, Shah, & Jonides, 2014). Reactions to stress will also show variability between adolescents. Friendship support, for instance, has been shown to increase resilience to stress (van Harmelen et al., 2017). The individual differences approach to development thus highlights exciting avenues for extending and refining our understanding of adolescence in future research (Brown, 2017; Kievit et al., 2017).

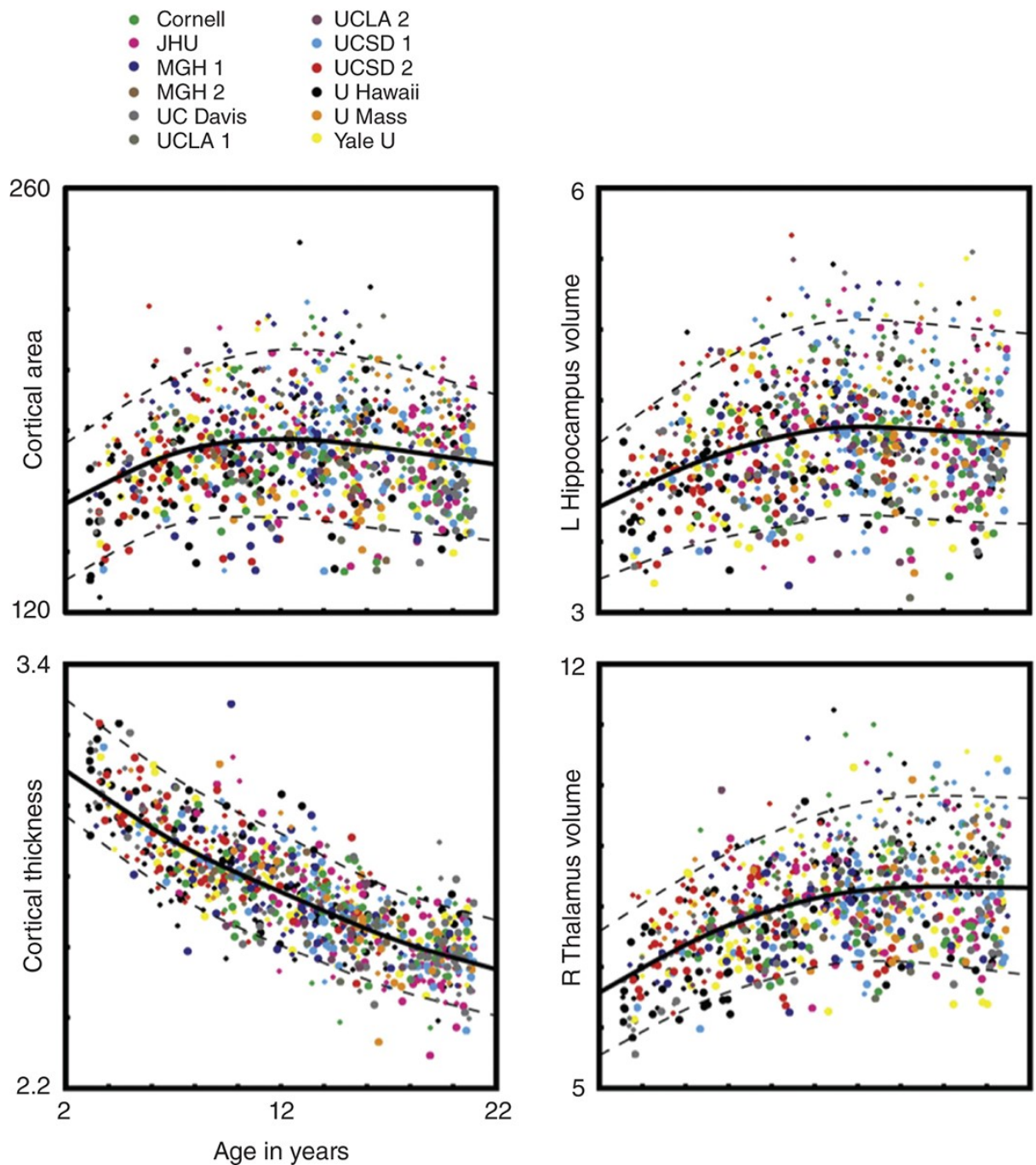


Figure 7.1. *Individual Differences in Brain Maturation*. Measures of brain structure of 886 participants aged 3 to 20 (reprinted with permission from Elsevier from Brown et al., 2012). The four panels show examples of different measures of brain morphology: total cortical area in mm^2 by thousands (upper left), mean cortical thickness in mm (lower left), volume of the left hippocampus in mm^3 by thousands (upper right), and volume of the right thalamus in cubic mm^3 by thousands (lower right). A spline-fit curve (solid line) with 95% confidence

intervals (dashed lines) is shown. Larger circles represent female participants, smaller circles male participants. Different sites and scanners are colour-coded.

7.2.6. Specificity of developmental trajectories

Based on work on plasticity in early childhood (Kuhl, 2004; Lewis & Maurer, 2005), and differences in the timing of maturation of different brain regions during adolescence (Tamnes et al., 2013), we predicted in Chapter 1.3.1 in the introduction that different cognitive domains should show variation in the on- and offset of heightened sensitivity to environmental input. The experimental findings from this thesis are consistent with this proposal.

Chapter 3 showed that developmental trajectories show similarities within domains and differences between domains. Both face memory and face perception, as core components of face cognition, mature at around 16 years of age, while reasoning, a skill unrelated to face cognition, shows a divergent developmental trajectory and continues to mature into adulthood.

Chapter 4 and 5 highlighted that sensitivity to negative environmental input and positive environmental input might differ, too. Chapter 4 showed that cognitive performance was particularly vulnerable to disruption by social exclusion during early adolescence (between ages 10 and 14), but not later in development. In contrast, Chapter 5 showed that environmental enrichment, in the form of cognitive training, was more effective in older age groups (above the age of 16) compared to younger age groups. This effect in and of itself showed specificity. Heightened learning in older age groups was restricted to maths-related skills

(numerosity discrimination and relational reasoning) and was not evident for face perception.

This highlights the notion that interventions need to carefully consider the specific developmental trajectories and periods of sensitivity of the cognitive skill they are interested in targeting (e.g. working memory, reasoning, self-control). Previous research on learning and plasticity of other skills may not be sufficiently informative of when in development interventions may be most effective.

7.3. Methodological Considerations

Apart from the specific methodological limitations associated with each study discussed in the experimental chapters and the summary of findings above, there are some general limitations of the experimental chapters of thesis that will need to be addressed by future studies. These limitations include the following issues: (I) the use of cross-sectional designs; (II) possible gender differences; (III) the need for naturalistic measures; and (IV) the need to investigate broader age ranges.

7.3.1. Cross-sectional versus longitudinal designs

As highlighted in the Chapter 3, the methods chapter, the choice of cross-sectional versus longitudinal designs affects the kind of conclusions we can draw from our data. While the training study presented in Chapter 5 had longitudinal components, Chapter 3, 4, and 6 were purely cross-sectional. We chose cross-sectional designs because their cost- and time-effectiveness. It should be noted,

however, that our ability to describe developmental change in cross-sectional studies is limited due to inter-individual variability (Grimm, Davoudzadeh, & Ram, 2017; Little, 2013).

The study on age-related differences in face cognition presented in Chapter 3, in particular, should be seen as a basis for designing future longitudinal studies that aim to investigate developmental change within the same set or subset of people. Similarly, longitudinal studies observing developmental changes in mean and variance of diligence (see Chapter 6) would be highly informative about developmental trends.

7.3.2. Gender differences

It is conventional for smaller developmental behavioural and neuroimaging studies to include only one gender. The studies presented in Chapter 4 and 6 of this thesis, for instance, included only girls. The main reason for this were the well-established differences in pubertal onset between the sexes (Sisk & Foster, 2004).

It should be noted, however, that there is relatively little reliable evidence of gender differences in cognition. Chapter 4 discussed and replicated gender-differences for one of the few cognitive domains that is known to show robust differences: face cognition (Hyde, 2016; Sommer et al., 2013). For many other cognitive functions, including verbal and mathematical skills, gender differences are not thought to be reliable (Hyde, 2016).

For brain development, too, there are only few clear sex-related differences, the main being that males have larger overall brain volumes than females (Giedd et al., 1999; Lenroot et al., 2007; Mills et al., 2016). Other differences reported by earlier studies (Giedd et al., 1999; Lenroot et al., 2007) have not replicated in recent, large-scale studies (Mills et al., 2016).

There are, however, well-established differences in the prevalence of different mental illnesses. Females have a higher prevalence of mood and anxiety disorders than males, for instance (Steel et al., 2014). It is therefore possible that the findings reported in Chapter 4 do not generalize to males: adolescent girls may be more susceptible to some of the negative effects of social stress (e.g. be more likely to ruminate or become anxious) than boys (see Hawes et al., 2012). We also cannot rule out that there may be gender differences in levels or developmental patterns of diligence (see Eskreis-Winkler, Shulman, Beal, & Duckworth, 2014) for some mixed evidence of gender differences in a diligence-related construct; Chapter 6). These possible gender differences will need to be investigated systematically by future studies with larger sample sizes.

7.3.3. The need for real-life measures

For both the cognitive training study (Chapter 5) and the imaging study (Chapter 6) we tried to collect academic grades to assess whether laboratory measures of training effects and diligence predicted real-life academic outcomes, but struggled with a number of difficulties.

First of all, UK school subjects and grading systems vary widely between schools, making it difficult to compare grades between participants. We attempted to address this issue for the study in Chapter 6 by collecting two waves of grades for the imaging study in order to predict how measures collected in the laboratory predict *changes* in grades rather than grades as such, which may be more reliable. However, only just over half of our participants responded to the second wave of data-collection, leaving us unable to analyse this data.

Future studies may therefore benefit from administering standardized tests of maths and reading ability instead of or alongside collecting school grades. Examples of such test include the Kaufman Test of Educational Attainment (Kaufman & Kaufman, 2004) or the Neale Analysis of Reading Ability test (Neale, 1999; see Dunning et al., 2013 for more examples). Research including such measures of educational attainment may help us assess the relevance of this thesis in the real world by assessing whether laboratory measures of cognitive performance (Chapter 4), reasoning or numeracy tasks (Chapter 5) and diligence tasks (Chapter 6) generalize to school performance.

Future research on social cognition would similarly benefit from investigating whether measures collected in the lab generalize to every-day behaviour. The study presented in Chapter 3, for instance, could be extended by assessing whether changes in face cognition skills during adolescence are related to how and with whom adolescents spend their time (Lam et al., 2014).

7.3.4. Modelling development

Our ability to gain a meaningful understanding of development from experimental studies is limited by a number of pragmatic design decisions: (I) which measures are used to assess development (e.g. age or puberty); (II) the range that is assessed for these measures (e.g. ages 11-33, or ages 14-15); and (III) how measures are modelled statistically (e.g. as a categorical or continuous variable).

The reader may refer to Chapter 2.4.2 for an extensive discussion of implications modelling age as a continuous or categorical variable (the latter was the case in this thesis). However, (I) and (II) warrant further consideration.

All studies in this thesis assessed age as a proxy for development. Age was chosen because it is a variable that can be assessed accurately, precisely and easily, while puberty is difficult to assess in self-report, and even using hormonal assays (Goddings et al, 2012). However, puberty may explain additional variance in development over and above age (Goddings, Mills, Clasen, Giedd, Viner & Blakemore, 2014) and is of particular relevance to sensitive periods. Theoretical work often posits that puberty, with its profound biological changes, triggers the onset of sensitive periods (Fawcett & Frankenhuis, 2015; Rodriguez de Fonseca et al., 1993). This suggestion remains untested at present, however, and will need to be followed up by future research.

Finally, the age ranges studied in this thesis were limited to adolescence and early adulthood. While other studies and reviews have provided excellent insights into earlier (Crone & Steinbeis, 2017; Röder, Ley, Shenoy, Kekunnaya, & Bottari, 2013; Sugita, 2008) and later (Baltes et al., 2006; Baltes, Staudinger, & Lindenberger,

1999; Thomas, 2008) development, we did not collect data on childhood or middle- and old age. This means we cannot compare development or learning directly between children, adolescents and adults and thus limits our ability draw conclusions about the on- and offset of sensitive periods of development.

Chapter 4, for instance, showed that 10 - 14 year olds were more affected by social stress than older age groups but we do not yet know how this age group compares with children. Similarly, Chapter 5 showed that adolescents and adults aged 15 - 33 benefited from cognitive training to a greater extent than younger age groups, but we do not yet know how older adults would respond to training. To be able to describe possible sensitive periods beyond childhood, future studies will need to investigate changes in plasticity and learning over the lifespan by including broader age ranges and elucidating whether the effects of the environment on human behaviour are quadratic (and thus consistent with a sensitive period, see Figure 1.5A in the Chapter 1) or linear (and therefore informative of age-related changes in plasticity but not consistent with sensitive periods, see Figure 1.5C).

7.4. Directions for Future Research

In addition to specific lines of inquiry proposed above to extend and improve the work presented in this thesis, this section explores three broader empirical, theoretical and methodological directions for future research. To better understand adolescent development, the following avenues of research could be insightful: (I) integrating research on adolescent development with insights from

lifespan psychology; (II) investigating mechanisms of learning other than plasticity; and (III) applying latent variable and multivariate models to large datasets.

7.4.1. Lifespan perspectives

This thesis has been concerned mainly with adolescent development. As alluded to in the previous section, future research could benefit from adopting a lifespan perspective and integrating the research presented here with research on earlier and later development.

The core tenants of lifespan psychology are that development is life-long, multi-dimensional, plastic up until old age, and that each developmental period has unique developmental tasks (Baltes et al., 2006; Baltes et al., 1999; Piaget, Grize, Szeminska, & Vinh, 1977; Tetens, 1777). Within this framework, childhood development is key for the acquisition of foundational skills like visual, perceptual and motor development (Inhelder & Piaget, 1958). Adolescence and early adulthood are characterized by a maturation of complex cognitive and social skills requiring high levels of cognitive control (Blakemore & Choudhury, 2006; Crone & Dahl, 2012). Later adulthood is a time of relative stability during which personality becomes consolidated, cognitive processes are optimized and both social and knowledge expertise continue to grow. Ageing, finally, is characterised by compensatory developments that mitigate physical and mental decline (Baltes et al., 2006; Baltes et al., 1999).

There are rich opportunities for future research in investigating lifespan changes in brain structure, cognitive functions and plasticity. Such studies may help us

understand how developmental phases affect one another. Some promising inroads have already been made, for instance, by linking the literature on automatic inhibitory control in childhood to the literature on deliberate inhibitory control development during adolescence (Crone & Steinbeis, 2017; Steinbeis & Crone, 2016). This has highlighted the possibility that some of the inconsistencies observed in studies on dual-systems and self-control in adolescence (see Chapter 6), may be due to different components of self-control developing at different rates. Late-maturing deliberate self-control tasks may be more apt to reveal developmental changes in adolescence than tasks relying on the automatic recruitment of self-control.

Another example of insightful lifespan work comes from Tamnes and colleagues, who carried out a set of studies linking structural brain development during adolescence to ageing. This research has shown that late-developing brain areas are more susceptible to atrophy during ageing (Douaud et al., 2014; Tamnes, Walhovd, Dale, et al., 2013; Walhovd et al., 2016). This highlights the relevance of adolescent development for health and well-being many decades later and more generally, shows that combining data and theory from different developmental phases creates insights into development that cannot be generated when each developmental phase is studied in isolation.

7.4.2. Learning beyond plasticity

The theoretical and empirical focus of this thesis has been understanding changes in plasticity as a mechanism for learning during adolescence. However, there

might be a role for mechanisms for learning, other than plasticity. For instance, the discussion of the cognitive training study in Chapter 5 highlighted that increased learning of relational reasoning in late adolescence and early adulthood could be due to strategy learning rather than plasticity (de Keysar & Larson-Hall, 2005; Goodwin & Johnson-Laird, 2005; Thomas, 2012).

Similarly, the peak in cognitive performance in adolescence and adulthood observed, for example, for working memory (Murre et al., 2013), episodic memory (Janssen, Chessa, & Murre, 2005; Janssen & Murre, 2008; Janssen et al., 2008), fluid intelligence (Li et al., 2004), numerical abilities (Halberda et al., 2012) or executive functions (Baum et al., 2017) may be not due to, or not only due to, increased plasticity, as proposed in Figure 1.5A in Chapter 1. Instead, a peak in cognitive performance and learning of complex cognitive skills may arise from an interaction between plasticity and other mechanisms of learning such as increasing declarative knowledge and a peak in optimization processes (e.g. devising and applying cognitive strategies, efficiently and flexibly allocating limited cognitive resources, integrating new and existing knowledge and skills) (see Figure 7.2).

In this theoretical model, plasticity is lower in adolescence than in childhood for many cognitive domains (Johnson, 2005; Thomas, 2012). However, learning of complex cognitive and social skills may still peak in adolescence and early adulthood because older age groups can access more declarative knowledge than children (Baltes et al., 2006; Li et al., 2004) and adolescents and young adults are better able to optimize cognitive resources and deploy strategies (see, for example, studies on optimal decision-making; Achterberg et al., 2016).

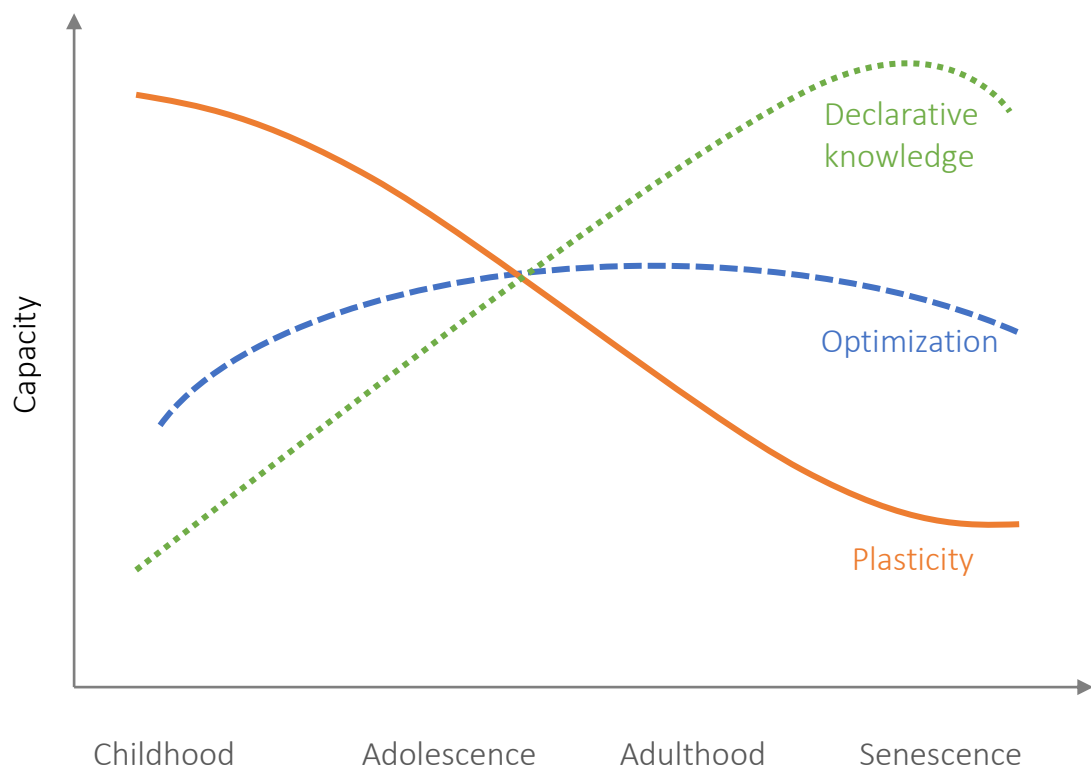


Figure 7.2. A Model of Lifespan Changes in Cognitive Performance and Learning. In this theoretical model learning is a product of plasticity (orange), optimization (blue) and declarative knowledge (green) (based on Baltes et al., 2006; Wenger & Lövdén, 2016).

This model predicts that there is not a single period during which learning is most efficient, but rather, that different types of learning may be more efficient at different ages. This prediction fits with the literature on second language learning and age differences in explicit and implicit learning (de Keysar & Larson-Hall, 2005; Thomas, 2012). Adults were often found to initially acquire a second language faster than children (de Keysar & Larson-Hall, 2005; Loewenthal & Bull, 1984). Explicit learning strategies and linguistic knowledge are thought to put them at an advantage. However, children eventually reach a level of proficiency often unattainable to adult learners because they perform better where learning involves complex and opaque linguistic material that is impervious to strategy use,

such as irregular pronunciation or grammar (Granena & Long, 2013; Hopp & Schmid, 2011). Importantly, children's implicit learning relies on massive, naturalistic exposure and is largely unreceptive to feedback or instruction (de Keysar & Larson-Hall, 2005), which limits the usefulness of training programs or interventions in childhood.

The model proposed in Figure 7.2 is speculative and remains to be tested. First of all, future research will need to establish whether plasticity, optimization and knowledge are separable components of learning ability. Optimization may, in fact, be best described in several sub-components rather than one unidimensional construct. Latent variable models may be useful here (see section 7.4.3). Then, future research will need to test when and how these components contribute to cognitive performance and learning and explore how they interact during development. As a third step, predictions from this model could be tested. We would expect, for instance, that training programs involving mass exposure are more effective early in development, while those recruiting strategy use may be more successful during adolescence. Interventions calling on extensive prior knowledge may be more effective later still, in adulthood.

7.4.3. Latent variable and multivariate models

This thesis, like most of cognitive neuroscience, employed univariate statistical tools (see Chapter 2). Such tools include univariate regression, ANOVA or GLMMs. However, cognitive neuroscience is an inherently multivariate discipline. We collect multiple dependent neuroimaging and behavioural variables to try to understand how brain and mind contribute to constructs such as perception,

language or executive function. Traditional multivariate approaches like MANOVA are limited in terms of their statistical power and ability to deal with incomplete data (McArdle, 2008). However, more powerful and flexible multivariate techniques like structural equation modelling (SEM) are now increasingly available open source and are rapidly gaining in popularity (Beaujean, 2014; Kievit et al., 2017; Rosseel, 2012). Future research may thus benefit from directly harnessing the multivariate nature of cognitive neuroscience data and may soon be able to address the relationship between brain structure, function and behaviour more explicitly (Kievit et al., 2011).

SEM is a combination of regression and factor analysis (Kievit et al., 2017). The relationship between multiple variables can be modelled as regression-like paths (e.g. frontal grey matter volume predicts cognitive performance, which predicts school grades). Using elements of factor analysis, SEM can also model how theoretical constructs (i.e. latent variables, e.g. cognitive performance) can be inferred from multiple observable variables (i.e. manifest variables, e.g. accuracy and reaction time on a working memory task) (Beaujean, 2014; Little, 2013; Newsom, 2015).

A major caveat of SEM is that it requires large sample sizes (Kievit et al., 2017; Little, 2013). It would not be possible to model data from smaller neuroimaging studies like the one presented in Chapter 6 with SEM, for instance. Nonetheless, it is hoped, that, as funding bodies begin to fund larger samples for neuroimaging studies and data sharing becomes more popular (Spires-Jones, Poirazi, & Grubb, 2016), SEMs can be utilized more.

Models like SEMs could help extend the work presented in Chapter 6 and provide a better understanding as to how brain structure, function and connectivity of different brain regions contribute to constructs such as self-control (Crone & Steinbeis, 2017; Huizinga, Dolan, & van der Molen, 2006). Longitudinal SEMs can also help us understand how different cognitive functions influence one another over development (Grimm et al., 2017; Little, 2013) or allow us to model individual differences in response to cognitive training or stress.

7.5. Conclusion

This thesis investigated learning and plasticity in adolescence. It highlights that many cognitive and socio-cognitive functions: (I) undergo protracted changes during adolescence; (II) are characterized by individual differences; (III) are susceptible to the advantageous effects of cognitive training; and (IV) are susceptible to the adverse effects of social stress. This underlines that adolescence is a key period of life for learning complex cognitive skills.

8. References

- Achterberg, M., Peper, J. S., van Duijvenvoorde, A. C., Mandl, R. C., & Crone, E. A. (2016). Frontostriatal white matter integrity predicts development of delay of gratification: A longitudinal study. *Journal of Neuroscience*, *36*(6), 1954-1961. doi:10.1523/jneurosci.3459-15.2016
- Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41-74. doi:10.1037/0033-2909.101.1.41
- Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, *4*(3), 165-178.
- Ahmed, S. P., Bittencourt-Hewitt, A., & Sebastian, C. L. (2015). Neurocognitive bases of emotion regulation development in adolescence. *Developmental Cognitive Neuroscience*, *15*, 11-25. doi:10.1016/j.dcn.2015.07.006
- Allen, D., & Smith, H. I. D. (2009). *Early intervention: Good parents, great kids, better citizens*. Retrieved from <http://www.centreforsocialjustice.org.uk/UserStorage/pdf/Pdf%20reports/EarlyInterventionFirstEdition.pdf>
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*(2), 606-621. doi:10.1111/j.1467-8624.2009.01282.x
- Andersen, S. L., & Teicher, M. H. (2008). Stress, sensitive periods and maturational events in adolescent depression. *Trends in Neurosciences*, *31*(4), 183-191. doi:10.1016/j.tins.2008.01.004

- Andersson, J., & Hutton, C. (2017). FieldMap toolbox. University College London, London, UK. Retrieved from <http://www.fil.ion.ucl.ac.uk/spm/toolbox/fieldmap/>
- Arseneault, L., Bowes, L., & Shakoor, S. (2009). Bullying victimization in youths and mental health problems: 'Much ado about nothing'? *Psychological Medicine, 40*(5), 717-729. doi:10.1017/S0033291709991383
- Arseneault, L., Walsh, E., Trzesniewski, K., Newcombe, R., Caspi, A., & Moffitt, T. E. (2006). Bullying victimization uniquely contributes to adjustment problems in young children: A nationally representative cohort study. *Pediatrics, 118*(1), 130-138. doi:10.1542/peds.2005-2388
- Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829-839. doi:10.1038/nrn1201
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.). *New York, NY: Academic Press.*
- Baltes, P. B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development, 11*(3), 145-171.
- Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (2006). Life span theory in developmental psychology. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology* (6th ed.). Hoboken, NJ: John Wiley & Sons.
- Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology, 50*(1), 471-507.
- Barnett, W. S. (2011). Effectiveness of Early Educational Intervention. *Science, 333*(6045), 975.

- Bates, D. J., Maechler, M., & Bolker, B. (2013). lme4: Linear-mixed effects models using S4 classes. Retrieved from <http://CRAN.R-project.org/package=lme4>
- Battistella, G., Fornari, E., Annoni, J. M., Chtioui, H., Dao, K., Fabritius, M., . . . Giroud, C. (2014). Long-term effects of cannabis on brain structure. *Neuropsychopharmacology*, *39*(9), 2041-2048. doi:10.1038/npp.2014.67
- Baum, G. L., Ciric, R., Roalf, D. R., Betzel, R. F., Moore, T. M., Shinohara, R. T., . . . Satterthwaite, T. D. (2017). Modular segregation of structural brain networks supports the development of executive function in youth. *Current Biology*. doi:10.1016/j.cub.2017.04.051
- Baumeister, R. F., Twenge, J. M., & Nuss, C. K. (2002). Effects of social exclusion on cognitive processes: Anticipated aloneness reduces intelligent thought. *Journal of Personality and Social Psychology*, *83*(4), 817-827.
- Bavelier, D., Levi, D. M., Li, R. W., Dan, Y., & Hensch, T. K. (2010). Removing brakes on adult brain plasticity: From molecular to behavioral interventions. *Journal of Neuroscience*, *30*(45), 14964-14971. doi:10.1523/jneurosci.4812-10.2010
- Bazargani, N., Hillebrandt, H., Christoff, K., & Dumontheil, I. (2014). Developmental changes in effective connectivity associated with relational reasoning. *Human Brain Mapping*, *35*(7), 3262-3276. doi:10.1002/hbm.22400
- Beaujean, A. (2014). *Latent variable modeling using R. A step-by-step guide*. New York, NY: Routledge.
- Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: Face-blind from birth. *Trends in Cognitive Sciences*, *9*(4), 180-187. doi:10.1016/j.tics.2005.02.011

- Berghuis, P., Rajnicek, A. M., Morozov, Y. M., Ross, R. A., Mulder, J., Urbán, G. M., . . . Harkany, T. (2007). Hardwiring the brain: Endocannabinoids shape neuronal connectivity. *Science*, *316*(5828), 1212-1216.
doi:10.1126/science.1137406
- Berrendero, F., Sepe, N., Ramos, J. A., Di Marzo, V., & Fernandez-Ruiz, J. J. (1999). Analysis of cannabinoid receptor binding and mRNA expression and endogenous cannabinoid contents in the developing rat brain during late gestation and early postnatal period. *Synapse*, *33*(3), 181-191.
doi:10.1002/(sici)1098-2396(19990901)33:3<181::aid-syn3>3.0.co;2-r
- Bieszczad, K. M., Bechay, K., Rusche, J. R., Jacques, V., Kudugunti, S., Miao, W., . . . Wood, M. A. (2015). Histone deacetylase inhibition via RGFP966 releases the brakes on sensory cortical plasticity and the specificity of memory formation. *Journal of Neuroscience*, *35*(38), 13124-13132.
doi:10.1523/jneurosci.0914-15.2015
- Bjork, J. M., Smith, A. R., Chen, G., & Hommer, D. W. (2010). Adolescents, adults and rewards: Comparing motivational neurocircuitry recruitment using fMRI. *PLoS One*, *5*(7), e11440. doi:10.1371/journal.pone.0011440
- Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, *9*(4), 267-277.
- Blakemore, S. J. (2010). The developing social brain: Implications for education. *Neuron*, *65*(6), 744-747. doi:10.1016/j.neuron.2010.03.004
- Blakemore, S. J. (2012). Development of the social brain in adolescence. *Journal of the Royal Society of Medicine*, *105*(3), 111-116.
doi:10.1258/jrsm.2011.110221

- Blakemore, S. J. (2013). Teenage kicks: Cannabis and the adolescent brain. *Lancet*, 381(9870), 888-889. doi:10.1016/s0140-6736(12)61578-5
- Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, 47(3-4), 296-312. doi:10.1111/j.1469-7610.2006.01611.x
- Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, 65, 187-207. doi:10.1146/annurev-psych-010213-115202
- Blakemore, S. J., & Robbins, T. W. (2012). Decision-making in the adolescent brain. *Nature Neuroscience*, 15(9), 1184-1191. doi:10.1038/nn.3177
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized Linear Mixed Models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24(3), 127-135.
- Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: the role of featural and configural face information. *Quarterly Journal of Experimental Psychology*, 66(12), 2426-2442. doi:10.1080/17470218.2013.789065
- Bourgeois, J. P., & Rakic, P. (1993). Changes of synaptic density in the primary visual cortex of the macaque monkey from fetal to adult stage. *Journal of Neuroscience*, 13(7), 2801-2820.
- 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(18), 7226.

Brain Development Cooperative Group (2012). Total and regional brain volumes in a population-based normative sample from 4 to 18 years: the NIH MRI Study of Normal Brain Development. *Cerebral Cortex*, 22(1), 1-12.
doi:10.1093/cercor/bhr018

Branstetter, S. A., Low, S., & Furman, W. (2011). The influence of parents and friends on adolescent substance use: A multidimensional approach. *Journal of Substance Use*, 16(2), 150-160.
doi:10.3109/14659891.2010.519421

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, JP.
<http://marsbar.sourceforge.net/>

Brown, T. T. (2017). Individual differences in human brain development. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1-2), e1389-n/a.
doi:10.1002/wcs.1389

Brown, T. T., Kuperman, Joshua M., Chung, Y., Erhart, M., McCabe, C., Hagler, Donald J., . . . Dale, Anders M. (2012). Neuroanatomical assessment of biological maturity. *Current Biology*, 22(18), 1693-1698.
doi:10.1016/j.cub.2012.07.002

Buckley, K. E., Winkel, R. E., & Leary, M. R. (2004). Emotional and behavioral responses to interpersonal rejection: Anger, sadness, hurt, and aggression. *Journal of Experimental Social Psychology*, 40(1), 14-28.

- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends in Cognitive Sciences*, *10*(4), 159-166. doi:10.1016/j.tics.2006.02.004
- Burke, A. R., McCormick, C. M., Pellis, S. M., & Lukkes, J. L. Impact of adolescent social experiences on behavior and neural circuits implicated in mental illnesses. *Neuroscience & Biobehavioral Reviews*. doi:10.1016/j.neubiorev.2017.01.018
- Burnett, S., Bird, G., Moll, J., Frith, C., & Blakemore, S. J. (2009). Development during adolescence of the neural processing of social emotion. *Journal of Cognitive Neuroscience*, *21*(9), 1736-1750. doi:10.1162/jocn.2009.21121
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365-376. doi:10.1038/nrn3475
- Buwalda, B., Geerdink, M., Vidal, J., & Koolhaas, J. M. (2011). Social behavior and social stress in adolescence: A focus on animal models. *Neuroscience and Biobehavioral Reviews*, *35*(8), 1713-1721. doi:10.1016/j.neubiorev.2010.10.004
- Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biology*, *4*(5), e125. doi:10.1371/journal.pbio.0040125
- Cappelletti, M., Pikkat, H., Upstill, E., Speekenbrink, M., & Walsh, V. (2015). Learning to integrate versus inhibiting information is modulated by age. *Journal of Neuroscience*, *35*(5), 2213.

- Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition: A maturational component? *Developmental Psychology, 16*(4), 257-269.
doi:10.1037/0012-1649.16.4.257
- Carroll, S. P., & Corneli, P. S. (1995). Divergence in male mating tactics between two populations of the soapberry bug: II. Genetic change and the evolution of a plastic reaction norm in a variable social environment. *Behavioural Ecology, 6*. doi:10.1093/beheco/6.1.46
- Casey, B. J. (2013). The teenage brain: An overview. *Current Directions in Psychological Science, 22*(2), 80-81. doi:10.1177/0963721413486971
- Casey, B. J., Galván, A., & Somerville, L. H. (2016). Beyond simple models of adolescence to an integrated circuit-based account: A commentary. *Developmental Cognitive Neuroscience, 17*, 128-130.
doi:10.1016/j.dcn.2015.12.006
- Casey, B. J., Getz, S., & Galván, A. (2008). The adolescent brain. *Developmental Review, 28*(1), 62-77. doi:10.1016/j.dr.2007.08.003
- Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage, 14*(5), 1136-1149.
doi:10.1006/nimg.2001.0922
- Chuderski, A. (2014). The relational integration task explains fluid reasoning above and beyond other working memory tasks. *Memory & Cognition, 42*(3), 448-463. doi:10.3758/s13421-013-0366-x
- Coffman, M. C., Trubanova, A., Richey, J. A., White, S. W., Kim-Spoon, J., Ollendick, T. H., & Pine, D. S. (2015). Validation of the NIMH-ChEFS adolescent face stimulus set in an adolescent, parent, and health professional sample.

International Journal of Methods in Psychiatric Research, 24(4), 275-286.

doi:10.1002/mpr.1490

Cogent 2000 Team (2015). Cogent 2000. London, UK: Functional Neuroimaging

Laboratory, University College London. Retrieved from

http://www.vislab.ucl.ac.uk/cogent_2000_terms_and_conditions.php

Cohen Kadosh, K. (2011). Differing processing abilities for specific face properties in mid-childhood and adulthood. *Frontiers of Psychology*, 2, 400.

doi:10.3389/fpsyg.2011.00400

Cohen Kadosh, K., Johnson, M. H., Dick, F., Cohen Kadosh, R., & Blakemore, S. J.

(2013). Effects of age, task performance, and structural brain development on face processing. *Cerebral Cortex*, 23(7), 1630-1642.

doi:10.1093/cercor/bhs150

Cohen Kadosh, K., Johnson, M. H., Henson, R. N., Dick, F., & Blakemore, S. J.

(2013). Differential face-network adaptation in children, adolescents and adults. *NeuroImage*, 69, 11-20. doi:10.1016/j.neuroimage.2012.11.060

Cohen Kadosh, K., Walsh, V., & Cohen Kadosh, R. (2010). Investigating face-

property specific processing in the right OFA. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsq015

Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory

performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology*, 31(1), 103-128. doi:10.1207/s15326942dn3101_6

Conway, M. A., Wang, Q., Hanyu, K., & Haque, S. (2005). A cross-cultural

investigation of autobiographical memory: On the universality and cultural

investigation of autobiographical memory: On the universality and cultural

- variation of the reminiscence bump. *Journal of Cross-Cultural Psychology*, 36(6), 739-749. doi:10.1177/0022022105280512
- Cook, J., & Bird, G. (2011). Social attitudes differentially modulate imitation in adolescents and adults. *Experimental Brain Research*, 211(3-4), 601-612. doi:10.1007/s00221-011-2584-4
- Credé, M., Tynan, M. C., & Harms, P. D. (2016). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, 113(3), 492-511. doi:10.1037/pspp0000102
- Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, 13(9), 636-650. doi:10.1038/nrn3313
- Crone, E. A., & Steinbeis, N. (2017). Neural perspectives on cognitive control development during childhood and adolescence. *Trends in Cognitive Sciences*, 21(3), 205-215. doi:10.1016/j.tics.2017.01.003
- Crone, E. A., Wendelken, C., van Leijenhorst, L., Honomichl, R. D., Christoff, K., & Bunge, S. A. (2009). Neurocognitive development of relational reasoning. *Developmental Science*, 12(1), 55-66. doi:10.1111/j.1467-7687.2008.00743.x
- Crookes, K., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition*, 111(2), 219-247. doi:10.1016/j.cognition.2009.02.004
- Curci, A., Lanciano, T., Soletti, E., & Rimé, B. (2013). Negative emotional experiences arouse rumination and affect working memory capacity. *Emotion*, 13(5), 867-880. doi:10.1037/a0032492

- Dahnke, R., & Gaser, C. (2016). Computational Anatomy Toolbox - CAT12.
University of Jena: Structural Brain Mapping Group. Retrieved from
<http://www.neuro.uni-jena.de/vbm/download/>
- Damon, W. (2004). In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed.). Hoboken, NJ: John Wiley & Sons.
- de Haan, M. (2014). Attention and executive control. In D. Mareschal, B. Butterworth, & A. Tolmie (Eds.), *Educational Neuroscience*. Chichester, UK: John Wiley & Sons.
- de Keysar, R., & Larson-Hall, J. (2005). What does the critical period really mean? In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 88-108). Oxford, UK: Oxford University Press.
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*(3), 487-506.
doi:10.1080/02643290244000239
- Dekel, E., & Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. *Nature*, *436*(7050), 588-592. doi:10.1038/nature03842
- Deshpande, A., Desrochers, A., Ksoll, C., & Shonchoy, A. S. (2017). The impact of a computer-based adult literacy program on literacy and numeracy: Evidence from India. *World Development*, *96*, 451-473.
doi:10.1016/j.worlddev.2017.03.029
- DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: Effects of feedback and training. *Frontiers in Human Neuroscience*, *6*, 68. doi:10.3389/fnhum.2012.00068

- Dews, P. B., & Wiesel, T. N. (1970). Consequences of monocular deprivation on visual behaviour in kittens. *The Journal of Physiology*, *206*(2), 437-455.
- DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience: Brain structure and the Big Five. *Psychological Science*, *21*(6), 820-828.
doi:10.1177/0956797610370159
- Diamond, R., Carey, S., & Back, K. J. (1983). Genetic influences on the development of spatial skills during early adolescence. *Cognition*, *13*(2), 167-185.
- Dobbing, J., & Sands, J. (1973). Quantitative growth and development of human brain. *Archives of Disease in Childhood*, *48*(10), 757-767.
- Dolzicka, D., Herzmann, G., Sommer, W., & Wilhelm, O. (2014). Can training enhance face cognition abilities in middle-aged adults? *PLoS One*, *9*(3), e90249.
- Donders, A. R. T., van der Heijden, G. J. M. G., Stijnen, T., & Moons, K. G. M. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, *59*(10), 1087-1091.
doi:10.1016/j.jclinepi.2006.01.014
- Douaud, G., Groves, A. R., Tamnes, C. K., Westlye, L. T., Duff, E. P., Engvig, A., . . . Johansen-Berg, H. (2014). A common brain network links development, aging, and vulnerability to disease. *Proceedings of the National Academy of Sciences*, *111*(49), 17648-17653. doi:10.1073/pnas.1410378111
- Dubow, E. F., Boxer, P., & Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: Mediation

- by family interactions, child aggression, and teenage aspirations. *Merrill-Palmer Quarterly*, 55(3), 224-249. doi:10.1353/mpq.0.0030
- Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, 44(4), 576-585. doi:10.1016/j.neuropsychologia.2005.07.001
- Duckworth, A. L., & Gross, J. J. (2014). Self-control and grit. *Current Directions in Psychological Science*, 23(5), 319-325. doi:10.1177/0963721414541462
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087-1101. doi:10.1037/0022-3514.92.6.1087
- Duckworth, A. L., & Steinberg, L. (2015). Unpacking self-control. *Child Development Perspectives*, 9(1), 32-37. doi:10.1111/cdep.12107
- Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, 13(2), 331-338. doi:10.1111/j.1467-7687.2009.00888.x
- Dumontheil, I., Houlton, R., Christoff, K., & Blakemore, S. J. (2010). Development of relational reasoning during adolescence. *Developmental Science*, 13(6), F15-24. doi:10.1111/j.1467-7687.2010.01014.x
- Dumontheil, I., Kuster, O., Apperly, I. A., & Blakemore, S. J. (2010). Taking perspective into account in a communicative task. *NeuroImage*, 52(4), 1574-1583. doi:10.1016/j.neuroimage.2010.05.056
- Dunning, D. L., Holmes, J., & Gathercole, S. E. (2013). Does working memory training lead to generalized improvements in children with low working

memory? A randomized controlled trial. *Developmental Science*, 16(6), 915-925. doi:10.1111/desc.12068

Eaton, D. K., Kann, L., Kinchen, S., Shanklin, S., Flint, K. H., Hawkins, J., . . . Chyen, D. (2012). Youth risk behavior surveillance - United States, 2011. *Morbidity and Mortality Weekly Report. Surveillance Summaries*, 61(4), 1-162.

Ebbinghaus, H. (1885). *Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie*. Leipzig, DE: Duncker & Humblot.

Ehrenreich, H., Rinn, T., Kunert, H. J., Moeller, M. R., Poser, W., Schilling, L., . . . Hoehe, M. R. (1999). Specific attentional dysfunction in adults following early start of cannabis use. *Psychopharmacology*, 142(3), 295-301.

Einon, D. F., & Morgan, M. J. (1977). A critical period for social isolation in the rat. *Developmental Psychobiology*, 10(2), 123-132. doi:10.1002/dev.420100205

Ellgren, M., Spano, S. M., & Hurd, Y. L. (2007). Adolescent cannabis exposure alters opiate intake and opioid limbic neuronal populations in adult rats. *Neuropsychopharmacology*, 32(3), 607-615. doi:10.1038/sj.npp.1301127

Eluvathingal, T. J., Chugani, H. T., Behen, M. E., Juhasz, C., Muzik, O., Maqbool, M., . . . Makki, M. (2006). Abnormal brain connectivity in children after early severe socioemotional deprivation: a diffusion tensor imaging study. *Pediatrics*, 117(6), 2093-2100. doi:10.1542/peds.2005-1727

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology*, 128(3), 309-331.

Ernst, M., Nelson, E. E., Jazbec, S., McClure, E. B., Monk, C. S., Leibenluft, E., . . . Pine, D. S. (2005). Amygdala and nucleus accumbens in responses to

- receipt and omission of gains in adults and adolescents. *NeuroImage*, 25(4), 1279-1291. doi:10.1016/j.neuroimage.2004.12.038
- Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers of Psychology*, 5, 36. doi:10.3389/fpsyg.2014.00036
- European Monitoring Centre for Drugs and Drug Addiction (2011). *Annual report 2011: The state of the drugs problem in Europe - Cannabis*. Retrieved from <http://www.emcdda.europa.eu/online/annual-report/2011/cannabis/3>
- Fawcett, T. W., & Frankenhuis, W. E. (2015). Adaptive explanations for sensitive windows in development. *Frontiers in Zoology*, 12(1), S3. doi:10.1186/1742-9994-12-S1-S3
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science*, 8(7), 307-314. doi:10.1016/j.tics.2004.05.002
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. East Norwalk, CT: Appleton-Century-Crofts.
- Fields, R. D. (2014). Myelin - More than insulation. *Science*, 344(6181), 264-266. doi:10.1126/science.1253851
- Fitzsimons, G. J. (2008). Editorial: A death to dichotomizing. *Journal of Consumer Research*, 35(1), 5-8. doi:10.1086/589561
- Forbes, C. E., Poore, J. C., Krueger, F., Barbey, A. K., Solomon, J., & Grafman, J. (2014). The role of executive function and the dorsolateral prefrontal cortex in the expression of neuroticism and conscientiousness. *Social Neuroscience*, 9(2), 139-151. doi:10.1080/17470919.2013.871333

- Forero, R., McLellan, L., Rissel, C., & Bauman, A. (1999). Bullying behaviour and psychosocial health among school students in New South Wales, Australia: Cross sectional survey. *British Medical Journal*, *319*(7206), 344.
- Foulkes, L., & Blakemore, S. J. (2016). Is there heightened sensitivity to social reward in adolescence? *Current Opinion in Neurobiology*, *40*, 81-85.
doi:10.1016/j.conb.2016.06.016
- Fuhrmann, D., Knoll, L. J., & Blakemore, S. J. (2015). Adolescence as a sensitive period of brain development. *Trends in Cognitive Sciences*.
doi:10.1016/j.tics.2015.07.008
- Fuhrmann, D., Knoll, L. J., Sakhardande, A. L., Speekenbrink, M., Kadosh, K. C., & Blakemore, S. J. (2016). Perception and recognition of faces in adolescence. *Scientific Reports*, *6*, 33497. doi:10.1038/srep33497
- Furman, W., & Buhrmester, D. (1992). Age and sex differences in perceptions of networks of personal relationships. *Child Development*, *63*(1), 103-115.
- Galbraith, S., Bowden, J., & Mander, A. (2017). Accelerated longitudinal designs: An overview of modelling, power, costs and handling missing data. *Stat Methods in Medical Research*, *26*(1), 374-398.
doi:10.1177/0962280214547150
- Galla, B. M., Plummer, B. D., White, R. E., Meketon, D., D'Mello, S. K., & Duckworth, A. L. (2014). The Academic Diligence Task (ADT): Assessing individual differences in effort on tedious but important schoolwork. *Contemporary Educational Psychology*, *39*(4), 314-325.
doi:10.1016/j.cedpsych.2014.08.001
- Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental

study. *Developmental Psychology*, 41(4), 625-635. doi:10.1037/0012-1649.41.4.625

- Gathercole, S. E., Brown, L., & Pickering, S. J. (2003). Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. *Educational and Child Psychology*, 20(3), 109-122.
- Geier, C. F., Terwilliger, R., Teslovich, T., Velanova, K., & Luna, B. (2010). Immaturities in reward processing and its influence on inhibitory control in adolescence. *Cerebral Cortex*, 20(7), 1613-1629. doi:10.1093/cercor/bhp225
- Germine, L. T., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition*, 118(2), 201-210. doi:10.1016/j.cognition.2010.11.002
- Gervain, J., Vines, B., Chen, L., Seo, R., Hensch, T., Werker, J., & Young, A. (2013). Valproate reopens critical-period learning of absolute pitch. *Frontiers in Systems Neuroscience*, 7, 102.
- Gevins, A., & Cuttillo, B. (1993). Spatiotemporal dynamics of component processes in human working memory. *Electroencephalography and Clinical Neurophysiology*, 87(3), 128-143. doi:10.1016/0013-4694(93)90119-G
- Gianotti, L. R., Knoch, D., Faber, P. L., Lehmann, D., Pascual-Marqui, R. D., Diezi, C., . . . Fehr, E. (2009). Tonic activity level in the right prefrontal cortex predicts individuals' risk taking. *Psychological Science*, 20(1), 33-38. doi:10.1111/j.1467-9280.2008.02260.x
- 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(10), 861-863. doi:10.1038/13158

GIMP Team (2013). GNU image manipulation program (Version 2.8.8). Retrieved from <http://www.gimp.org>

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(1), 94-103.

doi:10.1016/j.neuroimage.2009.08.003

Goddings, A. L., Burnett Heyes, S., Bird, G., Viner, R. M., & Blakemore, S. J. (2012).

The relationship between puberty and social emotion processing.

Developmental Science, 15(6), 801-811. doi:10.1111/j.1467-

7687.2012.01174.x

Goddings, A.-L., Mills, K. L., Clasen, L. S., Giedd, J. N., Viner, R. M., & Blakemore, S.-

J. (2014). The influence of puberty on subcortical brain development.

NeuroImage, 88, 242–251. doi: 10.1016/j.neuroimage.2013.09.073

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, 101(21), 8174-8179.

doi:10.1073/pnas.0402680101

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal

models. *Child Development*, 58(1), 80-92.

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations.

Psychological Review, 112(2), 468-493. doi:10.1037/0033-295x.112.2.468

- Goswami, U. (2006). Neuroscience and education: From research to practice?
Nature Reviews Neuroscience, 7(5), 406-413.
- Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311-343.
- Greenough, W. T., Black, J. E., & Wallace, C. S. (1987). Experience and brain development. *Child Development*, 58(3), 539-559. doi:10.2307/1130197
- Grimm, K. J., Davoudzadeh, P., & Ram, N. (2017). Developments in the analysis of longitudinal data. *Monographs of the Society for Research in Child Development*, 82(2), 46-66.
- Grydeland, H., Walhovd, K. B., Tamnes, C. K., Westlye, L. T., & Fjell, A. M. (2013). Intracortical myelin links with performance variability across the human lifespan: Results from T1- and T2-weighted MRI myelin mapping and diffusion tensor imaging. *The Journal of Neuroscience*, 33(47), 18618.
- Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., . . . Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. *Neuropsychology*, 26(2), 251-265. doi:10.1037/a0026712
- Haber, S. N. (2016). Corticostriatal circuitry. *Dialogues in Clinical Neuroscience*, 18(1), 7-21.
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44(5), 1457-1465. doi:10.1037/a0012682

- Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences, 109*(28), 11116-11120. doi:10.1073/pnas.1200196109
- Halberda, J., Mazocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*(7213), 665-668. doi:10.1038/nature07246
- Hall, G. S. (1904). *Adolescence: Its Psychology and its relations to Physiology, Anthropology, Sociology, sex, crime, religion and education*. New York, NY: D. Appleton and Company.
- Harkany, T., Keimpema, E., Barabás, K., & Mulder, J. (2008). Endocannabinoid functions controlling neuronal specification during brain development. *Molecular and Cellular Endocrinology, 286*(1-2, Supplement 1), S84-S90. doi:10.1016/j.mce.2008.02.011
- Hartgerink, C. H. J., van Beest, I., Wicherts, J. M., & Williams, K. D. (2015). The ordinal effects of ostracism: A meta-analysis of 120 Cyberball studies. *PLoS One, 10*(5), e0127002. doi:10.1371/journal.pone.0127002
- Hastie, T., & Tibshirani, R. (2006). Generalized additive models. *Encyclopedia of Statistical Sciences*.
- Hawes, D. J., Zadro, L., Fink, E., Richardson, R., O'Moore, K., Griffiths, B., . . . Williams, K. D. (2012). The effects of peer ostracism on children's cognitive processes. *European Journal of Developmental Psychology, 9*(5), 599-613. doi:10.1080/17405629.2011.638815
- Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics, 54*(1), 3-56. doi:10.1006/reec.1999.0225

- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900-1902.
- Heisz, J. J., Pottruff, M. M., & Shore, D. I. (2013). Females scan more than males: A potential mechanism for sex differences in recognition memory. *Psychological Science*, *24*(7), 1157-1163. doi:10.1177/0956797612468281
- Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. *Nature Reviews Neuroscience*, *6*(11), 877-888.
- Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: a meta-analytic review. *Visual Cognition*, *21*(9-10), 1306-1336.
- Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174-180.
doi:10.1016/j.tics.2012.01.006
- Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26*(3), 339-366.
- Hopp, H., & Schmid, M. S. (2011). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, *34*(2), 361-394.
doi:10.1017/S0142716411000737
- Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., & Scheibe, S. (2016). Multcomp: Simultaneous inference in general parametric models. Retrieved from <https://cran.r-project.org/web/packages/multcomp/index.html>

- Howard-Jones, P. A., Washbrook, E. V., & Meadows, S. (2012). The timing of educational investment: A neuroscientific perspective. *Developmental Cognitive Neuroscience, 2, Supplement 1*, S18-S29.
doi:10.1016/j.dcn.2011.11.002
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160*(1), 106-154.
- Hubel, D. H., & Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *The Journal of Physiology, 206*(2), 419-436.
- Hubener, M., & Bonhoeffer, T. (2014). Neuronal plasticity: Beyond the critical period. *Cell, 159*(4), 727-737. doi:10.1016/j.cell.2014.10.035
- 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*(11), 2017-2036.
doi:10.1016/j.neuropsychologia.2006.01.010
- Hyde, J. S. (2016). Sex and cognition: Gender and cognitive functions. *Current Opinion in Neurobiology, 38*, 53-56. doi:10.1016/j.conb.2016.02.007
- Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015a). The median split: Robust, refined, and revived. *Journal of Consumer Psychology, 25*(4), 690-704. doi:10.1016/j.jcps.2015.06.014
- Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015b). Toward a more nuanced understanding of the statistical properties of a median split. *Journal of Consumer Psychology, 25*(4), 652-665. doi:10.1016/j.jcps.2014.12.002

- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. (Vol. 84). New York, NY: Basic Books.
- Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, *40*(3), 366-371.
- Isbell, E., Fukuda, K., Neville, H. J., & Vogel, E. K. (2015). Visual working memory continues to develop through adolescence. *Frontiers of Psychology*, *6*, 696. doi:10.3389/fpsyg.2015.00696
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081-10086. doi:10.1073/pnas.1103228108
- Jaeggi, S. M., Buschkuhl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition*, *42*(3), 464-480. doi:10.3758/s13421-013-0364-z
- Jagdhuber, R. (2016). *Multiple imputation in Generalized Linear Mixed Models*. Ludwig Maximilian University, Munich, DE.
- Jamieson, J. P., Harkins, S. G., & Williams, K. D. (2010). Need threat can motivate performance after ostracism. *Personality & Social Psychology Bulletin*, *36*(5), 690-702. doi:10.1177/0146167209358882
- Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (2005). The reminiscence bump in autobiographical memory: Effects of age, gender, education, and culture. *Memory*, *13*(6), 658-668. doi:10.1080/09658210444000322
- Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (2007). Temporal distribution of favourite books, movies, and records: Differential encoding and re-sampling. *Memory*, *15*(7), 755-767. doi:10.1080/09658210701539646

- Janssen, S. M. J., & Murre, J. M. (2008). Reminiscence bump in autobiographical memory: Unexplained by novelty, emotionality, valence, or importance of personal events. *Quarterly Journal of Experimental Psychology*, *61*(12), 1847-1860. doi:10.1080/17470210701774242
- Janssen, S. M. J., Murre, J. M. J., & Meeter, M. (2008). Reminiscence bump in memory for public events. *European Journal of Cognitive Psychology*, *20*(4), 738-764. doi:10.1080/09541440701554409
- Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*(1), 1-123.
doi:10.17763/haer.39.1.l3u15956627424k7
- Johnson, M. H. (2001). Functional brain development in humans. *Nature Review Neuroscience*, *2*(7), 475-483.
- Johnson, M. H. (2005). Sensitive periods in functional brain development: Problems and prospects. *Developmental Psychobiology*, *46*(3), 287-292.
doi:10.1002/dev.20057
- Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2013). *Monitoring the future. National survey results on drug use 1975-2012*. Retrieved from Ann Arbor, MI:
http://www.monitoringthefuture.org/pubs/monographs/vol1_2006.pdf
- Judd, C. M., McClelland, G. H., & Ryan, C. S. (2009). *Data analysis: A model comparison approach*. New York, NY: Routledge.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal*

of *Experimental Psychology*, 133(2), 189-217. doi:10.1037/0096-3445.133.2.189

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Test of Educational Attainment*. Minneapolis, MN: Pearson Assessments.

Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Ustun, T. B. (2007). Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry*, 20(4), 359-364. doi:10.1097/YCO.0b013e32816ebc8c

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 593-602. doi:10.1001/archpsyc.62.6.593

Kievit, R. A., Brandmaier, A., Ziegler, G., van Harmelen, A. L., de Mooij, S., Moutoussis, M., . . . Dolan, R. (2017). Developmental cognitive neuroscience using Latent Change Score models: A tutorial and applications. *bioRxiv*.

Kievit, R. A., Romeijn, J. W., Waldorp, L. J., Wicherts, J. M., Scholte, H. S., & Borsboom, D. (2011). Mind the gap: A psychometric approach to the reduction problem. *Psychological Inquiry*, 22(2), 67-87.

Kilford, E. J., Garrett, E., & Blakemore, S. J. (2016). The development of social cognition in adolescence: An integrated perspective. *Neuroscience & Biobehavioral Reviews*, 70, 106-120. doi:10.1016/j.neubiorev.2016.08.016

Kim-Cohen, J., Caspi, A., Moffitt, T. E., Harrington, H. L., Milne, B. J., & Poulton, R. (2003). Prior juvenile diagnoses in adults with mental disorder:

- Developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry*, 60(7), 709-717. doi:10.1001/archpsyc.60.7.709
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Science*, 14(7), 317-324. doi:10.1016/j.tics.2010.05.002
- Klomek, A. B., Sourander, A., Niemela, S., Kumpulainen, K., Piha, J., Tamminen, T., . . . Gould, M. S. (2009). Childhood bullying behaviors as a risk for suicide attempts and completed suicides: A population-based birth cohort study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(3), 254-261. doi:10.1097/CHI.0b013e318196b91f
- Knoch, D., Gianotti, L. R. R., Baumgartner, T., & Fehr, E. (2010). A neural marker of costly punishment behavior. *Psychological Science*, 21(3), 337-342. doi:10.1177/0956797609360750
- Knoll, L. J., Fuhrmann, D., Sakhardande, A. L., Stamp, F., Speekenbrink, M., & Blakemore, S. J. (2016). A window of opportunity for cognitive training in adolescence. *Psychological Science*, 27(12), 1620-1631. doi:10.1177/0956797616671327
- Knoll, L. J., Magis-Weinberg, L., Speekenbrink, M., & Blakemore, S. J. (2015). Social influence on risk perception during adolescence. *Psychological Science*, 26(5), 583-592. doi:10.1177/0956797615569578
- Knowland, V. C. P., & Thomas, M. S. C. (2014). Educating the adult brain: How the neuroscience of learning can inform educational policy. *International Review of Education*, 60(1), 99-122.
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, 16(8), 1412-1425. doi:10.1162/0898929042304796

- Knuth, E., Kalish, C., Ellis, A., Williams, C., & Felton, M. D. (2012). Adolescent reasoning in mathematical and nonmathematical domains: Exploring the paradox. In V. F. Reyna, S. B. Chapman, M. R. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 183-209). Washington, DC, US: American Psychological Association.
- Koutoumanou, E., & Wade, A. (2015a). Introduction to dealing with missing data. University College London, London, UK: Centre for Applied Statistics.
- Koutoumanou, E., & Wade, A. (2015b). Introduction to logistic regression. University College London, London, UK: Centre for Applied Statistics.
- Kraemer, H. C., Yesavage, J. A., Taylor, J. L., & Kupfer, D. (2000). How can we learn about developmental processes from cross-sectional studies, or can we? *American Journal of Psychiatry*, *157*(2), 163-171.
doi:10.1176/appi.ajp.157.2.163
- Krawczyk, D. C. (2012). The cognition and neuroscience of relational reasoning. *Brain Research*, *1428*, 13-23. doi:10.1016/j.brainres.2010.11.080
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831-843. doi:10.1038/nrn1533
- Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron*, *67*(5), 713-727. doi:10.1016/j.neuron.2010.08.038
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in Linear Mixed Effects Models. Retrieved from <https://cran.r-project.org/web/packages/lmerTest/index.html>
- Laland, K. N., & Brown, G. R. (2011). Sense and nonsense: Evolutionary perspectives on human behaviour (2nd ed.). Oxford, UK: Oxford University Press.

- Lam, C. B., McHale, S. M., & Crouter, A. C. (2014). Time with peers from middle childhood to late adolescence: developmental course and adjustment correlates. *Child Development, 85*(4), 1677-1693. doi:10.1111/cdev.12235
- Larson, R. W., & Richards, M. H. (1991). Daily companionship in late childhood and early adolescence: Changing developmental contexts. *Child Development, 62*(2), 284-300.
- Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioral Reviews, 30*(6), 718-729. doi:10.1016/j.neubiorev.2006.06.001
- Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., . . . Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage, 36*(4), 1065-1073. doi:10.1016/j.neuroimage.2007.03.053
- Lenth, R. (2016). lsmeans: Least-Squares Means. Retrieved from <https://cran.r-project.org/web/packages/lmerTest/index.html>
- Leonard, J. S. (2000). Comments on James Heckman's "Policies to foster human capital". *Research in Economics, 54*(1), 61-64. doi:10.1006/reec.1999.0220
- Lewis, T. L., & Maurer, D. (2005). Multiple sensitive periods in human visual development: Evidence from visually deprived children. *Developmental Psychobiology, 46*(3), 163-183. doi:10.1002/dev.20055
- Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science, 15*(3), 155-163. doi:10.1111/j.0956-7976.2004.01503003.x

- Lillard, A. S., & Erisir, A. (2011). Old dogs learning new tricks: Neuroplasticity beyond the juvenile period. *Developmental Review, 31*(4), 207-239.
doi:10.1016/j.dr.2011.07.008
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.
- Livesey, D. J., & Dawson, R. G. (1981). The effect of early training age upon later savings in a go/no-go task. *Behavioral & Neural Biology, 31*(1), 8-18.
doi:10.1016/S0163-1047(81)91002-5
- Loewenthal, K., & Bull, D. (1984). Imitation of foreign sounds: What is the effect of age? *Language and Speech, 27*(1), 95-97.
doi:10.1177/002383098402700108
- Løhaugen, G. C. C., Antonsen, I., Håberg, A., Gramstad, A., Vik, T., Brubakk, A. M., & Skranes, J. (2011). Computerized working memory training improves function in adolescents born at extremely low birth weight. *The Journal of Pediatrics, 158*(4), 555-561.e554. doi:10.1016/j.jpeds.2010.09.060
- Lorenz, K. (1937). Über die Bildung des Instinkt Begriffes. *Naturwissenschaften, 25*(20), 307-318. doi:10.1007/BF01491749
- Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin, 136*(4), 659.
- Luciana, M., Collins, P. F., Olson, E. A., & Schissel, A. M. (2009). Tower of London performance in healthy adolescents: The development of planning skills and associations with self-reported inattention and impulsivity. *Developmental Neuropsychology, 34*(4), 461-475.
doi:10.1080/87565640902964540

- Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development, 76*(3), 697-712. doi:10.1111/j.1467-8624.2005.00872.x
- Lustenberger, D. E., & Jagacinski, C. M. (2010). Exploring the effects of ostracism on performance and intrinsic motivation. *Human Performance, 23*(4), 283-304. doi:10.1080/08959285.2010.501046
- Mackey, A. P. (2012). *Behavioral and neural effects of reasoning training*. (PhD), University of California Berkeley, Berkeley, CA. Retrieved from <http://www.escholarship.org/uc/item/2f99262s>
- Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science, 14*(3), 582-590. doi:10.1111/j.1467-7687.2010.01005.x
- Mackey, A. P., Miller Singley, A. T., & Bunge, S. A. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *Journal of Neuroscience, 33*(11), 4796-4803. doi:10.1523/jneurosci.4141-12.2013
- Mackey, A. P., Whitaker, K. J., & Bunge, S. A. (2012). Experience-dependent plasticity in white matter microstructure: Reasoning training alters structural connectivity. *Frontiers in Neuroanatomy, 6*, 32. doi:10.3389/fnana.2012.00032
- Malone, D. T., Hill, M. N., & Rubino, T. (2010). Adolescent cannabis use and psychosis: epidemiology and neurodevelopmental models. *British Journal of Pharmacology, 160*(3), 511-522. doi:10.1111/j.1476-5381.2010.00721.x

- Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13*(4), 330-334.
doi:10.1037/h0028434
- Marin, O. (2016). Developmental timing and critical windows for the treatment of psychiatric disorders. *Nature Medicine, 22*(11), 1229-1238.
doi:10.1038/nm.4225
- Maroun, M., Ioannides, P. J., Bergman, K. L., Kavushansky, A., Holmes, A., & Wellman, C. L. (2013). Fear extinction deficits following acute stress associate with increased spine density and dendritic retraction in basolateral amygdala neurons. *European Journal of Neuroscience, 38*(4), 2611-2620.
- Maurer, D., & Lewis, T. (2012). Human visual plasticity: Lessons from children treated for congenital cataracts *Plasticity in Sensory Systems* (pp. 75): Cambridge University Press.
- Maurer, D., & Mondloch, C. J. (2011). Sensitive periods in face perception. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception*. Oxford, UK: Oxford University Press.
- Maya Vetencourt, J. F., Sale, A., Viegi, A., Baroncelli, L., De Pasquale, R., O'Leary, O. F., . . . Maffei, L. (2008). The antidepressant fluoxetine restores plasticity in the adult visual cortex. *Science, 320*(5874), 385-388.
doi:10.1126/science.1150516
- McArdle, J. J. (2008). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605.
doi:10.1146/annurev.psych.60.110707.163612

- McClelland, G. H., Lynch, J. G., Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, Type II errors, and false–positive consumer psychology: Don't fight the power. *Journal of Consumer Psychology, 25*(4), 679-689. doi:10.1016/j.jcps.2015.05.006
- McGee, A. W., Yang, Y., Fischer, Q. S., Daw, N. W., & Strittmatter, S. M. (2005). Experience-driven plasticity of visual cortex limited by myelin and Nogo receptor. *Science, 309*(5744), 2222-2226. doi:10.1126/science.1114362
- McGivern, R. F., Andersen, J., Byrd, D., Mutter, K. L., & Reilly, J. (2002). Cognitive efficiency on a match to sample task decreases at the onset of puberty in children. *Brain and Cognition, 50*(1), 73-89.
- McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. *Cognitive Neuropsychology, 29*(1-2), 174-212. doi:10.1080/02643294.2012.660138
- Meier, M. H., Caspi, A., Ambler, A., Harrington, H., Houts, R., Keefe, R. S. E., . . . Moffitt, T. E. (2012). Persistent cannabis users show neuropsychological decline from childhood to midlife. *Proceedings of the National Academy of Sciences, 109*(40), E2657-E2664. doi:10.1073/pnas.1206820109
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270.
- Mery, F., & Kawecki, T. J. (2003). A fitness cost of learning ability in *Drosophila melanogaster*. *Proceedings of the Royal Society of London B: Biological Sciences, 270*(1532), 2465-2469.
- Mery, F., & Kawecki, T. J. (2005). A cost of long-term memory in *Drosophila*. *Science, 308*. doi:10.1126/science.1111331

- Meuwese, R., Crone, E. A., de Rooij, M., & Guroglu, B. (2015). Development of equity preferences in boys and girls across adolescence. *Child Development, 86*(1), 145-158. doi:10.1111/cdev.12290
- Michel, G. F., & Tyler, A. N. (2005). Critical period: A history of the transition from questions of when, to what, to how. *Developmental Psychobiology, 46*(3), 156-162. doi:10.1002/dev.20058
- Miller, D. J., Duka, T., Stimpson, C. D., Schapiro, S. J., Baze, W. B., McArthur, M. J., . . . Sherwood, C. C. (2012). Prolonged myelination in human neocortical evolution. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1117943109
- Mills, K. L., Goddings, A. L., Clasen, L. S., Giedd, J. N., & Blakemore, S. J. (2014). The developmental mismatch in structural brain maturation during adolescence. *Developmental Neuroscience, 36*(3-4), 147-160. doi:10.1159/000362328
- Mills, K. L., Goddings, A. L., Herting, M. M., Meuwese, R., Blakemore, S. J., Crone, E. A., . . . Tamnes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage, 141*, 273-281. doi:10.1016/j.neuroimage.2016.07.044
- Mills, K. L., & Tamnes, C. K. (2014). Methods and considerations for longitudinal structural brain imaging analysis across development. *Developmental Cognitive Neuroscience, 9*, 172-190. doi:10.1016/j.dcn.2014.04.004
- Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology, 9*(1), 90-100. doi:10.1001/archneur.1963.00460070100010

- Mondloch, C. J., Dobson, K. S., Parsons, J., & Maurer, D. (2004). Why 8-year-olds cannot tell the difference between Steve Martin and Paul Newman: Factors contributing to the slow development of sensitivity to the spacing of facial features. *Journal of Experimental Child Psychology, 89*(2), 159-181. doi:10.1016/j.jecp.2004.07.002
- Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*(5), 553-566. doi:10.1068/p3339
- Mondloch, C. J., Robbins, R., & Maurer, D. (2010). Discrimination of facial features by adults, 10-year-olds, and cataract-reversal patients. *Perception, 39*(2), 184-194.
- Murre, J. M., Janssen, S. M., Rouw, R., & Meeter, M. (2013). The rise and fall of immediate and delayed memory for verbal and visuospatial information from late childhood to late adulthood. *Acta Psychologica, 142*(1), 96-107. doi:10.1016/j.actpsy.2012.10.005
- Mussolin, C., Nys, J., Content, A., & Leybaert, J. (2014). Symbolic number abilities predict later Approximate Number System acuity in preschool children. *PLoS One, 9*(3), e91839. doi:10.1371/journal.pone.0091839
- Myers, C. A., Wang, C., Black, J. M., Bugescu, N., & Hoeft, F. (2016). The matter of motivation: Striatal resting-state connectivity is dissociable between grit and growth mindset. *Social Cognitive and Affective Neuroscience, 11*(10), 1521-1527. doi:10.1093/scan/nsw065
- Myers, T. A. (2011). Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handling Missing Data. *Communication*

Methods and Measures, 5(4), 297-310.

doi:10.1080/19312458.2011.624490

- Nakamoto, J., & Schwartz, D. (2010). Is peer victimization associated with academic achievement? A meta-analytic review. *Social Development*, 19(2), 221-242. doi:10.1111/j.1467-9507.2009.00539.x
- Neale, M. D. (1999). *Neale Analysis of Reading Ability (NARA)*. London, UK: GL Assessment.
- Nemmi, F., Nymberg, C., Helander, E., & Klingberg, T. (2016). Grit is associated with structure of nucleus accumbens and gains in cognitive training. *Journal of Cognitive Neuroscience*, 28(11), 1688-1699. doi:10.1162/jocn_a_01031
- Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, 2(3), 244-258.
- Newsom, J. T. (2015). *Longitudinal structural equation modelling - A comprehensive introduction*. New York, NY: Routledge.
- O'Brien, S. F., & Bierman, K. L. (1988). Conceptions and perceived influence of peer groups: interviews with preadolescents and adolescents. *Child Development*, 59(5), 1360-1365.
- Olsson, A., & Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences*, 12(2), 65-71. doi:10.1016/j.tics.2007.11.010
- Oudin, M. J., Gajendra, S., Williams, G., Hobbs, C., Lalli, G., & Doherty, P. (2011). Endocannabinoids regulate the migration of subventricular zone-derived neuroblasts in the postnatal brain. *Journal of Neuroscience*, 31(11), 4000-4011. doi:10.1523/JNEUROSCI.5483-10.2011

- Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., . . . Ballard, C. G. (2010). Putting brain training to the test. *Nature*, *465*(7299), 775-778. doi:10.1038/nature09042
- Paluck, E. L., Shepherd, H., & Aronow, P. M. (2016). Changing climates of conflict: A social network experiment in 56 schools. *Proceedings of the National Academy of Sciences*, *113*(3), 566-571.
- Panchanathan, K., & Frankenhuis, W. E. (2016). The evolution of sensitive periods in a model of incremental development. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1823).
- Park, J., & Brannon, E. M. (2013). Training the approximate number system improves Math proficiency. *Psychological Science*, *24*(10), 2013-2019. doi:10.1177/0956797613482944
- Parker, G. A., & Smith, J. M. (1990). Optimality theory in evolutionary biology. *Nature*, *348*(6296), 27-33.
- Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual Review of Neuroscience*, *28*, 377-401. doi:10.1146/annurev.neuro.27.070203.144216
- Pattwell, S. S., Duhoux, S., Hartley, C. A., Johnson, D. C., Jing, D., Elliott, M. D., . . . Lee, A. S. (2012). Altered fear learning across development in both mouse and human. *Proceedings of the National Academy of Sciences*, *109*(40), 16318-16323. doi:10.1073/pnas.1206834109
- Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences*, *9*(2), 60-68. doi:10.1016/j.tics.2004.12.008

- Peelle, J. E., Cusack, R., & Henson, R. N. A. (2012). Adjusting for global effects in voxel-based morphometry: Gray matter decline in normal aging. *NeuroImage*, *60*(2), 1503-1516. doi:10.1016/j.neuroimage.2011.12.086
- Penner, I. K., Vogt, A., Stocklin, M., Gschwind, L., Opwis, K., & Calabrese, P. (2012). Computerised working memory training in healthy adults: A comparison of two different training schedules. *Neuropsychological Rehabilitation*, *22*(5), 716-733. doi:10.1080/09602011.2012.686883
- Pfeifer, J. H., & Allen, N. B. (2012). Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. *Trends in Cognitive Science*, *16*(6), 322-329. doi:10.1016/j.tics.2012.04.011
- Pfeifer, J. H., Masten, C. L., Moore, W. E., 3rd, Oswald, T. M., Mazziotta, J. C., Iacoboni, M., & Dapretto, M. (2011). Entering adolescence: Resistance to peer influence, risky behavior, and neural changes in emotion reactivity. *Neuron*, *69*(5), 1029-1036. doi:10.1016/j.neuron.2011.02.019
- Pharo, H., Gross, J., Richardson, R., & Hayne, H. (2011). Age-related changes in the effect of ostracism. *Social Influence*, *6*(1), 22-38. doi:10.1080/15534510.2010.525852
- Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (2000). The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(10), 1090-1104.
- Piaget, J., Grize, J.-B., Szeminska, A., & Vinh, B. (1977). *Epistemology and psychology of functions*. (Transl. J. Castellanos & V. Anderson). Oxford, England: D. Reidel.

- Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, 14(12), 542-551.
doi:10.1016/j.tics.2010.09.008
- Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education enhances the acuity of the nonverbal approximate number system. *Psychological Science*, 24(6), 1037-1043. doi:10.1177/0956797612464057
- Piepers, D. W., & Robbins, R. A. (2012). A Review and Clarification of the Terms “holistic,” “configural,” and “relational” in the Face Perception Literature. *Frontiers of Psychology*, 3, 559. doi:10.3389/fpsyg.2012.00559
- Pope, H. G., Gruber, A. J., Hudson, J. I., Cohane, G., Huestis, M. A., & Yurgelun-Todd, D. (2003). Early-onset cannabis use and cognitive deficits: what is the nature of the association? *Drug and Alcohol Dependence*, 69(3), 303-310.
- Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., & Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of Experimental Child Psychology*, 108(3), 621-637.
doi:10.1016/j.jecp.2010.09.008
- Prinzie, P., & Onghena, P. (2014). Cohort Sequential Design *Wiley StatsRef: Statistics Reference Online*: John Wiley & Sons, Ltd.
- R Core Team (2015). R: A language and environment for statistical computing. Retrieved from <http://www.R-project.org/>
- Raven, J. C. (2009). The Raven Progressive Matrices and measuring aptitude constructs. *The International Journal of Educational and Psychological Assessment*, 2, 2-38.

- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017).
The human fetus preferentially engages with face-like visual stimuli.
Current Biology, 27(13), 2052. doi:10.1016/j.cub.2017.06.036
- Reiss, A. L., Abrams, M. T., Singer, H. S., Ross, J. L., & Denckla, M. B. (1996). Brain
development, gender and IQ in children. *Brain*, 119(5), 1763-1774.
- Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of
analogical reasoning: Insights from scene analogy problems. *Journal of
Experimental Child Psychology*, 94(3), 249-273.
doi:10.1016/j.jecp.2006.02.002
- Ridgway, G. (2007). get_totals. University College London, London, UK. Retrieved
from http://www0.cs.ucl.ac.uk/staff/gridgway/vbm/get_totals.m
- Rigby, K. (2000). Effects of peer victimization in schools and perceived social
support on adolescent well-being. *Journal of Adolescence*, 23(1), 57-68.
doi:10.1006/jado.1999.0289
- Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental
changes in mental arithmetic: Evidence for increased functional
specialization in the left inferior parietal cortex. *Cerebral Cortex*, 15(11),
1779-1790. doi:10.1093/cercor/bhi055
- Röder, B., Ley, P., Shenoy, B. H., Kekunnaya, R., & Bottari, D. (2013). Sensitive
periods for the functional specialization of the neural system for human
face processing. *Proceedings of the National Academy of Sciences*,
110(42), 16760-16765. doi:10.1073/pnas.1309963110
- Rodriguez de Fonseca, F., Ramos, J. A., Bonnin, A., & Fernandez-Ruiz, J. J. (1993).
Presence of cannabinoid binding sites in the brain from early postnatal
ages. *Neuroreport*, 4(2), 135-138.

- Roe, D. J., & Korn, E. L. (1993). Time-period effects in longitudinal studies measuring average rates of change. *Statistical Medicine*, *12*(9), 893-900.
- Romer, D., Reyna, V. F., & Satterthwaite, T. D. (2017). Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Developmental Cognitive Neuroscience*, *27*, 19-34.
doi:10.1016/j.dcn.2017.07.007
- Rosenthal, D. (1970). *Genetic theory and abnormal behavior*. New York, NY: McGraw-Hill.
- Rosseel, Y. (2012). lavaan: An R package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1-36.
- Rubin, D. C., & Schulkind, M. D. (1997). The distribution of autobiographical memories across the lifespan. *Memory & Cognition*, *25*(6), 859-866.
- Rucker, D. D., McShane, B. B., & Preacher, K. J. (2015). A researcher's guide to regression, discretization, and median splits of continuous variables. *Journal of Consumer Psychology*, *25*(4), 666-678.
doi:10.1016/j.jcps.2015.04.004
- Sabol, T. J., Chase-Lansdale, P. L., & Brooks-Gunn, J. (2015). Advancing the science of child and adolescent development: Do we need a new household panel survey? *Journal of Economic and Social Measurement*, *40*(1-4), 221-255.
- Sakai, K. L. (2005). Language Acquisition and Brain Development. *Science*, *310*(5749), 815-819. doi:10.1126/science.1113530
- Sampaio-Baptista, C., Scholz, J., Jenkinson, M., Thomas, A. G., Filippini, N., Smit, G., ... & Johansen-Berg, H. (2014). Gray matter volume is associated with rate of subsequent skill learning after a long term training intervention. *NeuroImage*, *96*, 158-166.

- Scherf, K. S., Behrmann, M., & Dahl, R. E. (2012). Facing changes and changing faces in adolescence: A new model for investigating adolescent-specific interactions between pubertal, brain and behavioral development. *Developmental Cognitive Neuroscience, 2*(2), 199-219.
doi:10.1016/j.dcn.2011.07.016
- Schneider, M. (2013). Adolescence as a vulnerable period to alter rodent behavior. *Cell and Tissue Research, 354*(1), 99-106. doi:10.1007/s00441-013-1581-2
- Schneider, M., Drews, E., & Koch, M. (2005). Behavioral effects in adult rats of chronic prepubertal treatment with the cannabinoid receptor agonist WIN 55,212-2. *Behavioural Pharmacology, 16*(5-6), 447-454.
- Schneider, M., & Koch, M. (2003). Chronic pubertal, but not adult chronic cannabinoid treatment impairs sensorimotor gating, recognition memory, and the performance in a progressive ratio task in adult rats. *Neuropsychopharmacology, 28*(10), 1760-1769.
doi:10.1038/sj.npp.1300225
- Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50*(2), 138-166.
- Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science, 16*(4), 197-201. doi:10.1111/j.1467-8721.2007.00503.x
- Sebastian, C. L., Fontaine, N. M. G., Bird, G., Blakemore, S. J., De Brito, S. A., McCrory, E. J. P., & Viding, E. (2012). Neural processing associated with

cognitive and affective Theory of Mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, 7(1), 53-63.

doi:10.1093/scan/nsr023

Sebastian, C. L., Tan, G. C., Roiser, J. P., Viding, E., Dumontheil, I., & Blakemore, S.

J. (2011). Developmental influences on the neural bases of responses to social rejection: Implications of social neuroscience for education.

NeuroImage, 57(3), 686-694. doi:10.1016/j.neuroimage.2010.09.063

Sebastian, C. L., Viding, E., Williams, K. D., & Blakemore, S. J. (2010). Social brain

development and the affective consequences of ostracism in adolescence.

Brain and Cognition, 72(1), 134-145. doi:10.1016/j.bandc.2009.06.008

Selemon, L. D. (2013). A role for synaptic plasticity in the adolescent development

of executive function. *Translational Psychiatry*, 3, e238.

doi:10.1038/tp.2013.7

Sharp, S. (1995). How much does bullying hurt? The effects of bullying on the

personal wellbeing and educational progress of secondary aged students.

Educational and Child Psychology, 12(2), 81-88.

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg,

L. (2016). The dual systems model: Review, reappraisal, and reaffirmation.

Developmental Cognitive Neuroscience, 17, 103-117.

doi:10.1016/j.dcn.2015.12.010

Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., &

Petersen, S. E. (2014). Statistical improvements in functional Magnetic

Resonance Imaging analyses produced by censoring high-motion data

points. *Human Brain Mapping*, 35(5), 1981-1996. doi:10.1002/hbm.22307

- Sigurdson, J. F., Undheim, A. M., Wallander, J. L., Lydersen, S., & Sund, A. M. (2015). The long-term effects of being bullied or a bully in adolescence on externalizing and internalizing mental health problems in adulthood. *Child and Adolescent Psychiatry and Mental Health, 9*, 42. doi:10.1186/s13034-015-0075-2
- Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia, 46*(1), 224-232. doi:10.1016/j.neuropsychologia.2007.07.015
- Sirin, S. R., Ryce, P., Gupta, T., & Rogers-Sirin, L. (2013). The role of acculturative stress on mental health symptoms for immigrant adolescents: A longitudinal investigation. *Developmental Psychology, 49*(4), 736-748. doi:10.1037/a0028398
- Sisk, C. L., & Foster, D. L. (2004). The neural basis of puberty and adolescence. *Nature Neuroscience, 7*(10), 1040-1047. doi:10.1038/nn1326
- Sisk, C. L., & Zehr, J. L. (2005). Pubertal hormones organize the adolescent brain and behavior. *Frontiers in Neuroendocrinology, 26*(3-4), 163-174. doi:10.1016/j.yfrne.2005.10.003
- Sisti, H. M., Glass, A. L., & Shors, T. J. (2007). Neurogenesis and the spacing effect: Learning over time enhances memory and the survival of new neurons. *Learning & Memory, 14*(5), 368-375. doi:10.1101/lm.488707
- Somerville, L. H., Hare, T., & Casey, B. J. (2011). Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. *Journal of Cognitive Neuroscience, 23*(9), 2123-2134. doi:10.1162/jocn.2010.21572

- Sommer, W., Hildebrandt, A., Kunina-Habenicht, O., Schacht, A., & Wilhelm, O. (2013). Sex differences in face cognition. *Acta Psychologica, 142*(1), 62-73. doi:10.1016/j.actpsy.2012.11.001
- Song, Y., Zhu, Q., Li, J., Wang, X., & Liu, J. (2015). Typical and atypical development of functional connectivity in the face network. *Journal of Neuroscience, 35*(43), 14624-14635. doi:10.1523/jneurosci.0969-15.2015
- Spear, L. P. (2013). Adolescent Neurodevelopment. *The Journal of Adolescent Health, 52*(2), S7-S13.
- Spires-Jones, T. L., Poirazi, P., & Grubb, M. S. (2016). Opening up: Open access publishing, data sharing, and how they can influence your neuroscience career. *European Journal of Neuroscience, 43*(11), 1413-1419.
- Spring, M. (2016). Would you get into grammar school? Try the 11-plus exam. Retrieved from <https://www.theguardian.com/education/2016/sep/09/would-you-get-into-grammar-school-take-our-11-plus-exam>
- Stamps, J. A., & Krishnan, V. V. (2014). Individual differences in the potential and realized developmental plasticity of personality traits. *Frontiers in Ecology & Evolution, 2*. doi:10.3389/fevo.2014.00069
- Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980–2013. *International Journal of Epidemiology, 43*(2), 476-493. doi:10.1093/ije/dyu038
- Steinbeis, N., & Crone, E. A. (2016). The link between cognitive control and decision-making across child and adolescent development. *Current*

Opinion in Behavioral Sciences, 10, 28-32.

doi:10.1016/j.cobeha.2016.04.009

Steinbeis, N., & Singer, T. (2013). The effects of social comparison on social emotions and behavior during childhood: The ontogeny of envy and Schadenfreude predicts developmental changes in equity-related decisions. *Journal of Experimental Child Psychology*, 115(1), 198-209.

doi:10.1016/j.jecp.2012.11.009

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking.

Developmental Review, 28(1), 78-106. doi:10.1016/j.dr.2007.08.002

Steinberg, L. (2014). *Age of opportunity - Lessons from the new science of adolescence*. Boston, MA: Houghton Mifflin Harcourt.

Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., Bacchini, D., . . .

Takash, H. M. S. (2017). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation.

Developmental Science, e12532-n/a. doi:10.1111/desc.12532

Steinberg, L., & Monahan, K. C. (2007). Age differences in resistance to peer influence. *Developmental Psychology*, 43(6), 1531-1543.

doi:10.1037/0012-1649.43.6.1531

Strang, N. M., Chein, J. M., & Steinberg, L. (2013). The value of the dual systems model of adolescent risk-taking. *Frontiers in Human Neuroscience*, 7, 223.

doi:10.3389/fnhum.2013.00223

Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces.

Proceedings of the National Academy of Sciences, 105(1), 394-398.

doi:10.1073/pnas.0706079105

- Sullivan, G. M. (2011). Getting off the “gold standard”: Randomized controlled trials and education research. *Journal of Graduate Medical Education*, 3(3), 285-289. doi:10.4300/JGME-D-11-00147.1
- Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities*, 39(3), 252-269. doi:10.1177/00222194060390030501
- Takesian, A. E., & Hensch, T. K. (2013). Balancing plasticity/stability across brain development. *Progress in Brain Research*, 207, 3-34. doi:10.1016/b978-0-444-63327-9.00001-1
- Takizawa, R., Maughan, B., & Arseneault, L. (2014). Adult health outcomes of childhood bullying victimization: Evidence from a five-decade longitudinal British birth cohort. *American Journal of Psychiatry*, 171(7), 777-784. doi:10.1176/appi.ajp.2014.13101401
- 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(10), 1231-1238. doi:10.1097/00004583-200210000-00013
- Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., . . . Mills, K. L. (2017). Development of the cerebral cortex across adolescence: A multisample study of interrelated longitudinal changes in cortical volume, surface area and thickness. *Journal of Neuroscience*. doi:10.1523/JNEUROSCI.3302-16.2017
- Tamnes, C. K., Walhovd, K. B., Dale, A. M., Ostby, Y., Grydeland, H., Richardson, G., . . . Fjell, A. M. (2013). Brain development and aging: Overlapping and

unique patterns of change. *NeuroImage*, 68, 63-74.

doi:10.1016/j.neuroimage.2012.11.039

Tamnes, C. K., Walhovd, K. B., Grydeland, H., Holland, D., Ostby, Y., Dale, A. M., & Fjell, A. M. (2013). Longitudinal working memory development is related to structural maturation of frontal and parietal cortices. *Journal of Cognitive Neuroscience*, 25(10), 1611-1623. doi:10.1162/jocn_a_00434

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective & Behavioral Neuroscience*, 9(1), 122-131. doi:10.3758/cabn.9.1.122

Tetens, J. N. (1777). *Philosophische Versuche über die menschliche Natur und ihre Entwicklung*: Bernasconi, S.E.

The MathWorks (2013). MATLAB and Statistics Toolbox Release. Natick, MA.

The NHS Information Centre (2011). *Statistics on drug misuse: England, 2011*.

Retrieved from

https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0ahUKEwj476LRz_fVAhXMEVAKHcbGAowQFgguMAE&url=https%3A%2F%2Fdigital.nhs.uk%2Fmedia%2F17223%2FStatistics-on-Drug-Misuse-England-2011-Report%2FAny%2Fdrug-misu-eng-2011-rep&usg=AFQjCNHNYvfZYb20DH2FC15SGb1jrOiiWw

The Royal Society (2011). *Neuroscience: Implications for education and lifelong learning*. Retrieved from <https://royalsociety.org/topics-policy/projects/brain-waves/education-lifelong-learning/>

Themanson, J. R., Ball, A. B., Khatcherian, S. M., & Rosen, P. J. (2014). The effects of social exclusion on the ERN and the cognitive control of action monitoring. *Psychophysiology*, 51(3), 215-225.

- Thomas, M. S. C. (2008). *Ageing, plasticity, and cognitive reserve in connectionist networks*. Paper presented at the 30th Annual Conference of the Cognitive Science Society, Austin, TX.
- Thomas, M. S. C. (2012). Brain plasticity and education. *British Journal of Educational Psychology-Monograph Series II: Educational Neuroscience, 8*, 142-156.
- Thomas, M. S. C., & Johnson, M. H. (2008). New advances in understanding sensitive periods in brain development. *Current Directions in Psychological Science, 17*(1), 1-5. doi:10.1111/j.1467-8721.2008.00537.x
- Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*(1), 106-113. doi:10.1111/j.1467-7687.2008.00745.x
- Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science, 10*(1), 121-125. doi:10.1111/j.1467-7687.2007.00573.x
- Ttofi, M. M., & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: a systematic and meta-analytic review. *Journal of Experimental Criminology, 7*(1), 27-56. doi:10.1007/s11292-010-9109-1
- Ungerleider, L. G., Doyon, J., & Karni, A. (2002). Imaging brain plasticity during motor skill learning. *Neurobiology of Learning and Memory, 78*(3), 553-564.
- Uttl, B. (2005). Measurement of individual differences: Lessons from memory assessment in research and clinical practice. *Psychological Science, 16*(6), 460-467. doi:10.1111/j.0956-7976.2005.01557.x

- van den Bos, W., Guroglu, B., van den Bulk, B. G., Rombouts, S. A., & 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. doi:10.3389/neuro.09.052.2009
- van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *Proceedings of the National Academy of Sciences*, *112*(29), E3765-E3774.
- van Harmelen, A. L., Kievit, R. A., Ioannidis, K., Neufeld, S., Jones, P. B., Bullmore, E., . . . Goodyer, I. (2017). Adolescent friendships predict later resilient functioning across psychosocial domains in a healthy community cohort. *Psychological Medicine*, 1-11. doi:10.1017/s0033291717000836
- van Leijenhorst, L., Moor, B. G., 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*(1), 345-355. doi:10.1016/j.neuroimage.2010.02.038
- Vara, A. S., Pang, E. W., Doyle-Thomas, K. A. R., Vidal, J., Taylor, M. J., & Anagnostou, E. (2014). Is inhibitory control a 'no-go' in adolescents with autism spectrum disorder? *Molecular Autism*, *5*, 6-6. doi:10.1186/2040-2392-5-6
- Vara, A. S., Pang, E. W., Vidal, J., Anagnostou, E., & Taylor, M. J. (2014). Neural mechanisms of inhibitory control continue to mature in adolescence. *Developmental Cognitive Neuroscience*, *10*, 129-139. doi:10.1016/j.dcn.2014.08.009

- 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*(11), 2505-2522. doi:10.1093/cercor/bhn012
- Venables, W. N., & Ripley, B. D. (2002). Random and Mixed Effects. In W. N. Venables & B. D. Ripley (Eds.), *Modern Applied Statistics with S* (pp. 271-300). New York, NY: Springer New York.
- Ver Hoeve, E. S., Kelly, G., Luz, S., Ghanshani, S., & Bhatnagar, S. (2013). Short-term and long-term effects of repeated social defeat during adolescence or adulthood in female rats. *Neuroscience*, *249*, 63-73. doi:10.1016/j.neuroscience.2013.01.073
- Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes*. Boston, MA: Harvard University Press.
- Walhovd, K. B., Krogsrud, S. K., Amlien, I. K., Bartsch, H., Bjornerud, A., Due-Tonnessen, P., . . . Fjell, A. M. (2016). Neurodevelopmental origins of lifespan changes in brain and cognition. *Proceedings of the National Academy of Sciences*, *113*(33), 9357-9362. doi:10.1073/pnas.1524259113
- Wang, J., Iannotti, R. J., Luk, J. W., & Nansel, T. R. (2010). Co-occurrence of victimization from five subtypes of bullying: Physical, verbal, social exclusion, spreading rumors, and cyber. *Journal of Pediatric Psychology*, *35*(10), 1103-1112. doi:10.1093/jpepsy/jsq048
- Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2009). Investigating ceiling effects in longitudinal data analysis. *Multivariate Behavioral Research*, *43*(3), 476-496. doi:10.1080/00273170802285941

- Wang, S., Zhou, M., Chen, T., Yang, X., Chen, G., Wang, M., & Gong, Q. (2016). Grit and the brain: Spontaneous activity of the dorsomedial prefrontal cortex mediates the relationship between the trait grit and academic performance. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsw145
- Wang, Z., Zhou, R., & Shah, P. (2014). Spaced cognitive training promotes training transfer. *Frontiers in Human Neuroscience*, 8, 217. doi:10.3389/fnhum.2014.00217
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. New York, NY.
- Wellcome Trust Centre for Neuroimaging (2014). SPM - Statistical Parametric Mapping. University College London, London, UK. Retrieved from <http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>
- Wenger, E., Brozzoli, C., Lindenberger, U., & Lövdén, M. (2017). Expansion and renormalization of human brain structure during acquisition. *Trends in Cognitive Sciences*, 21(12), 930-939.
- Wenger, E., & Lövdén, M. (2016). The Learning Hippocampus: Education and Experience-Dependent Plasticity. *Mind, Brain, and Education*, 10(3), 171-183.
- Wesselmann, E. D., Ren, D., Swim, E., & Williams, K. D. (2013). Rumination hinders recovery from ostracism. *International Journal of Developmental Science*, 7(1), 33-39.
- Wiesel, T. N., & Hubel, D. H. (1963). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology*, 26(6), 1003-1017.

- Wiesel, T. N., & Hubel, D. H. (1965a). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of Neurophysiology*, *28*(6), 1029-1040.
- Wiesel, T. N., & Hubel, D. H. (1965b). Extent of recovery from the effects of visual deprivation in kittens. *Journal of Neurophysiology*, *28*(6), 1060-1072.
- Wilhelm, O., Herzmann, G., Kunina, O., & Sommer, W. (2007). *Face cognition: A set of distinct mental abilities. Nature Proceedings.*
- Will, G. J., van Lier, P. A. C., Crone, E. A., & Güroğlu, B. (2016). Chronic childhood peer rejection is associated with heightened neural responses to social exclusion during adolescence. *Journal of Abnormal Child Psychology*, *44*, 43-55. doi:10.1007/s10802-015-9983-0
- Williams, K. D. (2007). Ostracism. *Annual Review of Psychology*, *58*, 425-452. doi:10.1146/annurev.psych.58.110405.085641
- Williams, K. D., Cheung, C. K. T., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, *79*(5), 748-762. doi:10.1037/0022-3514.79.5.748
- Williams, K. D., Yeager, D. S., Cheung, C. K. T., & Choi, W. (2012). Cyberball 4.0. Retrieved from <https://cyberball.wikispaces.com/>
- Wolf, L. K., Bazargani, N., Kilford, E. J., Dumontheil, I., & Blakemore, S.-J. (2015). The audience effect in adolescence depends on who's looking over your shoulder. *Journal of Adolescence*, *43*, 5-14. doi:10.1016/j.adolescence.2015.05.003
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749-750. doi:10.1038/358749a0

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives, 6*(4), 354-360.

Appendix 2.1: R script

This code was developed with L. J. Knoll and M. Speekenbrink for:

Fuhrmann, D., Knoll, L.J.*, Sakhardande, A., Stamp, F., Speekenbrink, M. & Blakemore, S-J. (2016). A window of opportunity for cognitive training in adolescence. Psychological Science, 27(12), 1620-1631. doi: 10.1177/0956797616671327 *Joint first authors.*

This script simulates a developmental dataset in R (R Core Team, 2015) and analyses it using Generalized Linear Mixed Models.

Suppose, we conduct a training study with three different age groups: 7 - 10 year-olds, 11 - 14 year-olds, and 15 - 18 year-olds. Participants are randomly allocated to a training or control condition. The researchers hypothesize that 11 - 14 year-olds will show higher training effects than the other two age groups. The dependent variable is a cognitive task with 40 trials. This variable is dichotomous (correct/incorrect). The researchers also want to control for IQ, in case age groups differ in this respect.

```
#####  
#####  
### Load the R packages you will need  
  
# Install packages - this only needs to be done once  
install.packages("car")  
install.packages("lme4")  
install.packages("lsmeans")  
install.packages("multcomp")  
install.packages("Hmisc")  
install.packages("Rmisc")  
install.packages("doBy")  
install.packages("ggplot2")  
  
# Load packages for use  
library(lme4)  
library(lsmeans)  
library(multcomp)  
library(Hmisc)
```

```

library(doby)
library(car)
library(Rmisc)
library(ggplot2)

#####
#####
### Simulate a simple dataset
set.seed(1111)

id= rep(1:600, times=40) # Create 600 participant IDs, repeat
them 40 times. We want to create 40 trials per participant.

x1 = sample(rep(7:18, length.out=600)) # Assign each participant
a random age between 7 and 18.
age = rep(x1, times=40)

x2 = sample(rep(1:4, length.out=600)) # Assign each participant
to one of 4 schools.
school = rep(x2, times=40)

x3 = sample(rep(1:2, length.out=600)) # Assign each participant
to either training or control group.
treatment_group = rep(x3, times=40)

x4 = sample(rnorm(n = 600, mean = 100, sd = 10)) # Assign each
participant an IQ score.
iq = rep(x4, times=40)

# Create 3, roughly even-sized categorical age groups
age_group <- as.numeric(cut2(age, g=3))

# Turn categorical variables into factors
id = as.factor(id)
age_group = as.factor(age_group)
school = as.factor(school)
treatment_group = as.factor(treatment_group)

# Give the levels of the categorical factors proper names
levels(age_group) <- c("7-10", "11-14", "15-18")
levels(school) <- c("School1", "School2", "School3", "School4")
levels(treatment_group) <- c("Training", "Control")

# Put all variables into one data frame called "data"
data = data.frame(id, iq, age, age_group, school,
treatment_group)

# Create a helper variable called int - the interaction of
age_group and treatment group
data$int = interaction(data$age_group, data$treatment_group)

```

```

# Create our dependent variable, fill it with 50% 1s and 0s
data$correct = sample(0:1, 24000, replace=T, prob=c(0.5,0.5))

# Give participants aged 11-10 who were in the training group a
higher number of correct trials
data[data$int=="11-14.Training",]$correct = sample(0:1, 3680,
replace=T, prob=c(0.45,0.55))

# Delete int - we won't need it any more
data$int = NULL

#####
#####
### Summarize the data

# This step is not necessary to run the GLMM but can speed up the
computation time considerably
data_summary <- summaryBy (correct ~ id + iq + age_group +
school + treatment_group, # Summarize the dependent variable
for each level of the other variables.
                           data = data,
                           FUN = c(sum, length, mean))
# sum: number of correct trials
# length: total number of trials
# mean: average accuracy - will be used for plotting

#####
#####
### Plot the data

# Summarize the data for each age and training group
plot_data <- summarySE(data_summary, measurevar="correct.mean",
groupvars=c("age_group", "treatment_group"))

# Convert the accuracy data to %
plot_data$correct.mean <- plot_data$correct.mean*100
plot_data$se <- plot_data$se*100

# Plot using ggplot
ggplot(data=plot_data, aes(x=age_group, y=correct.mean,
group=treatment_group)) +
  geom_bar(aes(fill=treatment_group),position = "dodge",
stat="identity")+
  geom_errorbar(aes(ymin=correct.mean-se,
ymax=correct.mean+se),
               width=.1,
               position=position_dodge(.9),stat="identity")+
  labs(x="Age group",y="Accuracy (%)")+
  coord_cartesian(ylim = c(40, 60))+
  theme(axis.ticks.x = element_blank()+
theme(axis.title.y = element_text(face="bold",size=12,

```

```

colour="black",vjust = 1.5)))+
  theme(axis.text.y = element_text(size=12))+
  theme(axis.text.x = element_text(face="bold",size=12,
colour="black")))+
  theme(axis.title.x = element_blank()+
  theme(legend.title = element_text(size=12, face="bold")))+
  theme(legend.text = element_text(size=12))+
  scale_fill_brewer(palette="Blues",name="Age group")

#####
#####
### Run the GLMM

# Set the contrast-coding scheme for our categorical fixed
effects
# The default is Dummy-coding, but it can be more useful to use
an orthogonal coding scheme like Helmert-coding. For more
information see http://stats.idre.ucla.edu/r/library/r-library-contrast-coding-systems-for-categorical-variables/
contrasts(data_summary$age_group)<-contr.helmert(3) # Set
Helmert-contrasts for the three levels of age group
contrasts(data_summary$treatment_group)<-contr.helmert(2) #
Set Helmert-contrasts for the two levels of training group.

# Z-score IQ
# Z-scoring can help interpret effects because the transforms
variable's mean will be 0. Z-scoring particularly useful when
continuous variables are part of interactions, because it can
reduce multicollinearity.
data_summary$iq = scale(data_summary$iq, center = TRUE, scale
= TRUE)

# Run the GLMM
model = glmer(cbind(correct.sum, correct.length-correct.sum)
~ # We specify dependent variable as the number of correct trials
(correct.sum) and the number of incorrect trials (correct.length-
correct.sum)

our fixed          age_group * treatment_group + # These are
effects of interest and their interaction

                  iq + # We also want to control for IQ

                  (1|school/id), # These are our nested
random effects. Participant ID is nested within school

                  data=data_summary,

                  family = binomial) # Use binomial as the
dependent variable is dichotomous

```



```
#####
#####
### Inspect the results

# Look at results in an ANOVA-style table
Anova(model,type=3)

# Inspect the intercept and the slopes ("Estimate") of our
effects
summary(model)

# The contrasts of the main effects can be accessed using lsmeans
lsmeans(model, pairwise ~ age_group)

# Some contrasts of the interaction can be inspected using
lsmeans
lsmeans(model, pairwise ~ age_group|treatment_group)

# More complex contrasts can be analysed using custom contrasts
dummydat <- aggregate(iq ~ age_group * treatment_group,
data=data_summary, mean) # Create a matrix that contains our
fixed effects
dummydat$iq <- 0 # Set IQ to 0, This means we are considering
effects of interest for average levels of IQ.
dummy=model.matrix( ~ age_group * treatment_group + iq,
data=dummydat) # Set up dummy codes to compare groups.

# Now code the contrasts using subtraction
contrasts <- rbind(
"Accuracy after training compared to control is different in age
group 7-10,
than in age group 11-14"=
((dummy[dummydat$treatment_group == "Training"
& dummydat$age_group == "7-10",]) -
(dummy[dummydat$treatment_group=="Control"
& dummydat$age_group == "7-10",])) -
((dummy[dummydat$treatment_group == "Training"
& dummydat$age_group == "11-14",]) -
(dummy[dummydat$treatment_group=="Control"
& dummydat$age_group == "11-14",])),
"Accuracy after training compared to control is different in age
group 7-10,
than in age group 15-18"=
((dummy[dummydat$treatment_group == "Training"
& dummydat$age_group == "7-10",]) -
(dummy[dummydat$treatment_group=="Control"
& dummydat$age_group == "7-10",])) -
((dummy[dummydat$treatment_group == "Training"
& dummydat$age_group == "15-18",]) -
(dummy[dummydat$treatment_group=="Control"
& dummydat$age_group == "15-18",])),
```

```

"Accuracy after training compared to control is different in
age group 11-14,
than in age group 15-18"=
((dummy[dummydat$treatment_group == "Training"
  & dummydat$age_group == "11-14",]) -
 (dummy[dummydat$treatment_group=="Control"
  & dummydat$age_group == "11-14",])) -
((dummy[dummydat$treatment_group == "Training"
  & dummydat$age_group == "15-18",]) -
 (dummy[dummydat$treatment_group=="Control"
  & dummydat$age_group == "15-18",]))

summary(glht(model, contrasts)) # Inspect results
#####

```