Investigating the mechanisms underlying fixation durations during the first year of life: a computational account

http://bbktheses.da.ulcc.ac.uk/148/

Version: Full Version

Citation: Saez de Urabain, Irati R. (2015) Investigating the mechanisms underlying fixation durations during the first year of life: a computational account. PhD thesis, Birkbeck, University of London.

Deposit guide

Contact: email

INVESTIGATING THE MECHANISMS UNDERLYING FIXATION DURATIONS DURING THE FIRST YEAR OF LIFE: A COMPUTATIONAL ACCOUNT

Irati R. Saez de Urabain

A thesis submitted for the degree of

Doctor of Philosophy

Department of Psychological Sciences

Birkbeck College

University of London

April 2015

Para aquellos que nos dejaron y los que llegaron,

para la abuela Asun y el pequeño Unax

Originality statement

"I, Irati R. Saez de Urabain, declare that the work submitted in this thesis is my own."

Signed: _____

April 15, 2015

Abstract

Infants' eye-movements provide a window onto the development of cognitive functions over the first years of life. Despite considerable advances in the past decade, studying the mechanisms underlying infant fixation duration and saccadic control remains a challenge due to practical and technical constraints in infant testing. This thesis addresses these issues and investigates infant oculomotor control by presenting novel software and methods for dealing with low-quality infant data (GraFIX), a series of behavioural studies involving novel gaze-contingent and scene-viewing paradigms, and computational modelling of fixation timing throughout development. In a cross-sectional study and two longitudinal studies, participants were eye-tracked while viewing dynamic and static complex scenes, and performed gap-overlap and double-step paradigms. Fixation data from these studies were modelled in a number of simulation studies with the CRISP model of fixation durations in adults in scene viewing. Empirical results showed how fixation durations decreased with age for all viewing conditions but at different rates. Individual differences between long- and short-lookers were found across visits and viewing conditions, with static images being the most stable viewing condition. Modelling results confirmed the CRISP theoretical framework's applicability to infant data and highlighted the influence of both cognitive processing and the developmental state of the visuo-motor system on fixation durations during the first few months of life. More specifically, while the present work suggests that infant fixation durations reflect on-line perceptual and cognitive activity similarly to adults, the individual developmental state of the visuo-motor system still affects this relationship until 10 months of age. Furthermore, results suggested that infants are already able to program saccades in two stages at 3.5 months: (1) an initial labile stage subject to cancellation and (2) a subsequent non-labile stage that cannot be cancelled. The length of the non-labile stage decreased relative to the labile stage especially from 3.5 to 5 months, indicating a greater ability to cancel saccade programs as infants grew older. In summary, the present work provides unprecedented insights into the development of fixation durations and saccadic control during the first year of life and demonstrates the benefits of mixing behavioural and computational approaches to investigate methodologically challenging research topics such as oculomotor control in infancy.

Acknowledgments

My PhD years at the CBCD were enriched by amazing colleagues willing to share their time, knowledge and experiences with me. Carina, always ready to help whatever the problem is; Kristen, who taught me all about handling young infants and caregivers during our joint testing sessions and showed me that it is possible to be a great mother and researcher at the same time(!); Katarina, who patiently taught me how to build a silicon hat for our attempts to build a baby head-mounted eye-tracker; Kostas, who made the whole research experience more enjoyable over occasional beers and great conversations; Sinead, whose joy and professionalism makes the Babylab a better place; Annette, a truly inspirational woman and amazing researcher whose life could perfectly be on a film; and of course Marian and Leslie, always there when you need them!

These years and also my life in London would not be the same without my dear friend, Bienam. We met when we were 8 and our paths kept crossing until it became obvious that we were predestined to be great friends! Our weekly meetings over beers, teas, and guarrindongadas are essential to survive the week, I am not sure what I would do without you!

And of course, the London crew: Marta, Jon, Filipe, Parag, Alex, Alba, Chris, Ellie, David, Juana,… thanks a million for all the great moments we have shared (and will share) in this amazing city!

I would not be where I am today without the support of my family, who believed in me since the day I was born. My parents, Luis Angel and Maria Jose; aunts, Txus and Ana; and brother, Ireber were always there when I needed them. They have always accepted and supported my (difficult) choice to build my life abroad, and for this I am incredibly grateful.

This thesis is dedicated to two very special people: my granny Asun, who couldn't stay to see what my life has become. I would not be the person that I am today without your example and love, you will always live in my heart. And my little son, Unax, who arrived to this world in the middle of Chapter 6 to fill our lives with joy and unconditional love!

Above all, I would like to thank you, Mats, my best friend and partner in crime, for filling my days with joy, love and positive thinking. No matter what with you the glass seems always full and the future bright! You make my life better at so many levels, and for this and many other things I will always love you.

List of publications

Publications in peer-reviewed journals

- **Saez de Urabain, I.R.**, Nuthmann, A., Johnson, M.H., Smith, T.J., (submitted). Disentangling the mechanisms underlying infant fixation durations scene perception: A computational account.

- Smith, T.J., **Saez de Urabain, I.R**., (in press) Eye-tracking. *The Cambridge encyclopaedia of child development*. Cambridge University Press.

- **Saez de Urabain, I.R.** , Johnson, M.H., Smith, T.J., (2014) *GraFIX: A semiautomatic approach for parsing low- and high-quality eye-tracking data.* Behavior Research Methods. doi:10.3758/s13428-014-0456-0

Presentations and conference proceedings

- **Saez de Urabain, I.R.**, Nuthmann, A., Johnson, M.H., Smith, T.J., (March 2015). Modelling longitudinal trajectories of fixation durations in infants. Talk presented at the Society for Research in Child Development (SRCD) Biennial Meeting, Philadelphia.

- **Saez de Urabain, I. R.**, Johnson, M.H. & Smith, T.J. (July 2014). Developing new approaches to deal with data quality in infant eye-tracking studies. Poster presented at the 19th Biennial International Conference on Infant Studies (ICIS), Berlin, Germany.

- **Saez de Urabain, I. R.**, Johnson, M.H. & Smith, T.J. (July 2014). Development of eye-movement control in complex scenes: A longitudinal study. Poster to be presented at the 19th Biennial International Conference on Infant Studies (ICIS), Berlin, Germany.

- **Saez de Urabain, I. R.**, Johnson, M.H. & Smith, T.J. (January 2014). Development of saccadic control in infancy in dynamic and static complex scenes: A longitudinal study. Poster presented at the 5th Annual CEU Conference on Cognitive Development, Budapest, Hungary.

- **Saez de Urabain, I. R.**, Johnson, M.H. & Smith, T.J. (August 2013). GraFIX: Developing a novel semi-automatic approach for detecting fixation durations in low quality data from infants and adults. Poster presented at the 37th European Conference on Visual Perception, Bremen, Germany(Link). Abstract published in Perception special issue.

- **Saez de Urabain, I. R.**, Nuthmann , A., Johnson, M.H. & Smith, T.J. (August 2013). Modelling fixation durations in infancy: a simulation study with 6 month-olds. Talk presented at the 17th European Conference of Eye Movements, Lund, Sweden. Abstract published in K.Holmqvist, F. Mulvey & R. Johansson (Eds.), Book of Abstracts of the 17th European Conference on Eye Movements, 11-16 August 2013, in Lund, Sweden. Journal of Eye Movement Research, 6(3), 53

- **Saez de Urabain, I. R.**, Johnson, M.H. & Smith, T.J. (May 2013). Investigating infants' inhibitory control and fixation durations in complex naturalistic and non-naturalistic scenes. Poster presented at the Vision Sciences Society 13th Annual meeting , Naples, Florida, USA. Abstract published at the Journal of Vision.

- **Saez de Urabain, I. R.**, Johnson, M.H. & Smith, T.J. (January 2013). Developmental changes in fixation durations in naturalistic and non-naturalistic complex scenes: a cross-sectional study. Poster presented at the 4th Annual CEU Conference on Cognitive Development, Budapest, Hungary.

- **Saez de Urabain, I. R.** , Smith, T.J. & Johnson, M.H. (September 2012). An analysis of the early saccadic system through a 'double step' saccade paradigm and an adaptation of the MOBS threshold method. Poster presented at the British Psychological Society Developmental Section Conference, Glasgow, Scotland.

Workshop preparation

- Designing eye-tracking experiments to measure infant and child behaviour (January 2014). Workshop co-organized with Tobii for the 5th Annual CEU Conference on Cognitive Development, Budapest, Hungary.

Media Coverage

- Wired magazine, March 2013 issue. Meet London's Babylab, where scientists experiment on babies' brains.

Table of Contents

Table of Figures

Table of Tables

List of abbreviations

ICC – Intra-class correlation coefficient

AOI – Area of interest

ANOVA – Analysis of variances

ADHD – Attention deficit hyperactivity disorder

RMS – The root mean square of inter-sample distances

MOBS – Modified binary search thresholding algorithm

MCMC – Markov chain Monte Carlo procedure

LMM – Linear mixed model

VOR – Vestibulo-ocular reflex

OKR – Optokinetic reflex

Thesis overview

Chapter 1 provides the general literature background for this thesis, discussing past research on vision, visual orienting and visual attention. It particularly focuses on reviewing previous empirical work and theoretical and computational models investigating visual attention and the mechanisms underlying oculomotor control and fixation durations in infants and adults.

Chapter 2 introduces the methods, experimental tools, and analytical techniques used in this thesis.

Chapter 3 presents GraFIX, a new method and software to detect fixations in low and high quality eye-tracking data. Furthermore, this chapter includes a validation analysis comparing this method to previous approaches.

Chapter 4 includes a cross-sectional study for which I analysed fixation durations and saccadic control in 6-month-old infants and adults that were presented with a battery of naturalistic and semi-naturalistic videos, and performed the gap-overlap and double-step paradigms.

Chapter 5 includes three simulation studies investigating the mechanisms underlying fixation durations and saccadic control in 6-month-olds. For this purpose the fixation data from Chapter 4 were modelled using the CRISP model of fixation durations in scene viewing (Nuthmann, Smith, Engbert, & Henderson, 2010).

Chapter 6 includes two longitudinal studies and a study with adults that aim to investigate the development of fixation durations and saccadic control during the first year of life. For this purpose participants were presented with dynamic stimuli (naturalistic and abstract videos) and static images, and performed the gap-overlap paradigm.

Chapter 7 includes three simulation studies that investigate the development of the mechanisms underlying fixation durations and saccadic control over the first year of life. For this purpose the CRISP model (Nuthmann et al., 2010) was used to model fixation data from Chapter 6.

Chapter 8 includes a critical discussion of the findings of my experiments from a broader perspective. Furthermore, it examines the limitations of this work and how these might be addressed and overcome in the future.

Chapter 1: Introduction

1.1 General introduction

From the first day of their life infants use their eyes to sample the world around and select the most relevant information in the environment, trying to disentangle all the mysteries that surround them. Even though at a first glance the process of moving our eyes from one point to another may seem a simple operation, it involves a number of attentional and cognitive processes, besides a very complex neural machinery that we have just started to understand.

Developmental psychology has widely investigated the development of attentional and cognitive processes -such as memory, object perception or language acquisition- through the analysis of different aspects related to visual attention. For instance, in the infant literature it is very common to find *habituation* and *familiarization* paradigms, where the subject is repeatedly or continuously exposed to an event before he/she is presented with an unfamiliar object. The mean looking times to these objects is then measured and compared. Along with these studies, others have applied *preferential looking* paradigms, where infants are presented with two simultaneous objects or patterns and the duration and location of each look is measured and analysed. All these looking paradigms have been extensively used to investigate cognitive development and the psychological processes going on inside the infant's brain.

Questions related to how do infants modulate and control eye-movements are, however, poorly understood: What is in a look? Why, where and when do they move their eyes? What are the inner mechanisms underlying oculomotor control in infancy? And how do they develop over the first year of life?

During active visual sampling eyes can remain stable at a visual location during what is known as a fixation, or can perform fast ballistic movements called saccadic eye-movements. It is during a fixation when visual encoding occurs while during a saccade visual sensitivity is suppressed (Matin, 1974; Ross, Morrone, Goldberg, & Burr, 2001). Interestingly, saccades are likely to be the most practiced motor skill throughout the first year of life. In fact, by 4 months infants have already made over 3 million eye-movements (M. H. Johnson, 2003), thus one could expect them to increase in efficiency rapidly during the first months of life.

Whilst most past research on eye-movements has been focused on where fixations land (e.g., Buswell, 1935; Itti & Koch, 2001; Loftus & Mackworth, 1978), there is a growing body of research investigating the durations of those fixations, particularly in human adults. For instance, a number of adult studies investigating fixation durations in scene perception have already shown how these types of eye-movements can be affected by factors such as the task (e.g., memorization vs. visual search; e.g., Henderson, Weeks Jr, & Hollingworth, 1999; Nuthmann et. al, 2010; Võ & Henderson, 2009), the visual characteristics of the stimulus (e.g., luminance, image degradation; e.g., Henderson, Nuthmann, & Luke, 2013; Loftus, 1985), the semantics of the scene (e.g., Henderson et al., 1999; Loftus & Mackworth, 1978; Võ & Henderson, 2009), or familiarity (e.g., Althoff & Cohen, 1999). These findings manifest the influence of perceptual and cognitive processing in adults' fixations and demonstrate how the exhaustive measurement and analysis of fixation durations can give rise to new metrics that can help answering questions such as those related to the assessment of attention and information processing in more ecologically valid settings.

Nevertheless, as a result of the numerous technical and practical constraints that testing infants involve (e.g., high degree of movement, lower data quality, not attending to the experimenter's instructions), infant fixation durations and saccadic control are much less understood. In the present thesis I will first overcome some of the issues related to gathering and analysing eye-tracking data from young infants, in order to then investigate the development of fixation durations and saccadic control during the first year of life.

1.2 Aims and objectives

In the present project I combine a series of behavioural cross-sectional and longitudinal studies with computational modelling in order to investigate the development of fixation durations and the mechanisms underlying fixation durations and saccadic control during the first year of life.

The specific aims for this project are:

- to define a precise protocol for gathering usable eye-tracking data from infants and develop a method to detect and analyse fixation durations in low and high quality eye-tracking data from infants as well as adults (Chapters 2 and 3);
- to investigate the factors influencing fixation durations and saccadic control in 6-month-olds and adults (Chapter 4);

- to investigate the unexplored mechanisms underlying fixation durations and saccadic control in infancy by determining the generalizability and potential limitations of extending the CRISP theoretical framework and computational model of fixation durations in adult scene-viewing (Nuthmann et al., 2010) to infants (Chapter 5);

- to test specific developmental theories of oculomotor control by applying the CRISP model to data from 6-month-old infants. More specifically, I will investigate the extent to which fixation durations are influenced by the developmental state of the visuo-motor system and by visual and cognitive processing (Chapter 5);

- to investigate the development of fixation durations and saccadic control during the first year of life through a series of longitudinal studies (Chapter 6);

- to explore the mechanisms underlying the development of fixation durations and saccadic control during the first year of life applying the CRISP model to longitudinal data (Chapter 7).

The remainder of this introductory chapter reviews the background literature on eye-movements from infants and adults as well as previous empirical work and theoretical models investigating the mechanisms underlying oculomotor control and fixation durations. This chapter then discusses current computational models of visual attention on infants and adults.

1.3 Vision, visual orienting, and visual attention

The concept of "attention" has been the focus of vast research. In fact, the number of definitions it has acquired over time are multiple and may slightly vary from one domain of study to another. For instance, Findlay and Gilchrist (2003) defined it as the preferential processing of some items to the detriment of others. Mesulam (1981) presented a functional approach to attention where he suggested three general attentional processes: the selection of a target, the engagement of attention, and the controls that allow us to maintain and shift attention as required. The present thesis will take this functional view of attention.

Attentional processes can be studied using a variety of methods such as behavioural approaches (analysis of eye-movements, analysis of facial expressions, reaction times, quantitative aspects of performance on a task, or counting the number of times the subject looks away from a central stimuli), physiological measures (heart rate, respiratory sinus arrhythmia, adrenocortical activity, or neurotransmitter activity), or neurological methods (ERP

studies, analysis of the effects of early cortical damage, or the study of developmental disorders of genetic origin). Additionally, since attention is the result of the interaction between a subject and its environment, it is essential to incorporate the context in which attention is observed into the equation (Ruff & Rothbart, 1996). For instance, it is important to differentiate the attentional responses measured in lab settings compared to those measured in naturalistic settings (Mesulam, 1981), or the attentional capabilities of a 6-month-old compared to a 12-month-old.

In the realm of cognitive psychology, the term visual attention is commonly associated with visual selectivity (Findlay & Gilchrist, 2003). Selectivity and the processes underlying it are crucial in order to manage the vast amount of information and stimulation that is available in the environment (Ruff & Rothbart, 1996). Furthermore, attentional selection can be made in two separate ways. *Overt attention* is the term that describes attention resulting from the physical movement of the sensory organs towards the source of the attended sensory information. In terms of visual attention this involves rotating the eyes, head or torso. On the other hand, adults are also able of shifting their attention *covertly*, without moving their eyes or other sensory receptors (Findlay & Gilchrist, 2003; M. H. Johnson, 2011). Even though there are a number of measures that can indicate if an infant is attending to a particular object or location (e.g., reaching and other motor responses), this thesis will only be concerned with overt shifts of attention where visual selectivity involves redirecting the gaze to a new location in the visual field, or in other words, with visual orienting and the mechanisms underlying it.

Visual orienting is the primary method for gathering information from the environment during the first year of life. From the moment they are born infants are able to move their eyes in order to select the most relevant information from their surroundings for further study and learning. By simply analysing infants' gaze researchers are thus able to learn a great deal about cognitive development. In fact, the vast majority of the literature studying mental processes in infancy involves measures related to visual attention and selectivity such as habituation, familiarization, or preferential looking paradigms. Additionally, the study of visual orienting in infants also allows the examination of the effects of brain development on the integration between sensory input and motor output (M. H. Johnson, 2011).

Despite the substantial literature on visual orienting in infants, the mechanisms underlying eye-movement control are still poorly understood. Surely, most of the literature on the neural basis

of saccades comes from neuropsychological and neuroimaging studies of human adults or single-cell recordings and lesion studies in non-human primates (for a review see Andersen, Batista, Snyder, Buneo, & Cohen, 2000). In this section I will review the relevant literature on vision and visual orienting from adults and infants.

1.3.1 Vision: The oculomotor system

The repertory of eye-movements that humans are capable of making is relatively small and well defined. Interestingly, we share the most consistent pattern of eye-movements -commonly referred as the 'saccade and fixate' strategy- with a number of other species coming from various evolutionary backgrounds such as fish, flies, or birds (Land, 1999). The fovea is the small region of high acuity in the eye (with a diameter of 0.3° - 2°), and the location at which visual activity is centred (Findlay & Gilchrist, 2003; Land, 2009). The accuracy of vision is more degraded the further away from the fovea one moves.

The angle of eccentricity refers to the angle between the visual axis where the fovea is directed and the peripheral location under consideration. Generally, the foveal region in adults is considered to have an angle of eccentricity of 1°, parafoveal region goes from 1° to 5°, and peripheral vision covers the rest of the visual field (Findlay & Gilchrist, 2003). The most important function of peripheral vision is to provide the appropriate information for subsequent orienting movements and foveal recognition (Findlay & Gilchrist, 2003). In infants, it has been a topic of debate whether the changes in visual abilities are due to limitations in the periphery – the structure of the eyes, lens or eye muscles-, or rather changes in the neural structures related to vision processing (M. H. Johnson, 2011). Nevertheless, in a study from Banks and Shannon (1993) comparing the morphology of neonatal photoreceptors and optics to that of adults, it was concluded that the development of the central nervous system pathway is an important contributing factor in the development of vision, even though much of the spatial and chromatic deficits found in neonates can be explained by optical and receptor immaturities. In this subsection I will review the current literature on the oculomotor system, eye-movements and oculomotor control in adults and infants.

1.3.1.1 Eye physiology, visual acuity and binocularity in adults and infants

The human eyeball is supported and moved by six muscles: four rectus muscles (lateral, medial, superior, and inferior) that rotate the eyeball in the horizontal and vertical planes, and the two oblique muscles (superior and inferior) that rotate the eyeball around the visual axis, i.e. torsional rotation. These six muscles always work together, which means that any type of eye-movement needs the action of all the eye muscles and a complex pattern of activation from the oculomotor nuclei.

The specialized foveal region of the retina contains the highest density of photoreceptors (cones), besides other visual cells namely bipolar, horizontal, amacrine and ganglion cells. At the very centre of the foveal region there is a rod-free region only composed by cones, that has a diameter of less than one degree (Hirsch & Curcio, 1989).

The photoreceptors are responsible for initiating the neural processing of the visual signal, sending it through the retina to the ganglion cells. It subsequently travels along the optic nerve to the visual cortex, through the lateral geniculate nucleus of the thalamus. Even though in primates this is considered to be the primary projection pathway, a number of other pathways - such as the one that goes to the superior colliculus (a mid-brain region responsible for saccade programming)- emerge at the stage of the optic tract.

Human newborns have poor eyesight which can be attributed to: poor contrast sensitivity – defined as the minimum luminance required for detecting a visual target– visual acuity –defined as the smallest detail that can be detected– and chromatic discrimination –defined as the ability to distinguish targets based on their wavelength composition– (Banks & Shannon, 1993). Interestingly, the eye suffers a significant growth from birth to adolescence, particularly during the first year of life. This growth includes the distance from the cornea to the retina, the pupil, or the fovea, which keeps undergoing morphological changes until at least 4 years of age (Yuodelis & Hendrickson, 1986). For instance, the diameter of the rod-free zone decreases from about 5.4° at birth to 1° at adulthood (Banks & Shannon, 1993; Hirsch & Curcio, 1989).

Visual acuity does not reach adult-like levels until at least the third year of life (Atkinson, 2000). In fact, newborns are believed to have a fixed depth of focus (Salapatek, Bechtold, & Bushnell, 1976), and in many cases, they even present some degree of astigmatism (Atkinson, 2000). However, it is during the first 3 to 4 months that the biggest changes in visual acuity, contrast

sensitivity, and focusing ability (accommodation) occur. This means that during the first months of life, infants may have problems accommodating as a function of target distance, and hence, they may not distinguish objects that are farther away than a certain distance (Salapatek et al., 1976).

Most primates, including humans, have a binocular central visual field that allows the integration of the information between the two eyes. This integration is thought to take place at the primary visual cortex, being the functional and anatomical structures observed in layer 4 important for this endeavour (M. H. Johnson, 2011). In infants, it has been argued that binocular vision does not develop until the end of the fourth month (Held, 1993), even though evidence for binocularity has been reported from around 3 months (for a review see Braddick & Atkinson, 2011).

1.3.1.2 Classification of eye-movements

In order to project the light reflected from an object of interest on to the fovea, it is essential to move our eyes. While our eyes are being relocated from one point to the next –during a *saccade*- we suffer a combination of blur and active suppression that makes visual perception difficult (Matin, 1974; Ross et al., 2001). Thus, it is essential to move our eyes as fast as possible, reaching speeds of 700°/sec in adulthood (Carpenter, 1988). It is only during the stationary eye-movements between saccades –called *fixations*- when visual information is processed. This is because the process of photoreception is slow, taking about 20 ms for a cone to respond to the light changes in the environment (Friedburg, Allen, Mason, & Lamb, 2004). In this subsection, I will provide a brief account of the repertory of the individual eye-movements available to humans.

1.3.1.2.1 The stabilizing mechanisms: Body, head and eye-movements

Target selection occurs as a consequence of the combination of body, head and eye-movements. Nevertheless, frequently the selected target or our own body and head are moving in space so that for a fixation to be maintained powerful compensatory mechanisms are required. Visual stabilization is promoted by two primary systems, which are involuntary and automatic: *the vesibulo-ocular* (VOR) and *optikinetic reflexes* (OKR). In VOR, head rotation is compensated by generating an eye-movement equal and opposite to this rotation. On the other hand, in OKR, motion of the world relative to the eye is compensated by moving the eyes in the same direction as the external motion (for more details see Carpenter, 1988). The present

project will not focus on these mechanisms and assumes that the head and torso are still relative to a static display.

1.3.1.2.2 Saccadic eye-movements

Saccades are the ballistic eye-movements that we make in order to relocate our gaze from one visual target to the next. Even though they can be voluntary, most saccades are performed below the level of conscious awareness (Findlay & Gilchrist, 2003). Further, they can be triggered by external events –such as the appearance of a novel stimulus- but also by internal instructions. The typical reaction times for saccades -or saccade latencies- that are externally triggered range between 150 and 200 ms for adults (Land, 2009). Nevertheless, it is also possible to find reaction times ranging from 90 to 130 ms that are triggered by some particular attentional events. This type of eye-movements are known as *express saccades* (Land, 2009; see also section 1.3.2.1).

 Saccade duration mainly depends on the saccadic amplitude (Carpenter, 1988; Hainline, Turkel, Abramov, Lemerise, & Harris, 1984), even though there may be other factors -such as age, drowsiness, or the use of certain drugs- influencing it (Findlay & Gilchrist, 2003).

Saccadic suppression is the phenomenon occurring during a saccade where visibility is reduced. In adults, it starts around 50 ms before the saccade is initiated and ends around 100 ms after the saccade has begun. This suppression is most evident on the magnocellular pathway (Burr, Morrone, & Ross, 1994) and helps preventing undesired activations as a consequence of the image motion generated by the saccade.

The very short involuntary saccades are known as *microsaccades* and typically occur during prolonged visual fixations. The role of this type of eye-movements has been the focus of extensive research, where it has been argued that microsaccades correct displacements in the eye position, and relate to the control of binocular fixation disparity and attentional shifts (for a review see Rolfs, 2009).

Infants are able to perform the 'saccade and fixate' strategy from their first day of life. Nevertheless, the saccadic dynamics do not reach adult-like levels until few months after birth (Atkinson, 2000). For instance, 1-month-olds have been reported to perform hypotonic saccades, or in other words, saccades that fall short of their targets after certain stimulus conditions (Aslin & Salapatek, 1975). Furthermore, the ability to perform a saccade in the

presence of distractors or multiple targets, and the latencies of these saccades change considerably with age and increase in efficiency rapidly during the first months of life.

1.3.1.2.3 Fixations

Fixations are defined as the periods between saccades when the eyes are stable and visual encoding occurs. Even though eyes are assumed to remain still during these periods, close inspection of the eyes demonstrates how they make continuous miniature movements (Ditchburn, 1973). *Drifts* are the slow motion movements that occur between microsaccades, while *tremor* is the small oscillation superimposed on drifts (Rolfs, 2009).

Measuring and reporting fixation durations is common practice in experimental psychology, and can be used as an index of online cognitive processing demands (e.g., Martinez-Conde, Macknik, & Hubel, 2004; Martinez-Conde, 2005; Nuthmann et al., 2010; Smith & Henderson, 2009; Tatler, Gilchrist, & Land, 2005). There is, in fact, a growing body of research that associates fixation durations with cognitive processes such as attention, information processing, memory and anticipation (Castelhano & Henderson, 2008; Kowler, 2011; Malcolm & Henderson, 2010; Rayner, Smith, Malcolm, & Henderson, 2009; Richardson, Dale, & Spivey, 2007). The study of fixation durations is becoming increasingly important when investigating populations unable to follow the experimenter's instructions such as infants (Bronson, 1994; Colombo & Cheatham, 2006; Frick, Colombo, & Saxon, 1999; Harris, Hainline, Abramov, Lemerise, & Camenzuli, 1988; Hunnius & Geuze, 2004b; Hunter & Richards, 2011; Richards & Holley, 1999) or other primates (Berg, Boehnke, Marino, Munoz, & Itti, 2009; Kano & Tomonaga, 2011a, 2011b). The past literature on fixations on human adults and infants will be reviewed in the sections 1.4 and 1.5 respectively.

1.3.1.2.4 Smooth pursuit eye-movements

We perform *smooth pursuit* eye-movements when tracking a small object moving relative to a static visual field, while OKR operates on large areas of the image (Land, 2009). Smooth pursuit is possible when the velocity of the object is under a threshold (15°/sec for adults), in which case the eyes perform smooth movements in order to keep the object of interest at the centre of the fovea. In case the velocity of the tracked object exceeds the threshold, the eyes will also perform saccades in order to compensate the object's motion (Land, 2009).

Smooth pursuit eye-movements develop during the first 6 months of life (Aslin, 1981; Richards & Holley, 1999). Infants younger than 1 month present a very poor tracking of visual stimuli, only possible at relatively slow stimulus speeds, that involves saccadic eye-movements rather than smooth pursuit eye-movements (e.g., Aslin, 1981). On the other hand, OKN eye-movements are present from birth, thus it can be argued that at least the kinetic mechanisms of smooth pursuit are reasonably mature in neonates (Atkinson, 2000). The present project will not focus on this type of eye-movements. For this reason the dynamic stimuli used in the experiments from Chapters 4 and 6 avoided elements eliciting this type of eye-movement.

1.3.1.2.5 Vergence eye-movements

The function of vergence eye-movements is to adjust the angle between the eyes to different distances in order to ensure that the image of an attended object is projected on to the corresponding parts of the left and right retina. These movements are smooth and typically slow, with a latency of about 160 ms in adults (Land, 2009). In infants, vergence eye-movements do not develop until 10 to 16 weeks of age (Aslin, 1993). The present project will not be focused on this type of eye-movement.

1.3.1.3 Mechanisms underlying saccadic control in adults

The oculomotor system is thought to work at three main levels: the cerebral cortex, responsible for voluntary eye-movement control; midbrain structures such as the superior colliculus, responsible for saccade generation and input-driven exogenous eye-movements; and the nuclei of the brain stem, responsible for connecting the motor output that controls eye-muscles (Carpenter, 1988).

The visual system is composed of different types of nerve cells. In the primate visual system, magnocellular (M) and parvocellular (P) cells separate into two distinct layers at the level of the lateral geniculate nucleus, originating magnocellular and parvocellolar parallel systems (Schiller, Logothetis, & Charles, 1991). Nevertheless, this separation of cells is also found in the retina, even though there is evidence suggesting that both streams converge in certain cortical areas (Ferrera, Nealey, & Maunsell, 1992).

Ungerleider and Mishkin (1982) proposed a simplified account where two primary pathways were defined: The dorsal stream, which goes from the occipital to parietal cortex, and the ventral stream that runs from occipital to temporal cortex. Based on data from lesion studies,

they suggested that the ventral stream is responsible for visual identification and recognition of objects (*what*), while the dorsal stream takes care of visuospatial awareness or the localization of those objects (*where*). Even though it has been suggested that the M and P pathways map onto these cortical routes (Livingstone & Hubel, 1987), later research has probed these claims controversial (Merigan & Maunsell, 1993).

Besides the main pathways discussed above, there are a number of other physiological pathways, such as the superior colliculus pathway, leaving the retina and projecting to different brain regions. Even though traditionally these secondary subcortical visual pathways have been thought to be responsible for 'reflex eye-movements', recent findings have demonstrated that in fact the fastest eye-movements make use of cortical pathways (Miles, 1998).

Physiological pathways involved in saccadic eye-movements have two main characteristics: First, the stimulation of the area will produce orienting movements of the eyes and second, cells in the area will discharge prior to the production of a saccadic eye-movement (Schall, 1991). Figure 1-1 displays a schematic diagram with the main areas involved in saccadic eye-movement generation (from Schall, 1995).
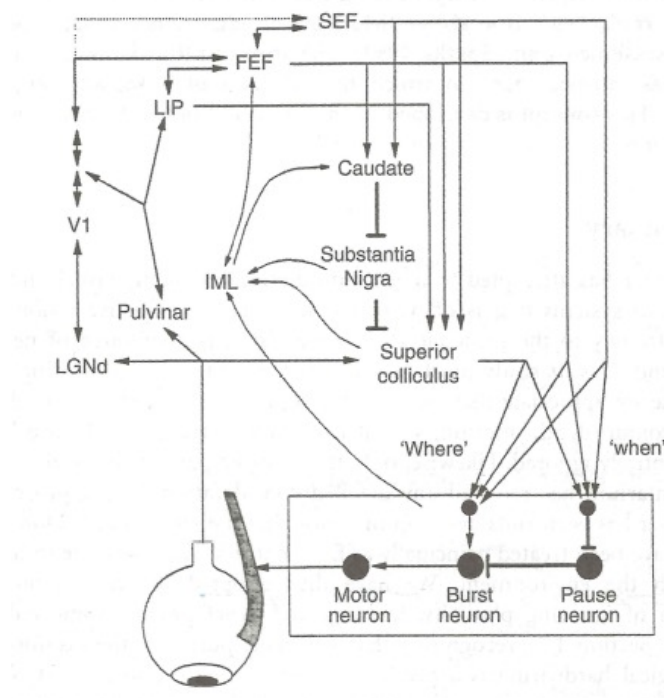
Figure 1-1 Oculomotor output pathways involved in generating saccadic eye-movements. V1, visual cortical area; LIP, lateral intra parietal area; FEF, frontal eye fields; SEF, supplementary eye fileds; IML, internal medullary lamina; LGNd, dorsal lateral geniculate nucleus. From (Schall, 1995).

Saccades can be elicited by electrical stimulation on cortical areas such as the lateral intraparietal area (LIP) in the posterior parietal cortex and the frontal eye fields (FEF) of the pre-motor frontal cortex. Even though lesion studies in primates have suggested that no area is essential for saccadic generation (Schiller, 1998), the removal of the superior colliculus and the frontal eye fields supresses the ability to generate saccades (Schiller, 1985) . Other studies have also demonstrated that lesions to both the superior colliculus and the occipital cortex (V1) elicit the same kind of difficulties (Mohler & Wurtz, 1977).

1.3.1.4 Mechanisms underlying saccadic control in infants: The development of visual pathways

Even though the technical advances in the last decades -such as PET or fMRI techniques- have allowed the direct investigation of cortical activity, much of what we know about human saccadic system comes from patients with damage to cortical areas of the brain, or lesion studies on animals (Findlay & Gilchrist, 2003). Nevertheless, an essential source of information comes from developmental cognitive neuroscience. It is assumed that certain visual areas and pathways develop and mature at different developmental stages due to an interaction between genetically prewired programming and environmentally dependent learning (Atkinson, 2000; M. H. Johnson, 2011). The emergence of parallel behavioural changes is then traced to the development of these neural structures. In the current subsection I will review the neuroanatomical models on the development of visual orienting and the literature on the mechanisms underlying saccadic control in infants.

1.3.1.4.1 Bronson's model

Few neurophysiological models have attempted to explain the substantial development that the visuo-motor system experiences during the first postnatal months. Bronson's (1974) account, which was supported by a number of electrophysiological, neuroanatomical and behavioural studies, claimed that eye-movement control is primarily driven through the subcortical retinotectal visual pathway, and that it is only from 2 to 3 months of postnatal age when the cortical primary visual pathway starts playing a role in infant visually guided behaviour. Nevertheless, later research has demonstrated that there is some -though limited- cortical activity in newborn infants. In addition, it seems more likely that the transition from subcortical to cortical activity occurs gradually rather than in a single developmental stage (M. H. Johnson, 2011). Bronson's view was later extended by other researchers in order to account for new

evidence from neurophysiological research on monkeys and neuropsychological research with human adults revealing that there are multiple pathways involved in oculomotor control and attention shifts in the primate brain (Atkinson, 2000; M. H. Johnson, 1990, 2011).

1.3.1.4.2 Johnson's model

M. H. Johnson (1990, 2011) presented a model grounded on Schiller's neuroanatomical model of saccadic control for primates (for more details see Schiller, 1985, 1998). Johnson proposed a developmental sequence for the onset for the different visual pathways traced to the onset of visual orienting components, a perspective also known as the maturational view to the development of human brain function. In other words, he proposed that the characteristics of visual orienting at a particular age are determined by which of the pathways are functional, and that this directly depends on the maturational state of the primary visual cortex.
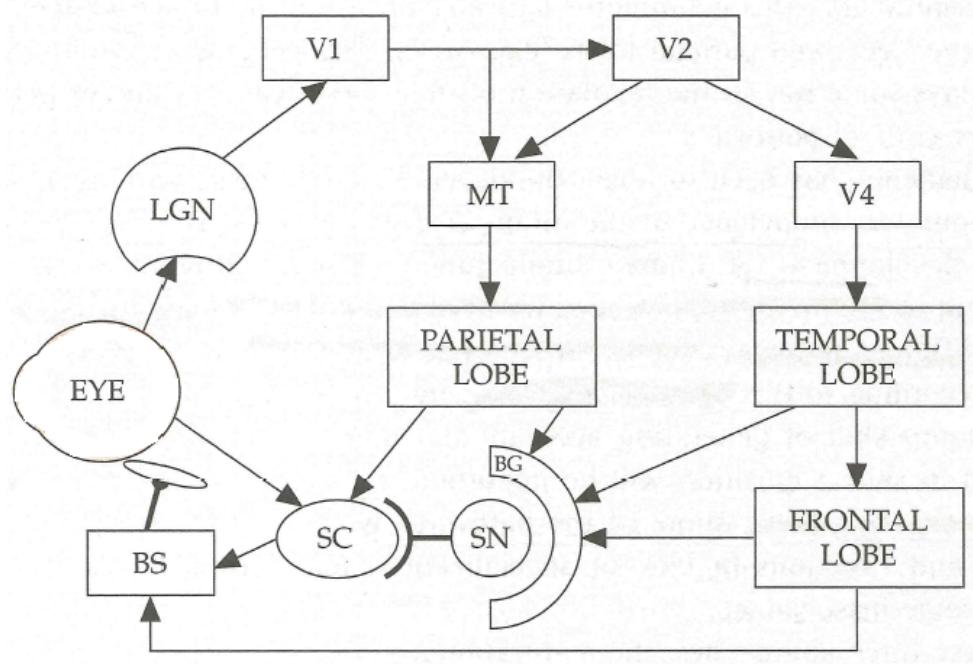
Figure 1-2 This diagram shows some of the primary neural pathways and structures involved in visual orienting and attention. BS, brain stem; LGN, lateral geniculate nucleus; V1, V2 and V4, visual cortical areas; MT, middle temporal area; SC, superior colliculus; SN, substantia nigra; BG, basal ganglia; FEF, frontal eye fields. From (M. H. Johnson, 2011)

Johnson discussed four brain pathways (see Figure 1-2) previously proposed by Schiller (1985):

(1) The pathway from the eye to the superior colliculus, involved in the generation of rapid input-driven eye-movements; (2) a cortical pathway that goes to the superior colliculus from the

primary visual cortex (V1) through the middle temporal area (MT), thought to be involved in motion detection and the smooth tracking of moving stimuli; (3) a cortical pathway that goes from V1 to other visual areas through the frontal eye fields (FEF), important for more complex forms of eye-movement planning such as anticipatory saccades or learning sequences of scanning patterns; and (4) a diffuse cortical projection to the superior colliculus via the basal ganglia and substantia nigra, thought to be involved in the regulation of the colliculus.

Johnson (1990) argued for the following developmental sequence of onset: First (1), followed by (4), followed by (3), and finally (2). These predictions were initially made based on previous literature on developmental neuroanatomy and then backed with behavioural experiments.

According to this account newborns' eye-movements are mainly driven by the subcortical pathway involving the superior colliculus, thought to be responsible for fast and input driven exogenous saccades. Evidence for this are for instance the studies on smooth pursuit in young infants, showing that newborns track moving objects by performing a series of short saccadic eye-movements that tend to land behind the movement of the stimulus (Aslin, 1981). Furthermore, newborn infants orient toward stimuli placed on the temporal visual field rather than on the nasal visual field (the half of the visual field of each eye closer to the nose; e.g., Lewis & Maurer, 1992). This behavioural evidence fits with subcortical control of orienting (M. H. Johnson, 1990, 2011).

At 1 month postnatal age the inhibitory pathway to the superior colliculus from the primary visual cortex that goes through the substantia nigra and basal ganglia prevents infants from consistently moving their eyes after a central stimulus is presented. In other words, as a consequence of this inhibitory mechanism infants find great difficulties disengaging their gaze from a stimulus in order to make an eye-movement to the next location, leading to very long fixation durations that could even last several minutes. This phenomenon is commonly known as "*sticky fixation*" or "*obligatory attention*" and has been frequently reported in the literature (e.g., Atkinson, 2000; Farroni, Simion, Umiltà, & Barba, 1999; Frick et al., 1999; M. H. Johnson, Posner, & Rothbart, 1991; M. H. Johnson, 2011).

From 2 months of age "sticky fixation" begins to gradually diminish as a result of increasing cortical control over saccades. In addition, at this age infants start to present periods of smooth visual tracking, and become more sensitive to stimuli placed at the nasal visual field (Aslin,

1981). Johnson (1990, 2011) proposed that the appearance of this behaviour concurs with the functioning of the cortical pathway involving MT.

It is not until infants are 4 months old that the premotor areas of the frontal lobes, which contain the frontal eye fields (FEF), are mature enough to be able to systematically release the superior colliculus from the inhibition produced by the substantia nigra. As a result, "sticky fixation" is thought to fade by this age, even though some studies have suggested that it may last longer when looking at dynamic stimuli (Bronson, 1990; Hunnius & Geuze, 2004a). In addition, the development of the FEF pathway increases the infant's ability to make anticipatory saccades (even when tracking a moving stimulus) and to learn sequences of looking patterns.

1.3.2 Visual orienting and visual attention

Visual orienting involves moving the eyes, head and body in response to, or in anticipation of, a new sensory stimulus (M. H. Johnson, 2011). As it was described before, it is the fovea the region of the eye that presents the highest visual acuity, thus the orienting process implicates moving the body, head and eyes in order to locate the focus of interest in the centre of the fovea. In adults, for objects located at visual eccentricities of less than 20°, orienting can be achieved by simply moving our eyes. This region is known as the *eye field*. The *head field* covers areas beyond eccentricities of 20° but lower than 90°. In these cases orienting will also involve a head movement. For covering areas outside the head and eye fields, body movements are also necessary (Sanders, 1963).

This thesis will only consider the orienting responses to objects in the eye field, or in other words, the orienting responses that exclusively involve saccadic eye-movements. The present section reviews the current literature on *target-elicited saccade forms*. This type of saccades occur when a target makes a sudden appearance in the parafoveal or peripheral regions of the visual fields generating a voluntary or automatic response (Findlay & Gilchrist, 2003).

1.3.2.1 Factors affecting the latency of orienting saccades

An extensive number of studies have investigated the orienting responses performed with the eyes alone. An essential way to measure the orienting response in target-elicited saccade paradigms is to evaluate *reaction times*, or in other words, the time lapses occurring between the appearance of the target and the moment when the eyes start to move. This periods are also known as *latencies* and represent the cumulative time taken by the brain processes that

enable orienting (Findlay & Gilchrist, 2003). Consequently, saccade latencies can help understanding the brain processes underlying orienting responses.

1.3.2.1.1 Measuring reaction times: Gap-overlap paradigm

Latencies for target-elicited saccades can be studied by analysing the reaction times when a target that is being followed makes an unpredictable step movement. Saslow (1967) realised that the time lapse between the disappearance of the currently fixated stimulus and the appearance of the visual target affected saccade latencies. In order to investigate this effect, he designed an experiment where these two events were separated by a temporal offset: during the gap trials, the peripheral target appeared preceded by a gap period (e.g., 200 ms) after the central target disappeared. On the other hand, during the overlap trials the central stimulus did not disappear, or disappeared only after a predefined time lapse during which the peripheral target was also present (see Figure 1-3).
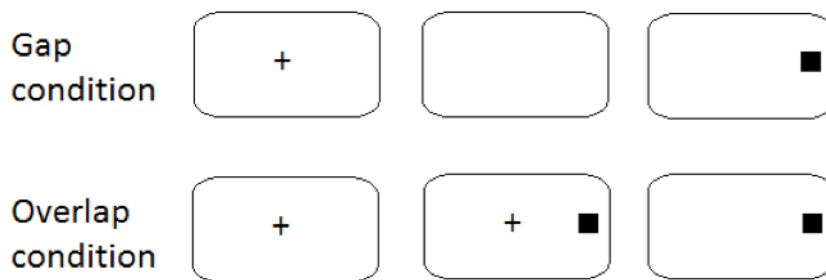
Figure 1-3 Gap and overlap trials from the gap-overlap paradigm. In the gap condition, a peripheral target appears after the temporal gap that follows the central target's disappearance. In the overlap condition, when the peripheral target appears the central target stays visible.

He found that latencies for the gap condition were progressively shorter as the temporal offset increased, while for the overlap condition they become longer as the overlap time lapse increased. This effect is known as the *gap effect* and has been reproduced in a multitude of experiments presenting different variants for this paradigm (Atkinson, Hood, Wattam-Bell, & Braddick, 1992; Blaga & Colombo, 2006; Butcher, Kalverboer, & Geuze, 2000; Domsch, Lohaus, & Thomas, 2010; Elsabbagh et al., 2009, 2013; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991; Kano, Hirata, Call, & Tomonaga, 2011; Kano & Tomonaga, 2011b; Kikuchi et al., 2011; Kopecz, 1995; Reuter-Lorenz, Hughes, & Fendrich, 1991).

It has been argued that the gap effect is the result of two general preparatory components: A general alerting component (present in any warning signal) and a disengagement effect (e.g., Reuter-Lorenz, Oonk, Barnes, & Hughes, 1995). In fact, the gap-overlap paradigm has been widely used in order to investigate the disengagement abilities and facilitation in young infants (Domsch et al., 2010; Farroni et al., 1999; M. H. Johnson et al., 1991). For more details about the gap-overlap paradigm go to Chapter 4.

1.3.2.1.2 Disengaging visual attention

The term "disengagement" has been defined as the difficulty generating an eye-movement after a fixation (M. H. Johnson, 1990). As it was discussed earlier, during the first months of life infants can present difficulties disengaging their gaze from a stimulus in order to move their eyes to another location ("sticky fixation"), which is manifested through longer reaction times and fixation durations. Thus, disengagement difficulties in infancy have been typically associated with a relative immaturity of the areas in the brain involved in eye-movement generation (e.g., the cortical pathway through the frontal eye fields, FEF; Butcher, Kalverboer, & Geuze, 2000; Matsuzawa & Shimojo, 1997), especially during the first 6 months of life when the neurological structures involved are thought to mature rapidly and get close to their adult form (Rothbart, Posner, & Rosicky, 1994). Nevertheless, disengagement can simply be the result of interrupting and shifting the current focal point (Kikuchi et al., 2011).

The ability to disengage from a central target to shift the gaze to a peripheral one has been traditionally evaluated using the gap-overlap paradigm (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991). As it was described earlier, reaction times are usually faster on gap trials as compared to overlap trials. The paradigm sometimes includes baseline trials, where the peripheral target appears immediately after the central target disappears (Elsabbagh et al., 2009; Wass, Porayska-Pomsta, & Johnson, 2011). In these cases disengagement latencies can be calculated by subtracting the baseline latencies from the overlap latencies.

Besides the developmental state of the visual system, disengagement latencies can also be affected by the visual characteristics of the central stimulus. Kikuchi and colleagues (2011) investigated the disengagement from faces and objects in children aged 9 to 17 years. They found larger disengagement latencies and saccade-related event-related potentials (ERPs)

when children disengaged from faces, suggesting that the encoding and processing of the foveated stimulus plays a role in the ability to shift the gaze from a central target. For more details about disengagement go to Chapter 4.

1.3.2.1.3 The properties of the central stimulus and the visual target

Saccade latencies can be affected by the visual properties of the target, or even by the target location (Findlay & Gilchrist, 2003). Few studies have reported shorter latencies when orienting towards a bright target compared to a blurred one (e.g., Kalesnykas & Hallett, 1994). For instance, Reuter-Lorenz and colleagues (1991) investigated the effect of the target luminance in saccadic latencies in the gap-overlap paradigm, and found a reduction in saccadic reaction times as a function of the target luminance.

1.3.2.1.4 Express and anticipatory saccades

Fischer and Boch (1983) trained monkeys in the gap-overlap paradigm and measured saccadic latencies. Results showed a bimodal distribution of saccadic latencies where apart from the expected population of reaction times there was also a sub-population of saccades that presented very short latencies (80- 100 ms). Such short saccadic latencies are known as *express saccades*, and have also been found in humans (Fischer & Ramsperger, 1984). However, in human adults the latencies were slightly larger (100- 130 ms) and the bimodality in reaction time distributions was not that well defined.

The properties and function of express saccades have been a focus of interest in the past years. Fischer and Weber (1993) argued that express saccades are visually guided by nature, but not anticipatory or predictive. On the other hand, *anticipatory saccades* were identified and distinguished from express saccades when applying a variety of paradigms such as the gap-overlap or the anti-saccade task. For instance, in the gap-overlap paradigm anticipatory saccades were defined as those with reactions times below 85-90 ms in gap trials in adult participants.

1.3.2.1.5 Variability in latencies

Saccades performed repeatedly from a location A to a location B will never have exactly the same latency (Smeets & Hooge, 2003). In fact, the cumulative distributions of reaction times are positively skewed approximating a Gaussian distribution (Carpenter & Williams, 1995). The sources for this variability have been explored in a few studies (e.g., Harris et al., 1988; Smeets

& Hooge, 2003) where it has been argued that besides a number of factors related to the variability in peak velocity and amplitude of saccades, the main causes for the variability in reaction times are the variations in the neural processing.

This phenomenon has also been approached from a computational modelling perspective (Carpenter & Williams, 1995): The cumulative distributions of reaction times can be explained with the LATER model (linear approach to threshold with ergodic rate), which includes a generative mechanism where a hypothetical variable increases at a linear rate until a given threshold is reached. The saccade will then be initiated being the time taken to reach the threshold the saccade latency. The model generates a distribution able to simulate typical reaction time distributions.

Interestingly, the variability in fixation durations during a free viewing task is much higher and unpredictable than the variability in reaction times gathered from oculomotor paradigms such as the gap-overlap (Harris et al., 1988). This can be due to the changes in the amount of processing demanded by the different features of the stimulus. Furthermore, it has been reported that infant saccades are more variable than adult saccades (Hainline et al., 1984), probably as a result of the underdevelopment of the neural structures related to visual and cognitive processing during the first months of life.

1.3.2.2 Saccade programming

The programming of a target-elicited saccade involves a series of mechanisms such as those related to target selection, target location, or inhibitory control. For instance, when moving the eyes to a single target at eccentricities below 10°, the most common pattern is the programming of a single saccade that goes directly to the target. However, occasionally the saccade will fall close to the target and a subsequent *corrective saccade* will be performed in order to correct the initial error (e.g., Henson, 1979). When there is more than one possible target, inhibitory mechanisms play an essential role in the programming of a saccade. This subsection will review some of the relevant literature on saccade programming and the paradigms that have been traditionally used for its study.

1.3.2.2.1 Labile and non-labile stage programming: The double-step paradigm

In the double-step paradigm participants are instructed to fixate two consecutive targets, in order to investigate the effects of the time lapse between the second target step and the onset

of the saccade. Typically the targets appear at varying directions and time intervals with the view to avoid the observer predicting the stimulus properties. Some designs also include one-step trials where a single target is presented.

This paradigm has been widely employed in a variety of contexts (Becker & Jurgens, 1979; Gilmore & Johnson, 1997; Heide, Blankenburg, Zimmermann, & Kömpf, 1995; Senju et al., 2011; Walker & McSorley, 2006). In research with adults, it has been typically applied in studies investigating saccade programming.

Double-step trials can have three different outcomes (see Figure 1-4): (1) the observer performs two consecutive saccades to the first and the second target. This happens when the time lapse between the first and the second target's appearances is long enough so that the saccade to the first target is already programmed when the second target appears. (2) The observer performs a single saccade directed to the second target. This occurs when the second target appears before the programming to the first target is completed. In these cases, the observer may still be able to cancel the saccade to the first target and start a new saccade program directed to the second one. (3) The observer makes a saccade to a location intermediate between the two target locations, followed by a second saccade directed to the second target.
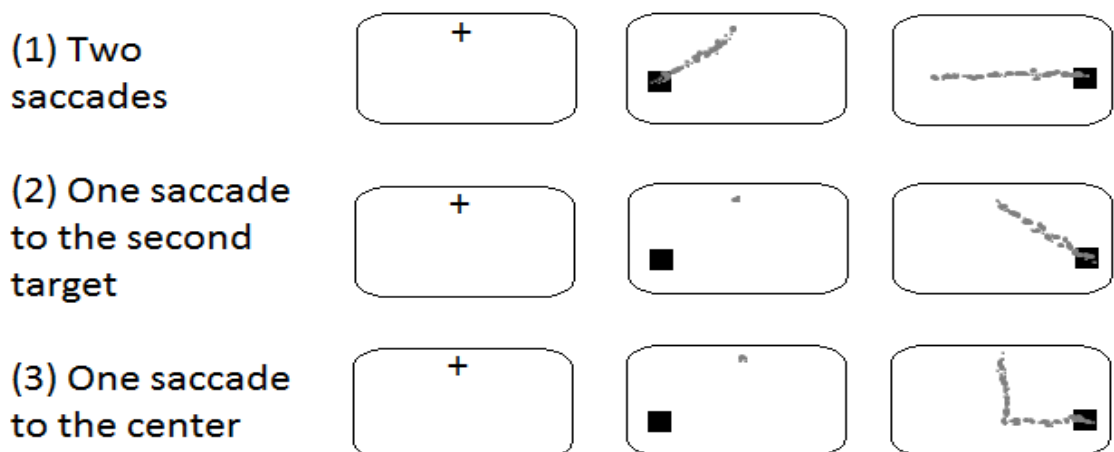
Figure 1-4 The picture above shows the three different possible outcomes from a double-step trial.

Becker and Jurgens (1979) demonstrated that the most important factor to determine the outcome of a trial was the time lapse (D) between the second target step and the onset of the

response. In fact, the delay D also depends on two variables: $D = R_1 - \tau$. While R_1 represents the reaction time, τ factors the interstep time between the first and the second target's appearances, and is the only variable under the control of the experimenter. For large τ values the observer tended to perform two consecutive saccades to the first and the second target, while for small τ values, one saccade to the second target was executed. Interestingly, for a certain range of τ values the first saccade landed in a location between the first and the second target, showing a smooth transition between the two locations.

Based on these findings, Becker and Jurgens (1979) proposed a two stage model of saccade generation: First, during the decision stage, the eyes move and their next location is determined; and second, the saccadic amplitude is calculated as a time average of the fixation error. Additionally, their results suggested that the preparatory processes of two different saccades may overlap in time.

Parallel saccade programming is thought to occur in two stages: an initial *labile stage* during which the saccade is subject to cancellation, and a subsequent *non-labile stage* in which the program cannot be cancelled any longer. In other words, if the program of a new saccade starts while another saccadic program is on its labile stage, the first program is cancelled and the second saccade is executed, resulting in a longer fixation. On the other hand, if the saccadic program begins during the non-labile stage of the previous saccade programming, both saccades are executed one after another. Even though parallel programming of saccades has not been specifically investigated in infants, a few studies have examined saccade suppression by applying paradigms such as the anti-saccade task (M. H. Johnson et al., 1991; M. H. Johnson, 1995b; Nakagawa & Sukigara, 2007)).

In infant literature, the double-step paradigm has been mainly used to identify the shift from retinotopic to spatiotopic coordinate systems (e.g., Gilmore & Johnson, 1997; Kaufman, Gilmore, & Johnson, 2006; Senju et al., 2011). For instance, Gilmore and Johnson (1997) applied the double-step paradigm to study the nature of the spatial representations that underlie simple visually guided actions in 3- and 7-month-old infants. They found that while 7-month-olds performed adult-like body-centred saccades, the younger group programmed the next saccade based on the target's retinotopic location without incorporating the effects of displacements originated by previous saccades. Based on these results Gilmore and Johnson theorized that

retinotopic representations are the most primitive ones and consequently develop first. Additionally, they argued that body-centred or spatiotopic representations emerge during the first months of life probably through experience. For more details about the double-step paradigm and saccade cancellation in infants go to Chapters 4 and 5.

1.3.2.2.2 Inhibitory control: The anti-saccade paradigm

A fundamental mechanism of visual orienting is the ability to inhibit the responses to distractors or events in the environment with no immediate relevance. This ability has been investigated in a number of studies employing a variety of tasks, such as the anti-saccade paradigm.

In the anti-saccade paradigm the subject is instructed to look to a position located in the opposite direction of the visual target being displayed. For this to happen, the subject needs to suppress or inhibit the natural tendency to make an orienting saccade to the target (Guitton, Buchtel, & Douglas, 1985). Thus, this task also serves to investigate the interaction of reflex and voluntary control of saccadic eye-movements.

Adult participants tend to increase in efficiency with training and reduce the number of involuntary saccades to the target –called *prosaccades*– the more trials they perform (Evdokimidis et al., 2002). In addition, studies have shown that patients with frontal lobe damage have difficulties suppressing prosaccades and making volitional eye-movements (Walker, Husain, Hodgson, Harrison, & Kennard, 1998). This suggests that *antisaccades* require the frontal system to send a signal to the superior colliculus in order to inhibit the natural reflexive saccade before the antisaccade is programmed. Nonetheless, some studies have proposed that both antisaccades and prosaccades are programmed in parallel (Mokler & Fischer, 1999). Likewise, Walker and McSorley (2006) suggested that endogenous ("voluntary") and stimulus-elicited ("reflexive") saccades can also be programmed in parallel on a common motor map.

Johnson (1995) developed a new version of the anti-saccade task for use with infants, in order to investigate the ability of 4-month-olds to inhibit automatic saccades. Evidently, infants do not respond to verbal instructions; thus, they were encouraged to look in the opposite direction of the cue by making the target stimulus visually more attractive than the first. After a number of training trials, the infants that are able to suppress saccades may start to inhibit their tendency to make a saccade to the cue (first stimulus) and program a saccade directed to the more

attractive second stimulus. Johnson found that 4-month-old infants were already able to inhibit saccades to a peripheral stimulus, implying maturation of the frontal eye fields (FEF) by this age.

1.4 Fixations in scene perception in adults

Gaze control in scene perception has been typically investigated through the study of two main aspects: the location where fixations land (fixation position or location) and the time-lapse during which the eyes remain still at a given location (fixation duration).

When analysing eye-movements, an essential question that needs to be addressed is whether gaze is driven by the visual features of the external world (*exogenous* or *bottom-up* factors) or by the viewer's internal thoughts or mechanisms (*endogenous* or *top-down* factors). Bottom-up accounts claim that eye-movements are guided by the properties of the visual input on the retina, and assume that oculomotor control ignores influences from higher cognitive processes and performs rather reflex-like eye-movements. Salience-based accounts of eye-movement control (reviewed later in this section) are based on this concept. On the other hand, top-down accounts assume that eye-movement control is driven by high level cognitive processes such as the viewing task or personal preferences (Land, 2009). Some researchers have proposed *mixed control theories* where eye-movements are driven by a combination of top-down and bottom-up factors (Connor, Egeth, & Yantis, 2004; Ogawa & Komatsu, 2004). Nevertheless, the extent to which each of these elements influence gaze control in certain settings is still a matter of debate (Land, 2009). In the following subsections I will review the current literature on gaze control in scene perception in human adults and show how top-down and bottom-up factors can influence both fixation locations and fixation durations.

1.4.1 Fixation locations

Much of the earliest investigations on saccades and fixations were concerned about the sequences or patterns of eye-movements and their role. For instance, Stratton (1902) investigated the sequence of eye-movements when viewing simple symmetrical shapes, finding asymmetrical patterns of eye-movements when viewing such shapes. In an acclaimed study where observers had to look at pictorial material, Buswell (1935), found that besides the huge individual differences in all the measures, eye fixations concentrated in particular areas of a picture, such as the human figures. Furthermore, he also reported different patterns of eye-

movements when attending to different instructions. Later on, Yarbus (1967) extended Buswell's work in a series of well-known studies. He gave observers different viewing tasks and investigated the scanning sequences they performed when viewing the same pictures. Interestingly, Yarbus found that scanning patterns were particularly sensitive to the mental task of the observer.

Much of the pioneering work from Buswell and Yarbus highlighted many aspects of eye-movement research –such as the influence of top-down and bottom-up factors in eye-movement control- that still remain the subject of contemporary research and hence form the basis for the studies described in this section.

1.4.1.1 Scanpaths

The study of sequences or patterns of eye-movements still awakens considerable interest in vision researchers. Some researchers have argued that particular eye-movement sequences could be associated with particular visual patterns (e.g., Noton & Stark, 1971). For instance, Yarbus (1967) investigated individual differences in scanpaths by analysing the sequences of fixations made by the same subject that was tested looking at the same picture on four different occasions. Interestingly, even though the scan patters were very similar –showing fixation clusters within the same key areas such as human faces-, they were still not identical. In another study, Yarbus tested different observers viewing the same picture, and noticed that while the differences in fixation sequences varied more from one participant to another, they still followed a similar pattern. Nevertheless, even though to an extent it is possible to find certain regularities in the patterns of eye scanning, this claim may not always be true (Mannan, Ruddock, & Wooding, 1997).

As a result of the research investigating scanpaths researchers developed some techniques to capture statistical regularities in the pattern of eye scanning have been developed (Ellis & Stark, 1986). For instance, in a Markov process, the properties of the immediate preceding saccade constraint the probabilities of the one currently being programmed (Findlay & Gilchrist, 2003).

1.4.1.2 Bottom-up control: Salience-based accounts of eye-movement control

The salient features of a scene are those that stand out and attract the attention of the viewer. Typically, an object or a part of a scene will be considered salient when it is bright, colourful,

contrasty, detail-containing, flashing, moving, or has other visual features that differentiate it from the rest of the scene (Land, 2009).

Saliency has been widely studied through computational modelling approaches, where the pre-attentive features of a scene region that contribute to its salience were computationally quantified and compared with empirical data (Borji & Itti, 2013; Itti & Koch, 2001; see the section *1.6.1*). Nevertheless, researchers have accepted that saliency is not the only factor that exclusively drives eye-movement control (e.g., Findlay & Gilchrist, 2003; Itti & Koch, 2001). In fact, a number of studies found that the empirical data containing fixation locations did not always coincide with the "hot spots" from a salience map (e.g., Parkhurst, Law, & Niebur, 2002). These findings suggest that even though saliency can still be a powerful index for predicting input-driven exogenous eye-movements there are still other higher level factors that can greatly influence oculomotor control.

1.4.1.3 Top-down control: The influence of cognitive processes in eye-movement control

The participants from the studies described in the previous subsection were not instructed to perform any viewing task while they watched the experimental stimuli. During a "*free viewing*" task it is assumed that the viewer is not applying any specific visual strategy that involves higher level cognitive processes (endogenous or top-down factors). However, it seems unlikely that the viewer can "switch off" their mind during the task and perform eye-movements free of any cognitive influence. In fact, the lack of instructions may just give them the freedom to select their own visual strategy.

As it was introduced earlier, Buswell (1935) and especially Yarbus (1967) focused much of their research on investigating the influence of top-down control in gaze allocation. They found that eye-movement control is not only influenced by the visual properties of the image, but also by top-down instructions from executive regions of the brain. More recent research have investigated the differences in eye-movement patterns between a free-viewing task and other tasks such as visual search (e.g., Henderson, Brockmole, Castelhano, & Mack, 2007; Underwood & Foulsham, 2006). These studies confirmed that while saliency models have some ability to predict eye-movement behaviour in a free-viewing task, they cannot predict the particularities of eye-movements under a visual strategy with the same ease.

Apart from the viewing task, there are other factors that can influence eye-movement control such as the scene context. A number of studies have shown the effect of the scene semantics (e.g., informative vs. non-informative objects, congruent vs. incongruent objects) in gaze allocation (e.g., Henderson et al., 1999; Loftus & Mackworth, 1978; Võ & Henderson, 2009). For instance, in a well-known study by Loftus and Mackworth (1978) it was reported that subjects fixated earlier, more often and with longer durations when there was an incongruent object in the scene (e.g., an octopus in a farm). These findings indicate that the scene is the responsible of providing the context for the objects on it and thus can be regarded as the schema that implements the framework in which objects are viewed.

In sum, these findings suggest that gaze allocation is driven by both bottom-up and top-down control. However, the extent to which each factor influences eye-movements is still a subject that requires further investigation.

1.4.2 Fixation durations

Whilst most past research on eye-movements has been focused on where fixations land (e.g., Buswell, 1935; Itti & Koch, 2001; Loftus & Mackworth, 1978), there is a growing body of research investigating the durations of those fixations. Fixation durations have been associated with cognitive processes such as attention, information processing, memory and anticipation (Castelhano & Henderson, 2008; Kowler, 2011; Malcolm & Henderson, 2010; Rayner et al., 2009; Richardson et al., 2007). While fixations may not be the only oculomotor events of interest to researchers measuring eye-movements (e.g., researchers investigating attentional shifts may be interested in raw-representations of saccadic or smooth pursuit trajectories), the majority of eye-movement research assumes that gaze allocation can be equated to visual encoding of high-spatial frequency foveal information and for this to happen the eyes need to be stable in a fixation (Rayner, 1998).

Reading research in particular, has paid very special attention to the relationship between fixation durations and visual and cognitive processing revealing how fixation durations can be affected by low-level non-linguistic factors such as word length, as well as lexical, syntactic or discourse factors (Rayner, 1998). This relationship is also evident in the scene perception literature, even though the research carried out in this area has not been as extensive. Once again Buswell (1935) was the first making detailed measurements of fixation durations during

picture scanning. He found that fixation durations tended to be shorter at the beginning of the presentation, and increased with higher exposure. This finding has been replicated several times. For instance Antes (1974) reported a mean fixation duration of 215 ms at the beginning of the exposure that after few seconds increased to 310 ms. According to Buswell's interpretation, viewers start looking at the most relevant areas of an image to subsequently analyse the details of localized regions. This view was later revised by Yarbus (1967), who claimed that the "additional time spent on perception is not used to examine the secondary elements, but to re-examine the most important elements".

Generally, when viewing pictures fixation durations form a positively skewed distribution with a mode at 230 ms, a mean of 330 ms, and fixation durations that go from 50 ms to 1000 ms approximately (Henderson & Hollingworth, 1999). As fixation locations, fixation durations can also be affected by both bottom-up and top-down control. For instance, a number of studies have shown that fixation durations can be affected by the visual characteristics of the stimulus such as luminance (Henderson et al., 2013; Loftus, 1985), or clutter (Henderson, Chanceaux, & Smith, 2009). In a recent study Henderson and colleagues (2013) applied a novel degradation gaze-contingent paradigm where the scenes were reduced in luminance only during saccades ending in critical fixations. Results demonstrated that fixation durations were at least partially under the direct influence of the point currently being fixated or under bottom-up control. All these findings validate the use of fixation durations as an index of online cognitive processing demands.

On the other hand, some other studies have highlighted the influence the viewing task (such as search or memorization) or personal preferences can have in fixation durations (Rayner, 1998; Yarbus, 1967). As it was reviewed earlier, while some of the first studies were focused on how viewing task affects gaze location (e.g. Buswell, 1935), recent articles have also pointed out that fixation durations are generally shorter during visual search compared to memorization (e.g., Henderson et al., 1999; Nuthmann et al., 2010; Võ & Henderson, 2009). Besides the viewing task, factors such as the semantics of the scene (e.g., informative vs. non-informative objects, congruent vs. incongruent objects; e.g., Henderson et al., 1999; Loftus & Mackworth, 1978; Võ & Henderson, 2009), or familiarity (Althoff & Cohen, 1999) can also greatly affect fixation durations. Nevertheless, as for fixation locations, the extent to which top-down and bottom-up factors affect fixation durations is still poorly understood.

Some studies have also found that individual differences in fixation durations remain stable across different viewing conditions. Significant relationships have been reported between fixation durations during the viewing of line drawings, photographs, computer rendered scenes and faces (Andrews & Coppola, 1999; Castelhano & Henderson, 2008).

Past research with animals as well as human adults have suggested that there are several oculomotor components that may individually affect fixation durations (Findlay & Walker, 1999; Nuthmann et al., 2010). One of these components is the oculomotor command to move to a peripheral target. It is thought to function based on the saliency computations that occur at the visual cortex and exerted via the brainstem circuitry including the superior colliculus (Becker & Jurgens, 1979; Findlay & Walker, 1999). Another component is related to the processing of visual information at the point being fixated, which is thought to function via the inhibitory control of eye-movements implemented in areas of the brain such as the frontal eye fields (FEF), the superior colliculus and the substantia nigra (Findlay & Walker, 1999). Additionally, past studies have suggested that internal saccade timing mechanisms also influence eye-movement control (Henderson & Pierce, 2008; Henderson & Smith, 2009; McAuley, Rothwell, & Marsden, 1999; Nuthmann et al., 2010). Some of these studies selectively manipulate global scene processing difficulty by employing the *scene onset delay paradigm*, whereby participants are presented with photographs of real-world scenes while performing a memory task. During the saccade just prior to the critical fixation the scene is substituted for an unpredictable amount of time with a blank screen. As the blank removes all visual content and the intended target of the prior saccade, if the eyes are under the direct control of the point currently being fixated they should remain in fixation until the scene returns. Results showed a bimodal distribution of durations of critical fixations where a population of fixations appeared to be longer despite the removal of the actual scene, while a second population terminated before the scene returned. These results support a mixed theory of eye-movement control showing how fixation durations can be under direct (i.e. demands of current fixation) and indirect control (e.g. attentional pre-setting, internal timing mechanisms and exogenous factors; Nuthmann et al., 2010). All these components have been satisfactorily modelled in human adults during simple psychophysical tasks (Carpenter & Williams, 1995), reading (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003) and scene viewing (Nuthmann et al., 2010; Trukenbrod & Engbert, 2014).

Some studies have also investigated within-participant variance in fixation durations. For instance, Benson, Castelhano, Au-Yeung, and Rayner (2012) showed how fixations from adults with autism spectrum disorder (ASD) were less influenced by the semantic content of the point being fixated.

1.5 Fixations in infancy

Looking has been widely used as an index of visual attention and a tool to study cognitive development in infants. For this purpose, researchers have traditionally analysed infants' looking behaviour by applying paradigms such as habituation, familiarization or preferential looking (see also Chapter 2). A number of these studies have analysed and reported *looking times* –and not fixation durations-, defined as the periods that the infant expends looking at a particular area of interest (e.g., a face, the right side of the screen, the left side of the screen). To date, only few studies have analysed fixation durations and oculomotor control in infants. In this section I will first provide a broad overview of what we know about infants' looking and visual attention and then I will describe the studies analysing fixation durations in infants.

1.5.1 Looking times vs. fixation durations

Looking time measures were first introduced by Robert Fantz in a series of pioneering studies (1956, 1963, 1964). Fantz (1963) found that infants are able to visually explore and selectively attend to particular stimuli (e.g., shapes or patterns) in their visual environment when they distinguish them. Moreover, Fantz (1964) found that infants become less interested (look less) in stimuli that are presented repeatedly while they showed more interest (look longer) in novel objects or events. These findings constituted the starting point for using looking time measures to investigate infant perceptual and cognitive abilities (Aslin, 2007). Since looking is a spontaneous behaviour easy to elicit in infants –unlike other techniques such as non-nutritive sucking rates-, this technique has become very popular and has been used for investigating a wide variety of subjects such as categorical learning (Quinn & Eimas, 1996), numerosity (Xu, 2003), or impossible/possible events (Baillargeon, 1987) to name a few.

Nevertheless, even though looking times measures helped shedding light into infants' perceptual and cognitive abilities, they do not allow investigating the micro-dynamics of visual and cognitive processing. Studies reporting fixation durations are becoming increasingly important as a result of their potential to explain the mechanisms underlying visual processing

or to answer questions related to the assessment of attention and information processing in spontaneous unconstrained settings. At this point it is important to mention that some of the studies using looking times as a measure have also used the term "fixation durations" to refer to them. These studies often use different criteria to define what a "fixation" constitutes (Colombo & Mitchell, 1990; Ruff & Rothbart, 1996). For instance, in some studies a fixation is defined as starting with any look to the stimulus and terminates with any clear look away from the stimulus (e.g., Courage, Reynolds, & Richards, 2006; Jankowski & Rose, 1997; McCall, Hogarty, Hamilton, & Vincent, 1973). Occasionally a relatively long minimum fixation length (e.g., 1 second, while the length for most fixations in infants range from 300 to 800 ms) was also established (e.g., Colombo, Mitchell, O'Brien, & Horowitz, 1987; Shaddy & Colombo, 2004). Due to the coarse spatial and temporal measurement, this design presents obvious limitations if we are interested in investigating the micro-dynamics of visual processing and attentional control. When the visual stimuli from which the looking time is derived fills more of the visual field than 2 degrees (e.g., a face), the infant will most likely move their eyes within the stimulus to see elements of it in detail (e.g., eyes, mouth, nose). Each "look" to the stimulus will therefore constitute multiple fixations as they are defined in adults (Saez de Urabain, Johnson, & Smith, 2014). To avoid confusion in this thesis I will only use the term "fixation duration" to refer to the periods between saccades when the eyes are stable.

1.5.2 The development of looking and visual attention during the first year

Despite the poor visual acuity of newborn infants (see section 1.3.1.1), their looking behaviour is appreciably organized and selective (Ruff & Rothbart, 1996). Fantz (1963, 1964) discovered that newborns prefer patterned stimulation rather than plain fields of colour, and that they look longer at some patterns –such as those with large features and high contrast- than others. Furthermore, other studies have suggested that during their early days infants prefer looking at curved patterns as opposed to straight lines (Fantz & Miranda, 1975).

Before 2 months of age, infants tend to visually scan certain parts of the external contours of stationary objects almost ignoring the internal features (Aslin & Salapatek, 1975). At this age, they may also find difficulties disengaging attention from a highly salient object (for more information about "sticky fixation" go to section 1.3.2.1.2). After 3 months these disengagement difficulties seem to fade out and infants start to be able to execute saccadic eye-movements more readily.

Around 2 to 3 months infants experience a big developmental transition during which their looking behaviour begins to change considerably (Ruff & Rothbart, 1996). From this age, infants start to stay awake for longer periods and hence spend significantly more time looking around (Wolff, 1987), gaining more and more experience at performing eye-movements. Some studies have shown that from 2 to 3 months infants are more likely to select particular patterns such as a face or bull's-eyes, regardless of the visual features of the alternative patterns (e.g., a bright geometrical shape; Ruff & Turkewitz, 1975). For instance, Salapatek (1975) found that infants younger than 2 months preferred to look longer to squares as opposed to lines –probably because squares were visually more salient-, regardless of whether squares represented the majority or minority of elements of the pattern. On the other hand, infants older than 2 months preferred to look at the areas in the pattern that were discrepant, despite them being composed by squares or lines. Furthermore, Salapatek (1975) also found that their scanning patterns are more likely to include the internal as well as the external features of an object. These findings indicate that around 2 to 3 months infants start showing more interest for the novel stimulus as opposed to the visually salient one, which suggests that while before that age infants' looking behaviour is predominantly driven by bottom-up or exogenous factors, from 2 to 3 months top-down or endogenous factors start influencing infants' visually guided actions.

The developmental transition that happens from 2 to 3 months also has some implications on infants' social behaviour. For instance, it has been reported that infants are more likely to make eye contact after 2 months of age (Keller & Gauda, 1987). As a result, their social experience increases being able to socially interact with other people such as their caregivers. Furthermore, it has been reported that during the first 4 months of life infants' looking at adults during social interactions increases significantly (Lamb, Morrison, & Malkin, 1987).

The transition at 2 to 3 months occurs in part due to the dramatic maturational changes in the visual and oculomotor system during the first 3 months of life (for more information go to section 1.3.1.4). Besides the improved visual acuity of 3-month-old infants, the shift from subcortical to cortical areas of visual processing allows for more complex forms of visual orienting.

The visual behaviour of infants from 3 to 9 months is increasingly influenced by experience. This is evident, for instance, in their ability to develop expectations based on the repetition of events (e.g., Johnson et al., 1991), or in the gradually decreasing looking times needed during

habituation and familiarization paradigms (Colombo, 1995). In addition, during this period infants' selective attention is markedly affected by the novelty of events and objects, and by what some researchers have called the orienting/investigative system or the first attentional system, which becomes functional early in life and governs attention during the first year (for a wider review see Ruff & Rothbart, 1996).

During the period from 3 to 9 months the visual and oculomotor systems continue their development. While visual acuity and binocularity have reached adult levels by 6 to 7 months (Aslin, 1987), other eye-movement related areas in the parietal cortex seem not to be fully functional until the end of the first year (Csibra, Tucker, & Johnson, 1998). Nevertheless, the posterior orienting network, though to be responsible for a number of processes including disengagement, shifting, and inhibition of return (M. H. Johnson et al., 1991; M. H. Johnson, 1990), is assumed to be fully functional at around 6 months of age (Hood, 1993).

The significant motor development that infants undergo from 3 to 9 months (Adolph & Berger, 2006) is also essential to understand the development of visual orienting and selectivity. For instance, at around 5 months infants are able to reach, grasp, and manipulate objects more readily. Due to their poor locomotor abilities, by this age they are more likely to engage their visual attention with objects located within their reach. By 9 months infants have acquired a great experience manipulating as well as visually exploring novel objects in their environment.

An important change that also happens around this time is the development of joint attention, which can be defined as "looking where someone else is looking" (Bakeman & Adamson, 1984; Butterworth, 2004; Ruff & Rothbart, 1996; Wurtz & Kirkham, 2010). Infants begin to share attention to the toys and objects with an adult, and learn to direct the gaze to a certain place based on social cues. Three to 4 months old infants are already able to follow "the looker's" gaze, a behaviour that 6-month-olds can perform reliably under simple ideal conditions (Butterworth, 2004; Senju & Csibra, 2008; Wurtz & Kirkham, 2010). Interestingly, when gaze is the only social cue infants are not able to reliably follow it until they are 18 months old (Moore & Corkum, 1998).

The next major transition point is considered to happen at around 9 months, as a result of the emergence of a number of new cognitive advances, including changes in memory (Fox, Kagan, & Weiskopf, 1979), the control of action (Diamond & Gilbert, 1989), or looking (Ruff & Rothbart,

1996). Several studies have reported a decrease in looking times –not fixation durations- from 3 to 13 months of age, mainly as a consequence of faster learning and habituation (Colombo & Mitchell, 1990; Kagan et al., 1971; Lewis et al., 1969; Mayes & Kessen, 1989) . Interestingly, some of these studies also reported an increase in looking at around 13 months, perhaps due to the emergence of executive functions. Looking times have also been investigated during infant playing with toys, where very different trends were found as compared to the studies investigating looking times to objects presented in an experimental display. Ruff, Saltarelli, Capozzoli, and Dubiner (1992) found no change in looking over time when 5 to 11 months infants were let to play with single objects. Nevertheless, some other studies have reported an increase in looking times over the same age range when several objects were given to the infant simultaneously (Bakeman & Adamson, 1984; Ruff & Saltarelli, 1993). These opposing results from both contexts may suggest that while some influences of attention decline, others increase, probably as a result of the development of new cognitive and motor skills (Ruff & Rothbart, 1996).

Moreover, the development of other cognitive domains such as social and emotional development can influence looking behaviour. For instance, from 9 to 12 months infants look more at their mothers' face when they are at some distance in order to search for information, a behaviour also known as social referencing (Sorce, Emde, Campos, & Klinnert, 1985; Walden & Ogan, 1988). This can also be interpreted as a search for comfort due to a boost in attachment (Baldwin & Moses, 1996). Around the same time infants also develop a transient fear of strangers (Bronson, 1972) and some behavioural inhibition to novelty or challenge (Rothbart, 1988).

During the second half of the first year it is possible to observe some rudimentary elements of executive control that become more noticeable towards the end of the first year and beginning of the second (Ruff & Rothbart, 1996). For instance, Diamond (1985) found an age effect in infants from 6 to 12 months that performed the "A not B" task. In this task, after the infant has found an object at a location A, the object is shifted to a location B. Successful reactions from the infant involve adapting the response instead of repeating the previous action. Diamond found that younger infants were unable to successfully execute this task, not only because presumably they were not able to remember the correct location, but also because they were failing to inhibit the prime response. In a later study, Diamond and Goldman-Rakic (1989)

showed the importance of the development of dorsolateral prefrontal brain structures for the "A not B" task. Executive functions continue their development throughout the second year of life together with other cognitive functions, eliciting more complex forms of looking patterns (Ruff & Rothbart, 1996).

1.5.3 Fixation durations in infancy

As described earlier, the relationship between looking times and the development of cognitive and perceptual processes has been widely investigated (Aslin, 2007; Colombo et al., 1995, 1991; Colombo, Mitchell, O'Brien, & Horowitz, 1987; Colombo, Richman, Shaddy, Greenhoot, & Maikranz, 2001; Colombo & Mitchell, 1990; Colombo, 1995; Sigman, Cohen, Beckwith, & Parmelee, 1986; Slater, 1995; see also Chapter 2). Nevertheless, not many studies have paid attention to the micro-dynamics of visual and cognitive processing, such as fixation durations and their underlying mechanisms. These investigations are becoming increasingly important especially in studies concerned with the assessment of attention and information processing in spontaneous unconstrained settings.

There are few studies that have investigated the relationship between infant fixation durations and cognitive and perceptual processing. For instance, Harris et al. (1988) tested infants and adults in a free-viewing task, revealing that fixation durations decreased as the stimulus size increased (see also Hainline et al., 1984). However, this study did not address the extent to which the size of the visual stimulus can affect infants' fixations at different developmental stages in infancy. Bronson (1990) investigated the development of visual scanning (and fixation durations) in a longitudinal study with infants from 2 to 14 weeks that were presented with a number of static, moving and flickering geometric figures. Besides finding more mature and accurate scanning strategies in older infants, he found that under flicker conditions older infants showed a higher proportion of longer fixations compared to other viewing conditions, which is a looking behaviour more common in younger infants and has been commonly associated with the maturation of volitional eye-movement control (Atkinson, 2000; Bronson, 1990; Hood & Atkinson, 1993). Furthermore, when viewing static stimuli, 1 to 2 month infants showed a series of long fixations that were closer together. By 3 to 4 months infants showed a greater proportion of short fixations (< 500 ms) when looking at static stimuli (Bronson, 1994). Bronson also found that by 6 weeks, infant fixations are influenced by whether the look falls on a stimulus contour or not.

More recent studies have investigated infant scanning abilities when presented with familiar and non-familiar complex dynamic stimuli. Hunnius and Geuze (2004) eye-tracked a group of infants in a longitudinal study between the ages of 6 and 26 weeks. Participants were presented with a video of their mother's face, and an abstract video. They found that infants only adapted their eye-movements according to the stimulus type from 14 weeks on, showing longer fixation durations for the abstract unfamiliar condition. Additionally, the median fixation duration did not stabilize before 18 weeks, which is slightly later than what has been reported for static stimuli (Bronson, 1990).

Individual differences have also been the focus of some of the work investigating fixation durations in infants. For instance, Bronson (1994) tested 6 to 13 week old infants and found that decreases in visual attentiveness were associated with longer fixation durations while looking at static figures. In another more recent study, Wass and Smith (2014) presented 11-month-old infants with a battery of complex dynamic stimuli and non-complex static stimuli. They reported that individual differences in fixation durations were stable across different stimulus types and testing sessions with infants who demonstrated short average fixation durations during one visit also displaying relatively short fixations in a second visit a week later. Furthermore, reaction times gathered from the gap-overlap paradigm correlated with dynamic stimuli, but not with static stimuli for which measures of cognitive control and arousal were most predictive of performance.

As looking behaviour (Fagan III, 1984; see also Slater, 1995), fixation durations in infancy have also been associated with later intellectual function in childhood. For instance, in a recent study Papageorgiou and colleagues (2014) demonstrated that individual differences in fixation durations in early infancy can predict individual differences in temperament in childhood. These findings could ultimately lead to early intervention practices aiming to improve executive attention and potentially identify infants at risk of attentional disorders such as the attention deficit hyperactivity disorder (ADHD).

These findings suggest that fixation durations in infancy can also reflect the cognitive and perceptual processing of the visual input, even though it is still unclear whether these factors have the same influence in infants and adults. In addition, fixation durations can be an indication of different cognitive or neural processes emerging at different developmental stages.

Pertaining to the type of stimuli used in experiments investigating eye-movements in infants, it is important to point out that, probably in an attempt to control the variables of the studies, most studies made use of very simple static or sometimes dynamic shapes as a stimuli (Bronson, 1990, 1994; Colombo et al., 1995; Frick et al., 1999; but see Mallin & Richards, 2012; Wass & Smith, 2014). This approach has been often criticized arguing that these stimuli are not adequately representative of real-world experiences and thus do not have the same effect as the more naturalistic ones (Hunnius & Geuze, 2004a; Hunter & Richards, 2011; Neisser, 1976; Schmuckler, 2001). In fact, recent research such as the study by Mallin and Richards (2012) demonstrated the importance of real world stimuli regarding performance characteristics in an experiment with infants at 14, 20 and 26 weeks whereby the stimuli were created by using "Sesame Street" video clips.

1.6 Modelling visual attention

Computational models allow us to describe, predict and explain data that is itself unobservable (Lewandowsky & Farrell, 2011). This approach does not supplant more traditional behavioural methods, but rather complements them by investigating the mechanisms that are not directly accessible through experimentation (Braitenberg, 1984; Schlesinger & McMurray, 2012). One of the main advantages of computational modelling is then that they allow the easy manipulation of the mechanisms built into the model permitting the researcher to observe the consequences of such transformations. By doing this we can investigate the functioning of hypothesized mechanisms that could not be explored otherwise (Schlesinger & McMurray, 2012). Furthermore, models can start as simple constructs and progress into more complex forms.

Apart from these advantages, Schlesinger and McMurray (2012) describe a number of reasons why psychological research, and developmental science in particular, can benefit from computational models. First, to build a model forces researchers to define theoretical constructs that could potentially be evaluated using behavioural studies. Secondly, models that do not work as expected can also be informative as they may reveal certain constraints that are important for the understanding of the phenomenon being studied (see Elman, 1993; McMurray, Aslin, & Toscano, 2009). Moreover, the most theoretically informative models can also be discarded if they fail to offer a clear explanation of their contribution. Thirdly, models can combine different levels of analysis. This point can be particularly important in developmental science since it is common to find explanations of developmental change that include

interactions between different domains such as genetics, neuroscience, or cognition (Elman et al., 1996; Mareschal et al., 2007; Westermann et al., 2007). Finally, models do not necessarily mimic real life. This means that models can be a great tool to investigate the effect of unnatural inputs, or other phenomena such as atypical development (Thomas, 2003).

Whilst It is possible to select the best fitting model based on quantitative evaluation and intellectual and scholarly judgment (Lewandowsky & Farrell, 2011), most of the time a data set can be explained using several alternative models based on different perfectly valid hypothesis. Additionally, recent scientific articles have presented extended reviews on computational models of visual attention in general (Frintrop, Rome, & Christensen, 2010; Heinke & Humphreys, 2005), infant cognitive development (Mareschal, 2010; Schlesinger & McMurray, 2012), or atypical development (Thomas, 2003), highlighting the importance of these methods for investigating visual attention and infant typical and atypical development.

Computational models can be classified according to different criteria. For instance, models can be divided into *symbolic* (Klahr & MacWhinney, 1998; Klahr & Siegler, 1978; Newell, 1994) and *sub-symbolic* (Rumelhart & McClelland, 1985) models. Symbolic models are those that encode the units of information in a language-like format (Schlesinger & McMurray, 2012; see the *production-system model*, Klahr, Langley, & Neches, 1984). *Bayesian networks* are a very particular type of symbolic models, where explicit knowledge is acquired using the laws of rational probabilistic inference (Gopnik & Tenenbaum, 2007; Perfors, Tenenbaum, Griffiths, & Xu, 2011). Sub-symbolic models on the other hand are based on distributed representations. For example*, connectionist models* belong to this category (Elman, 1993; Elman et al., 1996; Mareschal & Shultz, 1996; Shultz, Mareschal, & Schmidt, 1994; Westermann, Ruh, & Plunkett, 2009; Westermann & Ruh, 2012). *Artificial neural networks* are usually represented as a set of interconnected "neurons" or computing elements. Each of these elements receives an input that is weighed and transformed by a function (see Rumelhart, Hinton, & Williams, 1985) and then passed on to another neurons. The process terminates when the output neuron is activated and computes the resulting value. These network models could also be classified as *adaptive models*, since the internal structure of the model can change (in this case through changes in the strengths of the connections between units or changes in the architecture of the network; Schlesinger & McMurray, 2012), in opposition to *static models* where the internal structure remains the same. In the past few decades connectionist models have gained increased

popularity to study developmental change (Mareschal & Shultz, 1996; Mareschal, 2010; Schlesinger & McMurray, 2012; Westermann et al., 2009; Westermann & Ruh, 2012). For instance (Westermann & Ruh, 2012) used neural networks to investigate learning and processing the English past tense. In this adaptive model units and connections were added and removed during learning to develop a task specific final architecture.

Two important aspects of gaze control are the location where fixations land (the "where" decision) and how long these fixations last (the "when" decision). Most current computational models are concerned with fixation locations (Borji & Itti, 2013), whereas there are only few models accounting for fixation durations (Engbert et al., 2005; Nuthmann et al., 2010; Trukenbrod & Engbert, 2014). In the remainder of this chapter I will first review the most relevant models concerned with fixation locations and fixation durations, and then I will move on to describing current computational models on visual attention in infants.

1.6.1 Modelling fixation locations

The locations where fixations land is an overt behavioural manifestation of selective attention, and for this reason has been extensively investigated in behavioural studies (Baddeley & Tatler, 2006; Buswell, 1935; Findlay & Gilchrist, 2003; Tatler, Baddeley, et al., 2005; Underwood & Foulsham, 2006; Yarbus, 1967; see the section *1.4.1*). Likewise, predicting where viewers fixate has been the focus of numerous studies in the realm of computational modelling (Borji & Itti, 2013). The most popular hypothesis in the computational vision literature is the saliency hypothesis, which states that it is the bottom-up stimulus-based information generated from the image that drives the viewer's attention and hence motivates the decision of where to fixate next (Itti & Koch, 2000, 2001; Land, 2009). In saliency models the pre-attentive features of a scene region that contribute to its salience are computationally quantified and validated against eye-movements of human observers. Itti and Koch (2000, 2001) developed a saliency model able to identify a number of low-level visual features –such as colour, brightness and orientation– that are known to be extracted at the early stages of visual processing. Each of these low-level features produces a feature map that is then combined additively with the other feature maps creating a single salience map. This salience map will have "hot areas" that correspond to regions where a lot of visually relevant things are going on. Therefore, this model is only able to predict gaze allocation based on the visual features of the image (exogenous or bottom-up factors), without taking into account higher lever factors such as the visual task (endogenous or

top-down factors). Even though saliency is the only component of saccade initiation that has been modelled in a computationally satisfactory way (Land, 2009), researchers have accepted that saliency is not the only factor that exclusively drives eye-movement control (e.g., Findlay & Gilchrist, 2003; Itti & Koch, 2001). A number of studies found that the empirical data containing fixation locations did not always coincide with the "hot spots" from a salience map (e.g., Parkhurst, Law, & Niebur, 2002). Furthermore, the correlations between fixation selection and particular saliency features vary depending on the study: while some have found stronger correlations for contrast and edge information (Tatler, Baddeley, et al., 2005), others have found them for colour (Jost, Ouerhani, Wartburg, Müri, & Hügli, 2005). In accordance with these findings, recent studies have proposed that top-down guidance plays the primary role in gaze allocation and replace the concept of salience with an unprioritized input representation (Henderson et al., 2007; Henderson, Malcolm, & Schandl, 2009).

Some other models, such as the contextual guidance model for object search in real-world scenes (Torralba, Oliva, Castelhano, & Henderson, 2006), combine both bottom-up and top-down information. In this model gaze allocation is guided from a bottom-up salience computation modulated by contextual priors or task constraints (see also Navalpakkam & Itti, 2005, 2006). In the Findlay and Walker (1999) saccade generation model, the final saliency map is computed from separate topographically mapped channels for intrinsic salience (same as in Itti and Koch, 2000, 2001), search selection, peripheral visual evens such as sudden unexpected intrusions, and spatial selection. Interestingly, this model includes the two separate parallel pathways thought to be involved in saccade generation (see Figure 1-5), concerned with the spatial programming (the where component), and the temporal programming (the when component) of the saccadic eye-movement. The model accounts for a number of well-established phenomena in target-elicited saccades, such as the gap effect, express saccades, the remote distractor effect, and the global effect (Findlay & Walker, 1999).
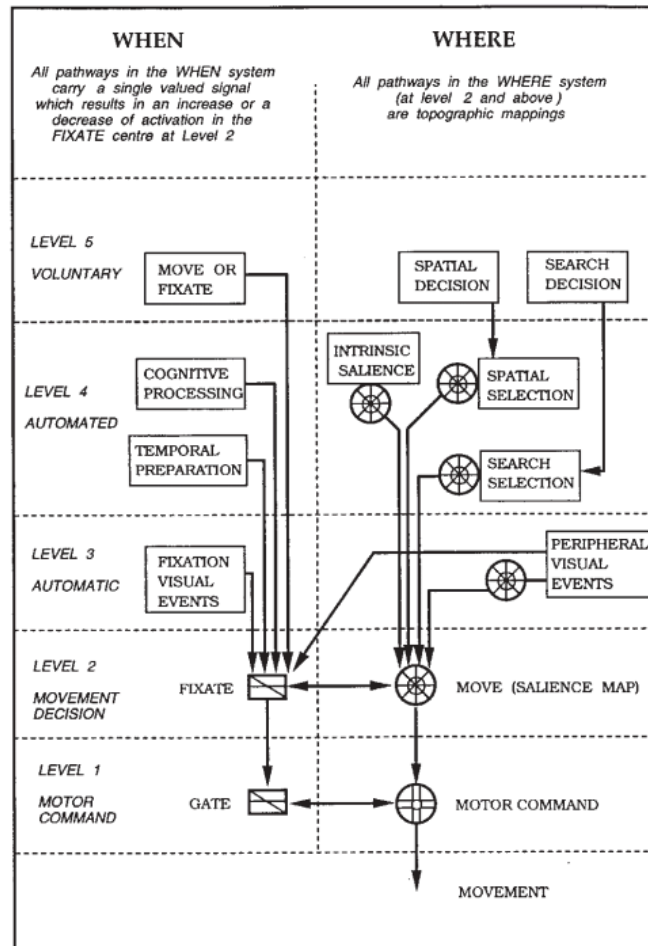
Figure 1-5 From Findlay and Walker (1999). Diagram indicating the information flow routes and competitive pathways in saccade generation.

For a wider review on computational models of visual attention that compute saliency maps refer to (Borji & Itti, 2013).

1.6.2 Modelling fixation durations

Whilst many models investigating visual attention have examined where fixations land, very few models provide control mechanisms for the duration of fixations or aim to shed light on the mechanisms underlying saccade generation and the encoding of visual information. Most of the models that do exist were developed for the task of reading. Here, the two most advanced mathematical models are the E-Z Reader model (Reichle et al., 1998, 2003) and the SWIFT model of saccade generation during reading (Engbert et al., 2002, 2005). The E-Z Reader (Reichle et al., 1998, 2003) is a processing model that extends the work from Morrison (1984), and provides a theoretical framework for understanding how word identification, visual processing, attention and oculomotor control jointly determine when and where the eyes move

during reading. The model's main two assumptions are: (1) a stage of word identification is the signal to move the eyes; and (2) attention is allocated from one word to the next in a strictly serial manner. The SWIFT model (Engbert et al., 2002, 2005), was developed to investigate the possibility of spatially distributed processing and to implement a general mechanism for all types of eye-movements observed in reading experiments (the E-Z reader is intended to be a "default" reading process since this model does not account for the many effects of higher-level linguistic processing on eye-movements). The SWIFT model is compatible with the general framework of generation of saccades developed by Findlay and Walker (1999) and includes concepts from the dynamic field theory of movement preparation by Erlhagen and Schöner (2002).

Both E-Z Reader and SWIFT models implement the notion of saccade programming being completed in two stages, as suggested by results from double-step experiments (Becker & Jurgens, 1979). However, the two models differ with regard to the mechanisms that control fixation durations. While in the E-Z-Reader model what triggers a new saccade program is conceptualized to be lexical processing, in the SWIFT model it is implemented as a random timer. In the SWIFT model, lexical processing difficulty of the currently fixated word modulates fixation durations by inhibiting the timer so that it delays the initiation of the next saccade program.

The mechanisms underlying saccadic control have also being explored by a few other models concerned with other domains of eye-movement research such as scene perception. For instance, the LATER model (linear approach to threshold with ergodic rate; Carpenter & Williams, 1995) attempted to explain cumulative distributions of reaction times. This model includes a generative mechanism where a hypothetical variable increases at a linear rate until a given threshold is reached. The saccade will then be initiated being the time taken to reach the threshold the saccade latency. The model generates a distribution able to simulate typical reaction time distributions.

Building upon the idea of a random saccade timer (Engbert et al., 2002) the CRISP model is a computational model and a theoretical framework that accounts for fixation durations in adult scene viewing (Nuthmann et al., 2010). CRISP is a timer (C)ontrolled (R)andom-walk with (I)nhibition for (S)accade (P)lanning model (see also the ICAT model; Trukenbrod & Engbert,

2014). The model architecture can be summarized with the following three main assumptions on saccade timing and programming: (1) The accumulation of activity to a saccade threshold is implemented via a random walk process (Boccignone & Ferraro, 2004) and is responsible for generating inter-saccadic intervals and hence variations in fixation durations; (2) saccade programming occurs in two stages: an initial, labile stage that is subject to cancellation, and a later, non-labile stage; and (3) processing difficulty can inhibit saccade timing and programming in a moment-to-moment fashion. The latter can happen in two ways. First, in case of increased processing demands the random walk saccade timer slows down, which delays the initiation of a new saccade program and eventually leads to longer fixation durations. Secondly, processing difficulties can even cancel an ongoing labile saccade program, which extends the duration of the current fixation. For more details about the CRISP model's architecture and theoretical assumptions go to Chapter 5.

1.6.3 Modelling visual attention in infants

Traditionally, developmental science has employed both observational methods as well as controlled empirical experiments in order to investigate child development. Nevertheless in the past decades computational modelling has become an essential tool to investigate developmental change in infancy (for reviews see Mareschal, 2010; Schlesinger & McMurray, 2012). Many of these models have focused on different aspects of cognitive development such as learning (Kemp, Perfors, & Tenenbaum, 2007; Klahr et al., 1984; Mareschal & Johnson, 2002), language development (McMurray, Horst, & Samuelson, 2012; Rumelhart & McClelland, 1985; Westermann et al., 2009; Westermann & Ruh, 2012) or memory (Mareschal & French, 1997).

Very few models, however, have investigated visual attention in infants. Some of these models have focused on habituation paradigms (Balkenius, 2000; Domsch et al., 2010; for a review see Sirois & Mareschal, 2002). For instance, Stanley (1976) used mathematical analysis and computer simulations in the form of functions in order to develop one of the first computational models of habituation. Many of the later models on habituation opted for the use of connectionist approaches (Sirois & Mareschal, 2002).

Computational modelling has also been used to investigate visual selective attention (Schlesinger, Amso, & Johnson, 2007), information processing (Domsch et al., 2010), or gaze

following (Bugajska, Trafton, Fransen, & Harrison, 2009; Carlson & Triesch, 2004; Triesch, Teuscher, Deák, & Carlson, 2006). Carlson and Triesch (2004) presented a model that aimed to investigate if gaze following is a skill that infants acquire because they learn that monitoring their caregiver's direction of gaze allows them to predict the interesting objects or events appearing in their environment. Results not only confirmed this hypothesis, but also showed how plausible modifications of model parameters motivated by findings on developmental disorders lead to impairments in the learning of gaze following. Similarly, a number of other studies have shown the effectiveness of using computational models to study atypical development (Thomas, 2003).

In another study by Mareschal, Plunkett, and Harris (1999) neural networks were also used to investigate why object permanence appears to develop earlier if measured by gaze behaviour as opposed to reaching. In this model, one sub-network controls gaze while another controls reaching. A third sub-network common to both controls object representations. Moreover, the Mareschal model includes dual-stream visual processing (dorsal and ventral pathways; Ungerleider & Mishkin, 1982), where the ventral sub-network simulates feature-based processing in the temporal cortex, and the dorsal sub-network simulates spatiotemporal processing in the parietal cortex. Results showed that the dorsal sub-network learns faster than the ventral one. Mareschal et al. hypothesized that the developmental advantage of gaze over reaching is a consequence of the earlier emergence of the dorsal pathway or the "where" system.

While these and other studies helped shedding light onto the development of infant visual attention, to date no computational model has attempted to explain the specific mechanisms underlying saccadic control in infancy.

1.7 Conclusion

In this chapter I have discussed the background literature on vision, visual orienting, and visual attention in infants and adults as well as previous empirical work and theoretical and computational models investigating visual attention and the mechanisms underlying oculomotor control and fixation durations.

As outlined earlier in this chapter, much of the research investigating gaze allocation (both behavioural and computational studies) has been focused on where fixations land (Baddeley &

Tatler, 2006; Borji & Itti, 2013; Buswell, 1935; Findlay & Gilchrist, 2003; Itti & Koch, 2000, 2001; Tatler, Baddeley, et al., 2005; Underwood & Foulsham, 2006; Yarbus, 1967), leaving the durations of those fixations aside. Nevertheless, a number of studies (also in the scene perception literature) and computational models (such as the CRISP model; Nuthmann et al., 2010) have illustrated the relationship between fixation durations and visual-perceptual and cognitive processing (e.g., Castelhano & Henderson, 2008; Kowler, 2011; Malcolm & Henderson, 2010; Rayner et al., 2009; Richardson et al., 2007). Furthermore, some behavioural work and computational models have also attempted to explain the mechanisms underlying saccadic control and fixation durations in human adults (Engbert et al., 2005; Nuthmann et al., 2010; Reichle et al., 1998, 2003; Trukenbrod & Engbert, 2014). In infants, however, as a result of the many technical and practical constraints, very few studies have attempted to explore these mechanisms and their development through the first year of life and thus the understanding of the development of infant oculomotor control is restricted (but see Bronson, 1990, 1994; Hainline et al., 1984; Harris et al., 1988; Wass & Smith, 2014).

In this thesis I will combine a series of behavioural cross-sectional and longitudinal studies with computational modelling in order to investigate the development of fixation durations and the mechanisms underlying fixation durations and saccadic control during the first year of life. For this purpose I will first analyse the factors influencing fixation durations and saccadic control in 6-months-olds and adults that viewed dynamic naturalistic and semi-naturalistic scenes and performed various oculomotor paradigms (Chapter 4). This data will then be modelled using the CRISP model of fixation durations in scene viewing (Nuthmann et al., 2010) in order to investigate the unexplored mechanisms underlying fixation durations and saccadic control in 6-month-olds (Chapter 5). Model simulations will allow investigating the extent to which the developmental state of the visuo-motor system and visual and cognitive processing influence fixation durations at this age. The longitudinal studies from Chapter 6 will investigate the development of fixation durations and saccadic control in groups of infants that were eye-tracked throughout their first year of life while viewing dynamic and static material and performed the gap-overlap paradigm. Model simulations on this longitudinal data (Chapter 7) will serve to analyse how the mechanisms identified in the previous cross-sectional study (Charter 5) may develop during this period. Furthermore, in the next two chapters I will review some of the technical and practical limitations associated with recording and analysing infants'

eye-movements, and I will introduce GraFIX (Saez de Urabain et al., 2014), a new method and software able to detect fixations in low and high quality eye-tracking data.

Chapter 2: General methodology: Eye-tracking in infancy

2.1 Introduction

In this thesis eye-movements where recorded using an eye-tracker. Even though eye-tracking is considered to be a precise measurement technique, infant testing entails a number of technical and practical constraints that need to be considered when performing an experiment. The first half of this chapter introduces general eye-tracking methodology and a description of the equipment and techniques used for recording the data presented on this thesis. The second half of this chapter describes the particularities and difficulties of measuring eye-movements in infants and the strategies used in this project to overcome some of these issues.

2.2 Eye-tracking: General methodology

The first attempts to measure eye-movements were developed in the late 1800s (Delabarre, 1898; Huey, 1898). Even though few of the early approaches already applied some of the techniques used in current eye-trackers, such as the principle of photographing the reflection of an external light source from the fovea (Dodge & Cline, 1901), they were intrusive and not suitable to use in non-compliant populations. Moreover, data analysis was time consuming, often involving the coding of frame-by-frame videos.

During the past decades and as a result of the development of video and computer technologies eye-tracking has seen an enormous progress. This has allowed researchers to address questions that could not be investigated before and thus has opened up entirely new approaches and areas of inquiry (Duchowski, 2007). In the realm of developmental psychology, researchers had to wait for the development of infant-friendly eye-tracking devices that allow a certain degree of head motion and record the participant's gaze in a non-invasive manner. Before this happened infant researchers evaluated infant's preferences, habituation or response to novelty by relying on human observers that had to code the duration (and location) of infant looking using stopwatches, event recorders or computer keyboards (Oakes, 2012). Current eye-tracking studies can provide new insights into the emerging of cognitive, social and emotional processing in infancy (Oakes, 2012; Richardson & Johnson, 2008). For instance, a number of eye-tracking experiments has investigated topics such as language development (Young, Merin, Rogers, & Ozonoff, 2009), perceptual learning (Johnson, Slemmer, & Amso, 2004), memory (Richmond & Nelson, 2009), or face processing (Liu et al., 2011) to name a few.

Current eye-tracking devices are able to record eye-movements as well as gaze allocation at sampling rates that can go from 30 to over 1000 frames per second (Holmqvist et al., 2011). Most modern eye-trackers measure the movement of the eye using video processing software – and even dedicated hardware- that implements the *corneal-reflection technique*, first developed by Salapatek, Kessen and Haith (Haith, 1969; Salapatek & Kessen, 1966). This technique is based on the recording of the location of near infra-red light reflected of the cornea (Feng, 2011; Gredebäck, Johnson, & von Hofsten, 2009). As a result of the spherical properties of the eyeball, the illumination of the infrared light creates a bright glint on the back of the cornea –also known as the *first Purkinje image*- that remains relatively stationary while the eye moves. It is this glint and its distance relative to the centre of the pupil that is used to estimate the gaze on the screen. Depending on the location of the light sources, the corneal reflection technique is classified as "black pupil" or "bright pupil". When the light source is located on-camera-axis, or in other words, in front of the participant, the light is reflected from the retina and the recorded image displays a bright pupil. On the other hand when the light comes from the sides the eye image displays a dark pupil (see Figure 2-1).
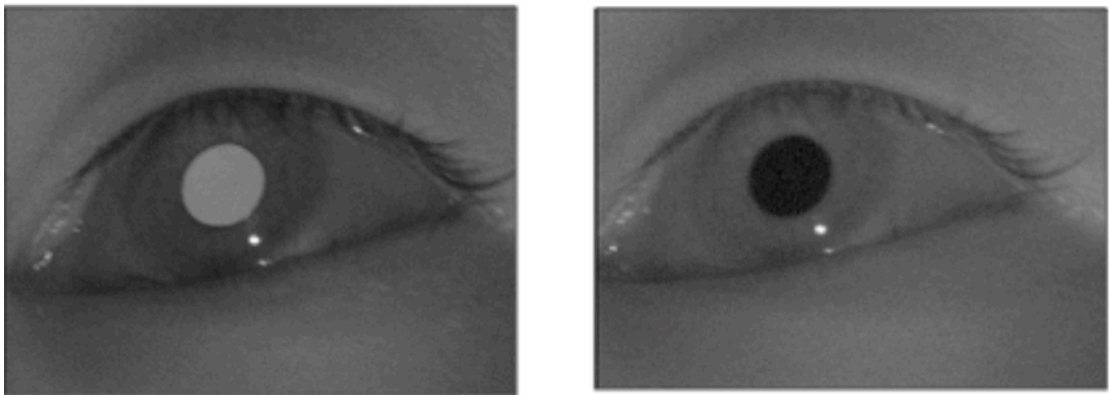
Figure 2-1 These images show how the eye image is captured when using corneal reflection techniques. The image on the left is from the bright pupil technique while the one on the right shows the dark pupil technique.

According to an editorial comment by Oakes published in *Infancy* (2010), articles reporting eye-tracking data should follow the next 8 guidelines: (1) describe the geometry of the testing situation and stimuli, (2) provide details about the eye-tracking system, (3) Fully describe the calibration procedure, (4) outline procedures to deal with missing data, (5) include details about how eye-movement data were processed, (6) specify data reduction procedures and

parameters, (7) provide information about how areas of interest (AOIs) were defined, and (8) provide complete information about exclusion criteria. These guidelines -with the exception of number 7 since this project does not include AOIs- will be addressed in this thesis, either in the present or in the following chapters.

2.2.1 Hardware

Eye-trackers are advanced physiological measuring systems (Holmqvist et al., 2011). Currently, there are a high variety of models from different manufacturers able to satisfy the user's needs. For instance, depending on how the eye-tracker contacts the user one can classify these devices into two categories: *Spatially fixed* or *static* eye-trackers (e.g., Tobii TX300, EyeLink 1000, SMI RED) and *head-mounted* eye-trackers (e.g., positive science eye-trackers, Ergoneers Dikablis glasses). While the first are widely used for lab-based experiments (e.g., Castelhano & Henderson, 2008; Mital, Smith, Hill, & Henderson, 2010; Nuthmann, Smith, Engbert, & Henderson, 2010; Papageorgiou et al., 2014; Tummeltshammer, Wu, Sobel, & Kirkham, 2014; Wu & Kirkham, 2010), head-mounted eye-tracking is becoming increasingly popular for doing research under more naturalistic and ecologically valid conditions. The difficulties of using these more intrusive devices go from measurement –the set-up often includes a hat with two attached scene and eye cameras, plus a belt or bag to carry the batteries- to analysis, which may involve a great deal of hand-coding (Aslin, 2012). Despite this there is an increasing number of articles that make use of head-mounted eye-trackers in human-adults (Land, Furneaux, & Gilchrist, 2002; Land & Lee, 1994; Land & Tatler, 2001), as well as in less compliant populations such as infants (Corbetta, Guan, & Williams, 2011; Franchak & Adolph, 2010; Franchak, Kretch, Soska, Babcock, & Adolph, 2010) or chimpanzees (Kano & Tomonaga, 2013).

Whilst some of the spatially fixed eye-trackers require the participants keep their head still – often by using a forehead and/or a chin-rest- during the measurement session, others, called *remote* eye-trackers, allow for head-movements within a certain area in front of the tracker also known as the *head-box* (Holmqvist et al., 2011). The state-of-the-art remote eye-trackers still provide lower-quality data than those that require head restrictions, mainly due to problems in gaze estimation models during head movement, or because the eye is generally filmed at lower resolutions (Holmqvist et al., 2011). These eye-trackers are, however, the only practical

alternative to test infants or other special populations that do not attend to experimental instructions.

The number of samples per second is an important property of eye-trackers that is measured in *hertz* (Hz). For instance, a 300 Hz eye-tracker will provide gaze data 300 times per second. While sampling frequency needs to be high enough to be able to calculate certain eye-tracking measures (e.g., studies investigating microsaccades have sampling frequencies no lower than 200 Hz), high-speed eye-trackers (e.g., > 500 Hz) are usually more intrusive for the participants, allow less head-movement, are more expensive, and produce larger data files (Holmqvist et al., 2011). Because of this reason many of the eye-trackers that are currently used for studies with infants do not have very high sampling frequencies.

Another critical feature of eye-trackers is the latency and temporal precision. *Eye-tracker latency* is the delay that happens from when the eye is recorded until the recording computer detects its signal. Even more important is the *stimulus-synchronization latency*, which arises in the interaction between the software being used for receiving the eye-tracker's signal and the one for presenting the stimuli on the experimental display. It is essential to keep these latencies as low as possible, especially for studies that require precise synchronization to external devices such as EEG or fMRI, or when using gaze-contingencies (Holmqvist et al., 2011). The *temporal precision*, on the other hand, refers to the standard deviation of the eye-tracker's latency, or in other words, the time interval between successive samples.

For the studies presented in this thesis participants' gaze was monitored using a Tobii TX300 eye-tracker running in a sampling rate of 120 Hz. The sampling frequency was lowered to 120Hz –rather than using the 300Hz that the Tobii TX300 allows- because it significantly improved the quality of the data. This particular eye-tracker model tolerates large head movements allowing the infants to move naturally in front of the stimuli presentation screen. More precisely, the freedom of head movement at 65 cm was within 37 x 17 cm (width x height), and the maximum head-movement speed was 50 cm/sec. The eye-tracker could recover gaze positions within 10-165 ms. The eye-tracker latency was reported to be less than 10ms. The stimuli were presented on a 23'' wide screen TFT monitor set to 1024 x 768 pixels and attached to the eye-tracker unit. The sounds were played through stereo external speakers located at

both sides of the screen. Participant were monitored and recorded through an external video camera located under the Tobii screen by using the ScreenFlow screen-casting software.

2.2.2 Software

Most lab-based eye-tracking studies present the stimuli (e.g., still or dynamic images) on a monitor. For this to happen it is required to prepare the monitor-based experiment using one of the options available for this purpose.

Many manufacturers provide software packages for the presentation of the stimuli (e.g., Tobii studio, SMI BeGaze) that are easy to integrate with their hardware. Nevertheless not even one of these packages provides support for all types of experiments. An alternative to the manufacturer's software is the professional stimuli presentation software, such as E-Prime, Presentation or PsyScope. The main benefits of this software are the large user community and the fact that the templates for many simple experiments are already available online. On the other hand, integrating them with the hardware may not be straightforward and, as the manufacturer's software, they do not support all types of experiments.

When the study requires having a good control over the experimental events (e.g., developing gaze-contingencies; integrating the eye-tracking experiment with EEG equipment), the researcher can use a programming environment such as Matlab with the Psychophysics Toolbox (Brainard, 1997; Shukla, Wen, White, & Aslin, 2011) or Python with PsychoPy (Peirce, 2007).

The stimuli used in this thesis were presented using a MATLAB program (MATLAB version R2010a 32-bit) specially developed for the purposes of the experiments. This program made use of the T2T (Talk to Tobii) package and the Psychophysics Toolbox Version 3 (PTB-3; Brainard, 1997; Shukla et al., 2011). The stimulus synchronization latency when using T2T has been reported to be around 100 ms (Shukla et al., 2011).

2.2.3 Calibration

Before the experiment starts each participant needs to perform a calibration. Its main purpose is to adapt the parameters for the calculation of gaze direction to the participant's eye and the particularities of the current testing session, such as the luminance of the room or the participant's distance and angle with respect to the eye-tracker. Even in adult participants, the eyeball radius can vary by up to 10% and have different shapes, changing the geometrical

values necessary for calculating gaze direction (Hammoud, 2008). Differences between eyeballs can be even more obvious in infants, whose eyes are still undergoing development (Banks & Shannon, 1993; Yuodelis & Hendrickson, 1986).

In a typical calibration procedure the participant needs to look at a number of predefined calibration points that appear subsequently on the experimental screen. Alternatively, when using head-mounted eye-trackers or other set-ups that do not use an experimental screen, points can be located at certain regions in the scene. Even though some of the systems that are currently being developed only require 1 or even none calibration points, usually participants are presented with 5, 9, or even 16 points (see Figure 2-2). Ideally, calibration points should be small and animated (Holmqvist et al., 2011).
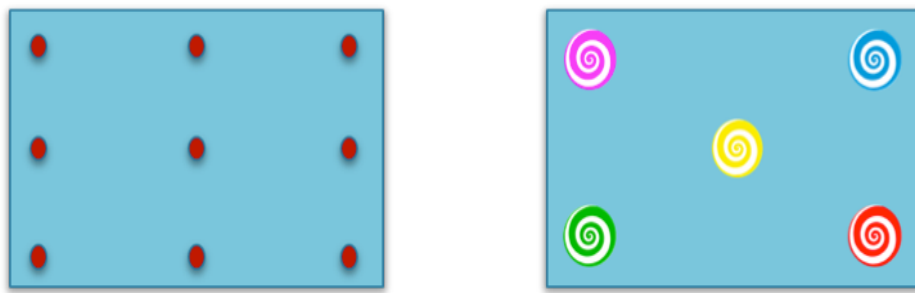
Figure 2-2 The image on the left-hand-side shows a typical 9 points adult calibration, while the image on the right-hand-side displays a typical infant 5 points calibration.

The first step for performing a typical calibration procedure is to detect the participant's eyes with the eye-tracker. Evidently, for this to happen the participant needs to be looking at the screen, which may not always be the case when testing non-compliant populations such as infants. It is essential for researchers to find a way to keep participants looking and happy during this process, which can occasionally take longer than expected (e.g., if the participant's eyes are not easily detected by the eye-tracker). For instance, when testing infants a baby-friendly video can be presented in the background while this procedure is performed (see Figure 2-3).

Figure 2-3 Before the calibration procedure starts, the participant's eyes need to be detected. When testing infants a baby-friendly video can be presented on the background while infants are placed in front of the eye-tracker and their eyes are detected.

Once the eyes are detected the calibration points are presented. Many calibration procedures display a diagram at the end of the procedure with the calibration points and the gaze data recorded during the proceeding. While this can be informative, it may not be sufficient to evaluate the calibration results accurately. Some researchers evaluate the calibration procedure by asking the participants to look at certain points in the screen and evaluating the offset between the points and the actual gaze. This is a technique that can work well when testing adults, but that cannot be used in populations that do not attend to experimental instructions. To learn more about the issues and strategies for performing and evaluating calibrations in infants continue reading through the next sections.

For the experiments presented in this thesis, infants' eyes were detected while watching a baby-friendly video (see Figure 2-3) and then they where calibrated using a standard 5 points calibration with points displayed at the corners and the centre of the screen (see Figure 2-2, right hand-side; e.g., Addyman & Mareschal, 2010; Senju & Csibra, 2008; Tummeltshammer et al., 2014; Wu & Kirkham, 2010). The calibration points where comprised of a set of colourful images with a clear central point. The experimenter was able to press a key to move onto the next calibration point, giving the infant enough time to fixate each point. The calibration procedure was repeated if the infant did not look at all calibration points or if the data quality at

the end of the procedure was not good enough. Nevertheless, if the infant seemed too tired the calibration was not repeated. Figure 2-4 shows the flow diagram with the calibration procedure that was applied for the studies described in the present thesis. Although this proceeding intended to add consistency in the evaluation of the calibration procedures, there was still certain subjectivity when deciding if a calibration had to be accepted or not.
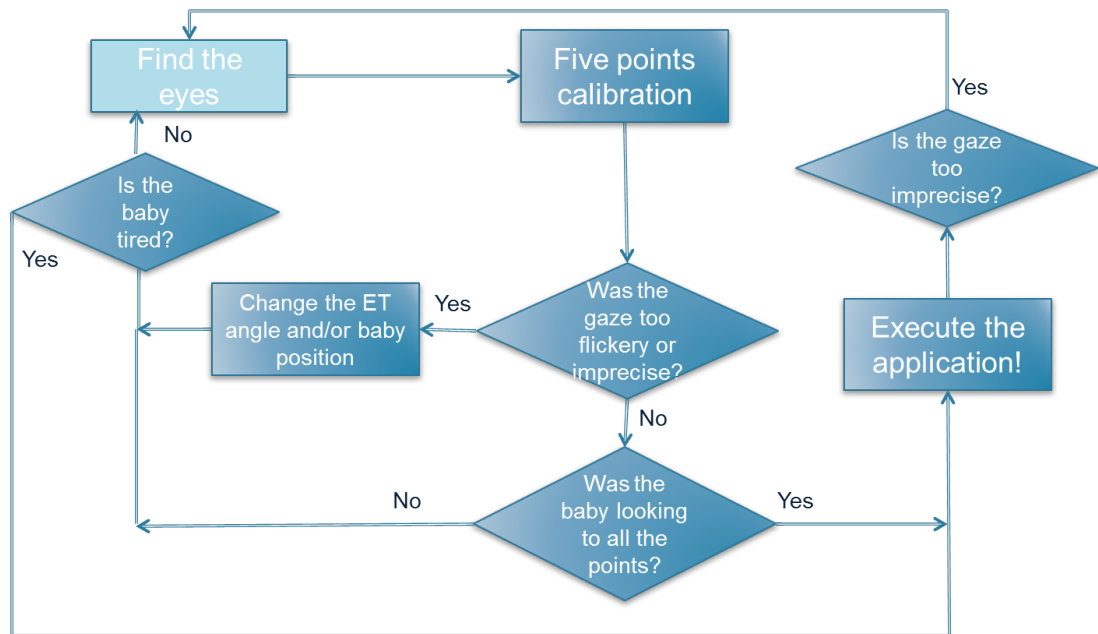
Figure 2-4 Calibration procedure. First, the infants' eyes were detected while they looked at a baby-friendly video. Infants were calibrated using a 5 points calibration. The calibration was repeated when infants did not look at the points consistently or if at the end of the procedure the data quality was not good enough. Nevertheless, if the infant seemed too tired the calibration was not repeated.

2.2.4 Room set up

The experiments described in this thesis were performed in a soundproof room with no natural light (see Figure 2-5). The space was divided into the control area and the infant area by using black panels. The control area included the computer where the MATLAB program was being executed, and the stimuli monitor, where the experimenter could see the same display the infant was seeing during the experiment. The infant area included the eye-tracker, the monitor attached to the eye-tracker, the stereo external speakers located at both sides of the screen, the video camera to monitor and record the participant, and the infant's chair.

Infants were positioned 60 cm away from the eye-tracker either sat on their caregiver's lap or strapped on a car seat, depending on what the caregiver reported was the infant's preference. If no preference was reported they were placed on the car seat. If infants became fussy while sitting on the car seat they were moved onto the caregiver's lap.
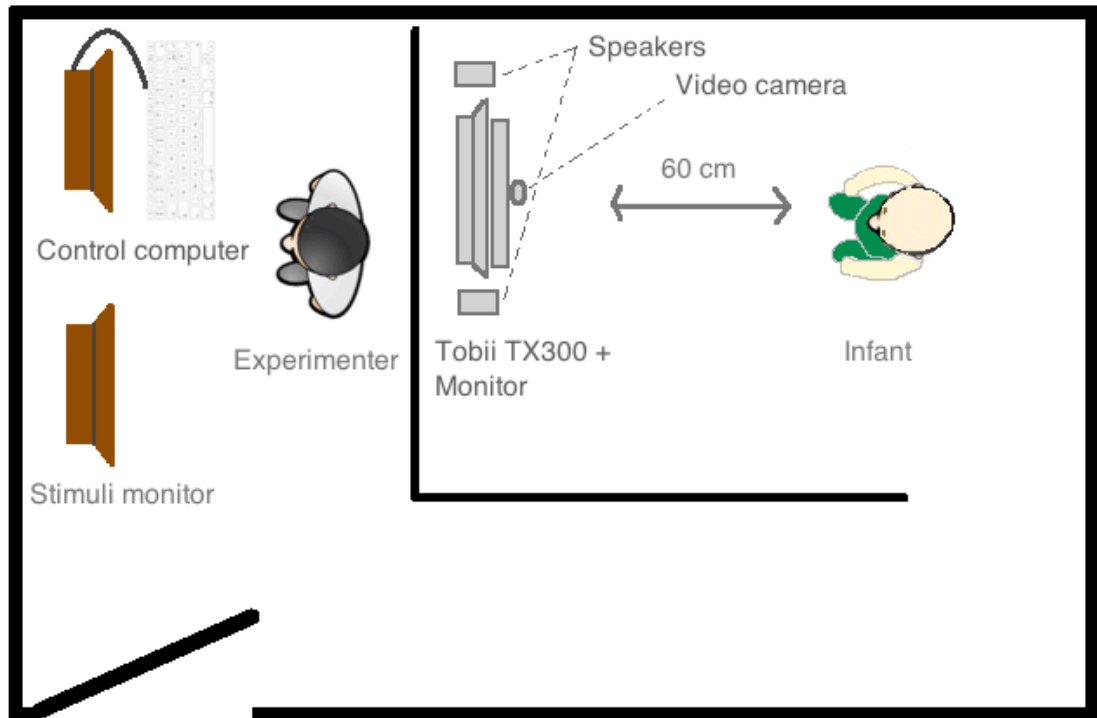
Figure 2-5 Room set-up. Infants were seated either on their mothers lap or on a high-chair at 60 cm from the eye-tracker. The experimenter stayed in a separated area of the room located behind the eye-tracker.

2.2.5 Recruitment

The infants were recruited via magazine advertisements, social networking media and flyers and added into the infant database from the Centre for Brain and Cognitive Development. Parents of the infants were called and asked to participate in the studies. Adult participants were recruited using the SONA system, where volunteers could sign themselves up for the studies. Each participant was assigned with a unique ID number. The author of this thesis tested all the subjects that participated in the experiments.

2.2.6 Experimental procedure

Infants came to the lab at times when the caregivers reported they would be happy, rested, and feed. Families were welcomed in a reception where infants could play with toys and familiarize

with the experimenter and the lab setting. Upon arrival, the experimenter gave the caregiver a brief explanation of the study and its procedure while interacting with the infant. Caregivers were instructed to refrain from interacting with the infants during the experiment. Before entering the experimental room, the caregiver was asked to sign two consent forms (one for the caregiver and one for the lab).

Upon entering the room infants were strapped into the car seat or sat on their caregiver's lap (see Figure 2-6). In the later scenario, caregivers were showed how to hold the infant to prevent them from moving excessively and were asked to maintain this position throughout the experiment. Infant's eyes were detected with the eye-tracker while they looked at a baby-friendly video displayed on the experimental screen (see Figure 2-3). Often this process involved changing the infants' position and/or the angle of the infant with respect to the eye-tracker until the eyes were accurately detected. Afterwards, infants performed a 5 points calibration (please see the previous section for more details about this procedure).

After the calibration, the experiment started. If during the experiment infants became inattentive the experimented shook a noisy toy located behind the eye-tracker (in the control area of the room) in order to attract their attention towards the experimental screen. If infants became fussy the experiment was stopped.

When the session finished, the experimenter showed the caregiver what the infant fixated during the study. Families were given baby t-shirts or bags as gifts for their participation, and their travel expenses were reimbursed.
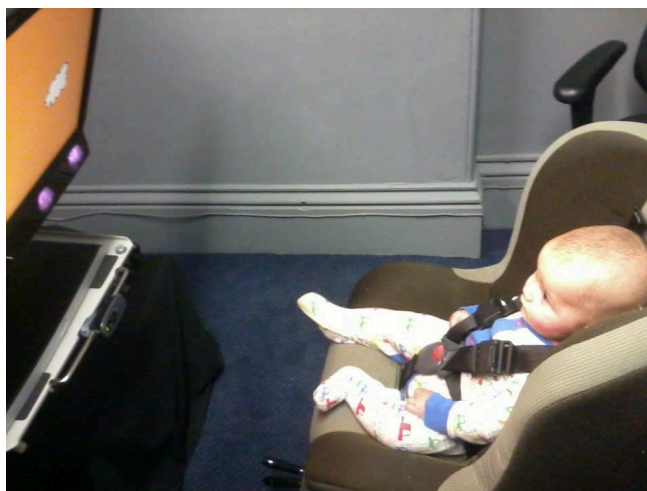
Figure 2-6 Infant participating in the study.

2.3 Measuring eye-movements in infancy: Issues and strategies for testing young infants

Recording the eye-movements of special populations such as infants or participants suffering from certain disorders (Karatekin, 2007) entail a number of issues that can seriously affect the quality of the eye-tracking data (for an extended review about eye-tracking data quality go to Chapter 3). For instance, participants suffering from Autism may present a high degree of head movement (Kelly, Walker, & Norbury, 2013), which can lower the spatial precision and the accuracy of the data. Similarly, in some populations –such as in Parkinson patients – the head motion can be constant and consistent across all the participants in the study.

Infants constitute a group that can be particularly difficult to test in eye-tracking studies (Aslin & McMurray, 2004; Aslin, 2012; Gredebäck et al., 2009; Oakes, 2012; Richardson & Johnson, 2008; Rommelse, der Stigchel, & Sergeant, 2008; Saez de Urabain et al., 2014). The issues that make this group especially challenging to test are numerous: for instance, they do not respond to experimental instructions, they do not sit still, or their eye physiologies are still underdeveloped and can considerably vary from one subject to another. All these problems can interfere with the detection process affecting the quality of the eye-tracking data and thus the experimental results (Holmqvist, Nyström, & Mulvey, 2012). In this section I will review some of these issues and describe the strategies used within this thesis to minimize the impact of these problems in data quality. Nevertheless, even after using these testing strategies in some cases data quality was far from perfect. In Chapter 3 I introduce I new method and software able to deal with variable quality eye-tracking data.

As other special populations, infants can also present a high degree of head movement, especially from 7 to 8 months of age when their locomotor abilities are rapidly improving (Adolph & Berger, 2006). For this reason systems that require a chin-rest or that provide a small head-box are not suitable for infants. Remote eye-trackers are thus the most popular alternative for infants. To avoid the use of a chin-rest these eye-trackers include some extra algorithms for the head position calculations that can interfere with the gaze estimation and hence affect spatial precision (Holmqvist et al., 2011; Kolakowski & Pelz, 2006).

Nevertheless many of the eye-tracking difficulties and quality problems in infants derive from poor calibration procedures (Duchowski, 2007; Feng, 2011; Oakes, 2012). As described earlier,

traditionally infants are calibrated using 5 points calibrations where a colourful puppet is presented in each corner and centre of the screen, while adults usually perform a 9 points calibration following a series of small dots (see Figure 2-2).

The first obvious problem occurs when the infant does not look at the calibration points when they are presented (Oakes, 2012). As a consequence, the offset calculation will be incorrect and the spatial accuracy will be affected in one or more areas of the screen. After a calibration is performed the eye-tracker software generates a diagram displaying the calibration points together with the gaze points recorded during the calibration procedure. Whilst in some cases this diagram can be informative, it still assumes the subject was gazing the calibration points when they where presented, which may not always be the case when testing infants. Frequently infants move their eyes towards a target with certain delay (even several seconds) after it was presented. A strategy to control for this issue is to play the gaze data online while the calibration is being performed.

Adult participants are frequently asked to repeat the calibration procedure until the results are satisfactory. Infants' interest in the points may decrease very quickly, and thus it may not be possible to perform as many calibrations as needed. To minimize these issues the researcher can regularly change the calibration stimuli (e.g., varying it in colours and/or shapes) and accompany the points' presentation with attractive sounds. At other times, the calibration points are presented too briefly for the infant to move his/her eyes from one point to the next. This problem is more obvious in subjects younger than 4 months, as the neural structures implicated in oculomotor control are still underdeveloped (M. H. Johnson et al., 1991; M. H. Johnson, 2011) and thus may present long disengagement latencies ("sticky fixations"; e.g., Hood & Atkinson, 1993). This tendency lessens as the infant develops and by around 4 months of age, they are able to shift their attention more rapidly and accurately from one point to another. Thus during the first months of life it is more likely infants won't follow the calibration points if they are too quick. This problem can be solved by giving the researcher the possibility to decide when the next calibration point should appear, e.g., the researcher presses a button when the infant is looking to the current calibration point.

Visual acuity does not reach adult like levels before the third year of life (Atkinson, 2000; see Chapter 1). In fact, during the first 3 to 4 months of life infants may have problems

accommodating as a function of target distance and hence they may not see objects (i.e. the calibration points) that are further away than a certain distance (Salapatek et al., 1976). Thus, for subjects younger than 3-4 months, the distance between the infant and the presentation screen needs to be considerably shorter (around 30-40 cm; Salapatek et al., 1976) than the distance that is recommended for most eye-tracking systems (around 60 cm; Holmqvist et al., 2011). Changing the recommended distance will obviously affect the quality of the data.

Another problem related to the calibration procedure is the size of the calibration points. As it was previously explained, infant calibration targets are usually bigger and visually more complex than the traditional adult targets. This means that even though it is possible to tell if the infant is foveating the correct object, we still cannot know which part of the object they are gazing. Once again this can lead to imprecise offset calculations. A way of minimizing this problem is to use calibration points that even though are bigger than the typical dots for adults, have a clear central point that is more likely to be gazed (e.g., a colourful spiral) or decreases in size down to a point (see Figure 2-2).

A final problem related to the calibration procedure is the viewing angle between the eye-tracker and the participant. Wrong angles may produce higher offsets and inaccurate binocular disparities. Usually this problem can be solved repeating the calibration procedure once again changing the participant's or the eye-tracker's position each time. Nevertheless, infants do not respond to instructions and their attention span is considerably lower than adults'. Thus, repeating this procedure may not always be feasible.

To date, even though some procedures have been developed for calibrating infants (Gredebäck et al., 2009), there is are still no standards for performing and evaluating a calibration procedure (Oakes, 2012; see also Frank, Vul, & Johnson, 2009) and thus it is up to the researcher to decide whether a calibration procedure was successful or not.

As described earlier, most of the current eye-tracking systems use corneal reflection to estimate gaze, where the pupil is detected using bright or dark pupil techniques (Holmqvist et al., 2011). There are various reasons why the glint or the pupil can be unreliably detected or not detected at all (e.g., poor lighting conditions, different eye physiologies, wrong distance and/or angle between the participant and the eye-tracker). As a consequence the data quality can be seriously affected (poor spatial precision, missing points) and hence, it can complicate the

process of identifying fixations (Holmqvist et al., 2012). Infants' eyelids can be particularly watery, especially during the first few months of life, and this can considerably interfere with the glint detection process. Usually bright pupil techniques are considered to be more accurate than dark pupil ones when dealing with certain eye physiologies like bright eyes or watery eyelids (Gredebäck et al., 2009). For this reason, when testing young infants, it is always a better choice to use eye-tracking systems that also include a bright pupil corneal reflection technique. Despite of this, for this thesis all the studies were performed using the black pupil technique, due to limitations with the eye-tracking system (Tobii TX300) that was used to test the participants.

2.4 Conclusions

In this chapter I have introduced general eye-tracking methodology and a description of the equipment and techniques used for recording the data presented on this thesis. Furthermore, in the second half of this chapter I described some of the particularities and difficulties of measuring eye-movements in infants and the strategies that were used throughout this project to overcome some of these issues. Nevertheless, whilst the actions taken during the testing sessions did improve the quality of the data gathered from the eye-tracker, it was still far from being optimal.

The studies described in this thesis are directed to study the mechanisms underlying saccadic control and fixation durations in infants. Moreover, besides presenting a series of behavioural studies I use a computational modelling approach, for which it is essential to gather as many valid fixations as possible from the eye-tracking recordings. As a result of the highly variable data quality, existing fixation detection approaches can be highly time consuming (hand-coding) or imprecise (automatic detection). In the next chapter I present GraFIX, a new method and software that attempts to address this problem by using a two-step process in which eye-tracking data is initially parsed by using velocity-based algorithms whose input parameters are adapted by the user, and then manipulated using the graphical interface, allowing accurate and rapid adjustments of the algorithms' outcome.

Chapter 3: GraFIX, a new method for detecting fixations in low and high quality eye-tracking data

The data from this chapter has been published in the article:

3.1 Introduction

Recent articles (Holmqvist, Nyström, & Mulvey, 2012; Wass, Smith, & Johnson, 2013) have highlighted the substantial impact that low quality data can have on experimental measures. A poor eye-tracking recording can affect the validity of results, and sadly, it is still not common practice to report data quality measures or deeper descriptions of the fixation detection methods used (Holmqvist et al., 2011, 2012). This can alter the viability of research results and hence lead to problems replicating previous studies.

The present thesis investigates fixation durations and the mechanisms underlying saccadic control in infants. Therefore, it is an essential part of this project to detect fixations as accurately as possible to avoid misleading results. Furthermore, the use of a computational modelling approach to analyse the data requires a high number of fixations per participant, which can be particularly challenging when testing infant populations, in part due to the low quality of their data. As it was described in Chapter 2, testing infants entails a number of difficulties (e.g., high degree of movement, unreliable eye detection, low spatial precision) that result in highly variable data quality and render existing fixation detection approaches highly time consuming (hand-coding) or imprecise (automatic detection). To address this problem I developed GraFIX, a novel semi-automatic method consisting of a two-step process in which eye-tracking data is initially parsed by using velocity-based algorithms whose input parameters are adapted by the user, and then manipulated using the graphical interface, allowing accurate and rapid adjustments of the algorithms' outcome. The present algorithms (1) smooth the raw data, (2) interpolate missing data points, and (3) apply a number of criteria to automatically evaluate and remove artifactual fixations. The input parameters (e.g., velocity threshold, interpolation latency) can be easily manually adapted to fit each participant. Furthermore, the present software

includes visualization tools that facilitate the manual coding of fixations. I assessed this method by performing an inter-coder reliability analysis in two groups of infants presenting low and high quality data, and compared it with previous methods. Results revealed that the present approach gives rise to more reliable and stable measures in low and high quality data.

To download the code, go to: http://sourceforge.net/projects/grafixfixationscoder/

3.2 What is data quality and why it is so important

The raw data recovered from any eye-tracker includes a time stamp and the x and y coordinates for one eye (monocular systems) or both eyes (binocular systems). Fixations can be identified when these coordinates are relatively stable in a point (and hence the eyes' velocity, defined as the rate of change in x and y coordinates from one gaze point to the next, is low); whereas saccades are flagged when the x and y coordinates are more variable in the scene and the eyes' velocity exceeds a given threshold (see Figure 3-2). Additionally, other types of eye-movements can be detected in the raw data, such as smooth pursuit (Larsson, Nyström, & Stridh, 2013), or blinks (e.g., Morris, Blenkhorn, & Zaidi, 2002). In cases when the data quality and the sampling frequency are very high, it is even possible to identify very short fixational eye-movements such as microsaccades, glissades or tremor (Nyström & Holmqvist, 2010).

The quality of the raw data generated by the eye-tracker may vary depending on many different factors such as the eye-tracker model and manufacturer, the eye physiology, the calibration procedure, the position of the participant relative to the eye-tracker, the degree of head motion (Holmqvist et al., 2011, 2012), or even ethnicity (Blignaut & Wium, 2014). The term "data quality" entails different aspects affecting eye-tracking data, but not all these aspects will necessarily affect fixation detection equally.

Low data quality can have major effects on both spatial and temporal accuracy of gaze measurements (see Figure 3-1). Spatial accuracy or offset refers to the difference in space between the detected gaze and the real gaze, and can be an important issue when analysing areas of interest (AOIs; Holmqvist et al., 2012). Apart from the vertical and horizontal accuracy that the eye-tracking systems report, aspects such as binocular disparity should also be taken into account, especially when studying special populations. For instance, we know that binocular disparity in young infants may be markedly larger than in adults (Appel & Campos,

1977; Yonas, Arterberry, & Granrud, 1987). However, it is also common to find infant data with very large disparities as a consequence of poor calibrations or incorrect angles between the eye-tracker and the participant. Often, it is possible to minimize the effects of poor accuracy in AOIs analysis by simply enlarging the regions of interest, or by quantifying and correcting the offset for each participant (Holmqvist et al., 2011, 2012). Frank, Vul, & Saxe (2012) designed an offline procedure to correct errors in calibration in order to increase the accuracy and include a measure that evaluates it. They presented their participants (infants from 3 to 30 months) with calibration points that appeared during the experiment and that were subsequently used to correct the offset.
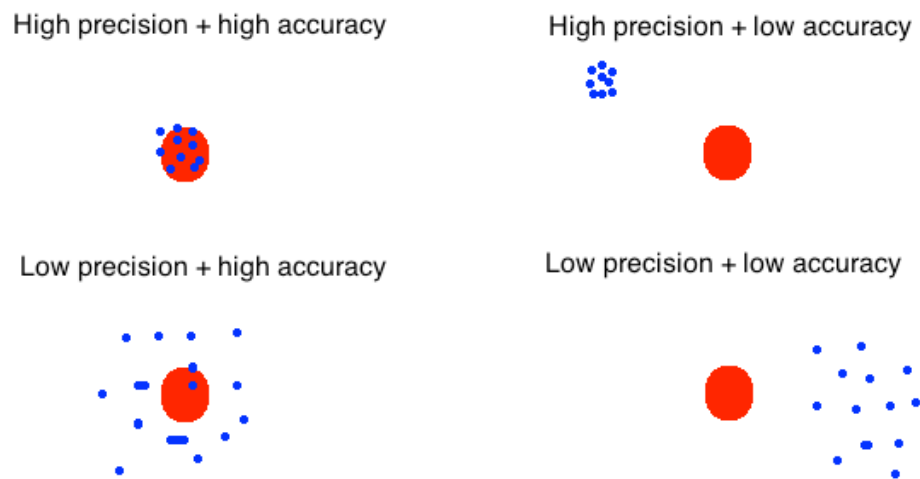
Figure 3-1 Spatial precision and accuracy. Spatial accuracy or offset refers to the difference in space between the detected gaze and the real gaze. On the other hand, spatial precision is defined as the consistency in detecting and calculating gaze points. The red point represents the visual target, while the blue dots are the gaze points.

Nonetheless, a data set with a large offset can also present high spatial precision and thus still be suitable to detect fixations accurately. We refer to spatial precision as the consistency in detecting and calculating gaze points (see Figure 3-2). Data sets with relatively low spatial precision will present higher gaze velocities as a result of noise in the data, and this will complicate the process of detecting fixations accurately. Precision can be affected by individual factors that vary on a participant basis, (such as different eye physiologies, or the position of the subject relative to the eye cameras), environmental factors that change according to the experimental design (such as the lighting conditions of the room where the participants are

being tested) or the eye-tracking hardware and software (Holmqvist et al., 2011, 2012). The default spatial precision for a particular eye-tracker can be calculated by using an artificial eye. Additionally, there are a number of methods to calculate spatial precision (Holmqvist et al., 2011, 2012), such as the root mean square (RMS) of inter-sample distances (commonly used by manufacturers) or the standard deviation, which measures the dispersion of each sample from a mean value. To see the effect that spatial precision has on the detection of fixations by a velocity-based algorithm, see Figure 3-2.
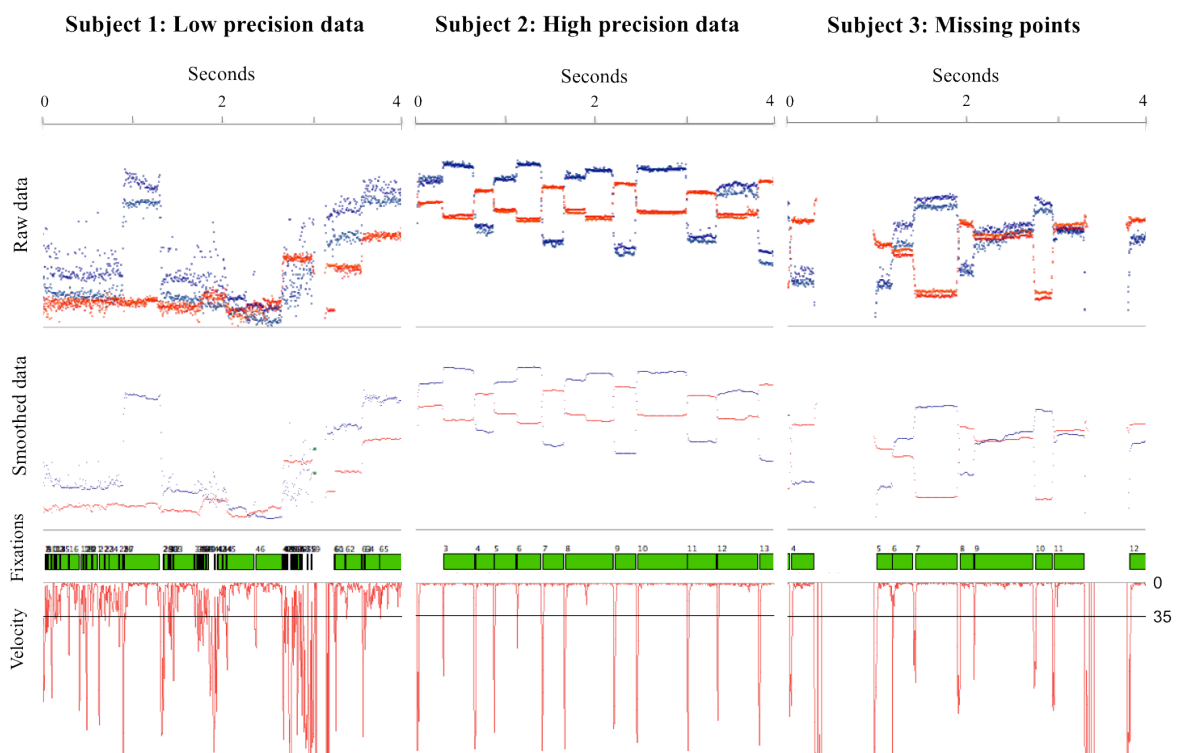
Figure 3-2 Data from 3 infant subjects recorded with a Tobii TX300 at 120 Hz. The first and the second row show the raw and the smoothed data respectively. The red and the blue lines represent the data from the left and the right eyes respectively. The third row displays the fixations detected by a velocity-based algorithm (velocity threshold = 35º/sec), and the fourth the velocity calculated from the smoothed data. Subject 1 shows low precision data, which is very common in young infants. As a consequence, the fixations-parsing algorithm detected a number of physiologically implausible artifactual fixations. Subject 2 displays high precision data from infants. Although the algorithm was more accurate, due to the high velocity threshold, it merged together fixations that had short saccades in between (e.g., fixation 8). Subject 3 shows a participant that presents frequent missing data points.

Likewise, data loss (often a consequence of unreliable detections of the pupil or the cornea reflection) is another issue that can considerably affect fixation detection (Holmqvist et al., 2011, 2012; Wass et al., 2013). This also is highly dependent on individual and environmental factors

as well as on the eye-tracker hardware and software. These individual variations in the recordings from different subjects can lead to very different levels of data quality (accuracy, precision and data loss), even when participants have performed the same study under the same experimental conditions. This can in fact be a problematic issue when trying to standardize the procedure to analyse the eye-tracking data: do we use exactly the same protocol and values to analyse the data regardless the noise that each participant presents? Or would it be more appropriate to adapt it somehow?

3.3 Previous methods for detecting fixations

Fixations can be detected by an algorithm or by a person on a basis of the visual inspection of the raw eye-tracking data. Many eye-tracking manufacturers already provide smoothing and event detection tools. However, what these algorithms do to the data may still be unclear for many users, especially for those not particularly familiar with event detection techniques. Additionally, when the user chooses arbitrary input parameters without considering issues like data quality, the sampling frequency, or other aspects of the experimental design (e.g., it is not the same to detect small fixational eye-movements in reading research or to detect saccades in infants), the detection results can be gravely affected and hence the validity of the experimental outcomes can be questioned (Holmqvist et al., 2012).

Event detection algorithms can be classified into two main groups: *dispersion and duration* algorithms; and *velocity and acceleration* algorithms (for more detailed reviews see Holmqvist et al., 2011). Dispersal-based algorithms use a minimum fixation duration threshold (e.g., 50 ms) and the positional information (dispersion) of the eye-tracking data in order to decide if consecutive points belong to the same fixation, in which case they are grouped together. If not they are assumed to be a saccade or a missing point. Dispersion can be measured according to different metrics (Blignaut, 2009) such as the distance between the points in the fixation that are the farthest apart (Salvucci & Goldberg, 2000), the distance between two random points in a fixation (e.g., Shic, Scassellati, & Chawarska, 2008), the distance between two points at the centre of a fixation (e.g., Shic et al., 2008), the standard deviation of x and y coordinates (e.g., Anliker, 1976), or a minimum spanning tree of the points in a fixation (e.g., Salvucci & Goldberg, 2000). Currently it is possible to find a number of commercial (e.g., SMI BeGaze) and non-commercial implementations for these algorithms (e.g., Salvucci & Goldberg, 2000), which are mostly used to parse low-sampling rate data (< 200 Hz). On the other hand, the algorithms from

the second group calculate the velocity and/or acceleration for each point in order to detect events on the data. Velocity-based algorithms in particular, flag all the points whose velocity are over a threshold (e.g., 10-70 °/sec) as saccades, and define the time between two saccades as a fixation. Once again there are a number of commercial (e.g., Tobii, EyeLink) and non-commercial variations (Nyström & Holmqvist, 2010; Smeets & Hooge, 2003; Stampe, 1993; Wass et al., 2013) for this type of event detection algorithms. These algorithms are commonly used in data collected at high sampling rates (e.g., >500 Hz). All these algorithms are very sensitive to noise and unless the collected data has a very high spatial precision the results will include a number of artifactual fixations.

The use of event detection algorithms implies decisions about which thresholds should be selected in order to obtain optimal results. However, how these decisions are made, the range of parameters that can be manipulated, and whether they are reported in published papers is not yet standardized making it difficult to compare or replicate results from different studies. Komogortsev, Gobert, Jayarathna, and Gowda (2010) compared the performance of different velocity and dispersal based algorithms and presented a standardized scoring system for selecting a reasonable threshold value (velocity or dispersion threshold) for different algorithms. Nevertheless, this article did not take into account the individual differences in data quality across participants and/or trials.

Most researchers tend to use the same input parameters for all the participants paying very little attention to these variations in data quality and the effects that selecting different thresholds may have on the data from different participants. Nyström and Holmqvist (2010) presented a new velocity-based algorithm for detecting fixations, saccades and glissades using an adaptive, data driven peak saccade detection threshold that selects the smallest velocity threshold that the noise level in data allows. The use of thresholds was motivated by physiological limitations of eye-movements. These algorithms already highlighted the importance of adapting the input parameters to different levels of noise. More recent approaches have attempted to remove the need of arbitrary thresholds by applying non-parametric methods that use cluster analysis for detecting fixations and saccades (König & Buffalo, 2014). Nevertheless these methods still do not solve the problem of accurately detecting fixations in data sets with relatively higher levels or noise, such as those from infants.

Wass and colleagues (2013) analysed standard dispersal based fixation detection algorithms and showed how results were highly influenced by inter-individual variations in data quality. Additionally, they went a step further to solve these problems developing new detection algorithms that include a number of post-hoc validation criteria to identify and eliminate fixations that may be artifactual. These algorithms already exclude many artifactual fixations that where included when using other velocity-based detection algorithms. However, any automatic approach for detecting fixations in data with a certain degree of noise is likely to produce artifactual fixations that are erroneously calculated and/or fixations that are not detected at all.

An alternative to using automatic algorithms is to hand-code eye-movements based on the visual inspection of the data. For instance, developmental psychologists have traditionally studied infants' attention and eye-movements by videotaping participants and hand-coding the direction of the gaze post-hoc (e.g., Elsabbagh et al., 2009; Harris, Hainline, Abramov, Lemerise, & Camenzuli, 1988). Also, it is a common practice when analysing the data from head-mounted eye-trackers to replay the scene and eye videos frame by frame and make annotations of the onsets and offsets of fixations on a separate file (e.g., Tatler et al., 2005). Obviously these techniques are highly time consuming and can limit the number of subjects that a researcher is able to test and code.

With a view to avoid these problems, some researchers suggest to exclude all the participants whose spatial precision is over a predefined threshold (Holmqvist et al., 2011, 2012). This way the use of automatic algorithms should be relatively safe, though not perfect. However, excluding participants according to their data quality is a luxury that not every study can afford, like is the case of the infant studies presented in this thesis. As previously explained, the data quality for many experiments studying high-cost populations such as infants or special populations to whom access is limited may be consistently low. Using data quality as an inclusion criterion might result in many participants (or even all) being excluded. In cases of special populations, the data can be too valuable to be discarded.

3.4 Introducing GraFIX

GraFIX implements a two-step approach where fixations are initially parsed by using an adaptive velocity-based algorithm, before the algorithm's outcome is hand-moderated using a graphical interface (see Figure 3-3). As it was described before, previous methods for detecting

fixations have adopted either a purely automatic approach or manual coding. Due to the high variability in data quality across participants and even within a single participant (e.g., as a result of moving the head throughout the eye-tracking session), the automatic detection algorithms can be remarkably unreliable. On the other hand, current hand-coding methods (e.g., coding fixations looking at the videos frame by frame) can be extremely time consuming and in some cases imprecise if coding low quality data sets.
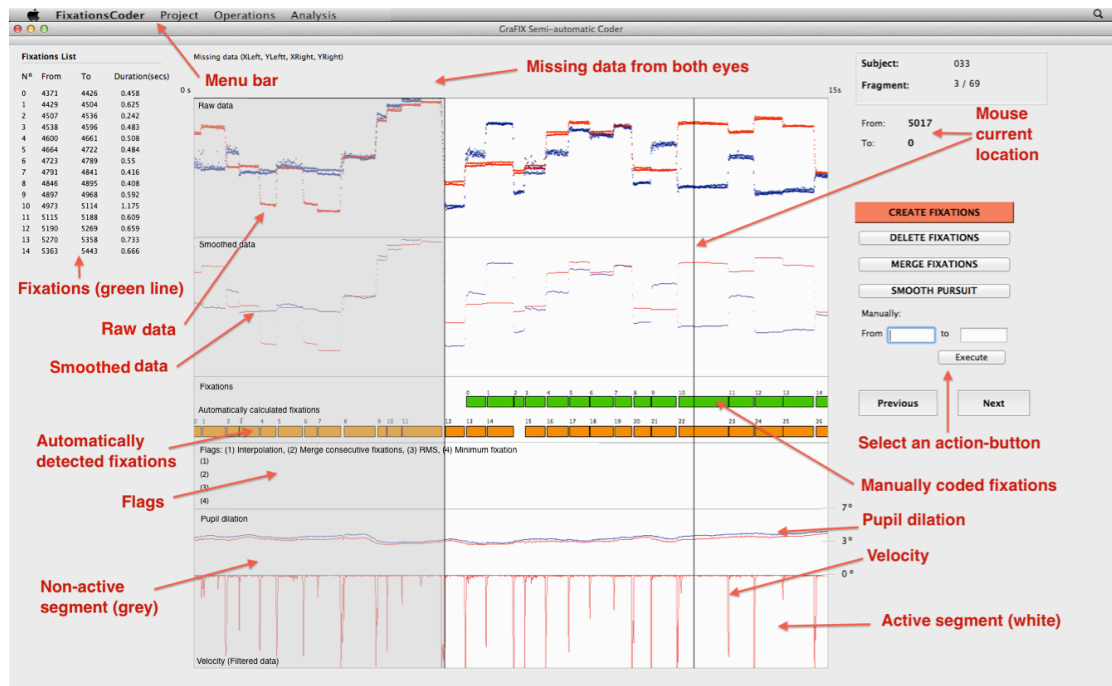
Figure 3-3 GraFIX application overview. This is the main window where the user is allowed to manipulate fixations. The eye-tracking data is displayed in its different formats (raw and smoothed data). Top section of window: X and y coordinates are presented on the vertical axis and time on the horizontal axis. Fixations can then be identified when both x and y coordinates do not present any displacement in the vertical axis, and saccades when there is a vertical displacement between two fixations accompanied by a velocity peak. Automatically detected fixations (orange rectangles) are displayed aligned with hand-moderated fixations (green rectangles), which are the ones that can be manipulated by selecting an action on the right-side of the screen (create, delete or merge fixations; code them as smooth pursuit), and mouse-clicking on them. Further, GraFIX allows defining and indicating the sections where the user is interested in detecting fixations by displaying them on white or grey.

The proposed method combines these two approaches together in order to detect fixations in a rapid manner, and obtain a fixation distribution with the lowest possible degree of noise. The current fixation detection algorithm includes a number of input parameters that can be easily adapted on a participant basis. Additionally it implements three post-hoc validation criteria that fix or remove many of the artifactual fixations generated by the velocity-based algorithm. The ultimate aim of adapting the input parameters and applying certain post-hoc validation criteria is

to obtain the most accurate outcome by the algorithms alone and thus reduce the hand-coding time during the subsequent step. Once the fixations have been automatically estimated, the researcher can evaluate them and fix those that were not accurately detected using the GraFIX graphical hand-coding tool.

3.4.1 Software description

GraFIX is a multiplatform application developed in C++ and QT frameworks that makes use of Armadillo C++ linear algebra library. It works with any binocular or monocular eye-tracking system that can record raw X/Y gaze coordinates, including SMI, EyeLink or Tobii eye-trackers.

3.4.1.1 User interface

GraFIX displays the eye-tracking coordinates in the raw and the smoothed data boxes (see Figure 3-3). It presents the x and y coordinates on the vertical axis and time on the horizontal axis. Fixations can then be identified when both x and y coordinates do not present any displacement in the vertical axis, and saccades when there is a vertical displacement between two fixations accompanied by a velocity peak (see Figure 3-3, velocity box). Occasionally our eyes move to smoothly pursue an object in the visual scene, and this type of eye-movement can be identified when there is a regular increasing or decreasing displacement in the vertical axis with low velocity and acceleration (not present in Figure 3-3).

The main window consists of a number of boxes where the eye-tracking data and the application's output is visualized in addition to all the necessary buttons used to manipulate fixations (see Figure 3-3). A brief explanation of these different components is included below.

- Menu Bar. The menu bar allows access to different dialogs such as *Project configuration*, *Visualizations* or *Automatic Detection of Fixations* dialogs.
- Raw data. The horizontal axis represents time while the vertical axis displays the position of subsequent gaze data-points (Each data-point consists of the x and y coordinates for the right and/or the left eyes.). For binocular systems the right and left eye coordinates are displayed (x-right = Red; x-left = Orange; y-right = Dark blue; y-left = Light blue). In case the system is monocular, the data from the second eye will simply not be displayed.
- Smoothed data. As for the raw data, the horizontal axis represents time while the vertical axis displays the position of subsequent smoothed gaze data-points. Once the data is smoothed the resulting x and y coordinates are displayed.

- Pupil dilation. In cases where the eye-tracking system provides pupil dilation data, it is displayed in this box. Additionally, the average pupil dilation is calculated for each fixation.

- Velocity. The horizontal axis represents time while the vertical axis displays the velocity for subsequent gaze data-points.

- Missing data from left and right eyes. If there is a data-point missing for either of the eyes it is displayed in this box.

- Automatically detected fixations. The orange boxes represent the fixations that are automatically detected by the automatic-detection algorithms.

- Flags from the automatic-detection algorithm. During the automatic detection the data is interpolated and the post-hoc validation criteria is applied. The flags indicate which data-points were affected by different algorithms.

- The cursor's current location. A horizontal line is displayed at the cursor's location facilitating the accurate coding of fixations' onsets and offsets. Additionally, the *From* label indicates the line in the raw file where the cursor is located.

- Manually coded fixations. The green boxes represent the fixations that can be manipulated manually. Each fixation has a fixation number that matches the numbers from the first column at the fixations list.

- Fixations list. This is a list of the fixations that are being manually coded (green boxes). The first column shows the fixation number, the second displays the line in the raw file where the fixation starts, the third displays the line in the raw file where the fixation ends, and finally the fourth column reveals the fixation duration in seconds.

- Select the action button. In order to perform an action the corresponding option has to be selected. Once it is selected one will need to click and drag the cursor in order to create, delete or merge fixations. Moreover, it is possible to target smooth-pursuit fixations by selecting this option and dragging the cursor on top of the fixations that need to be coded. In order to add more precision to the task it is possible to enter in the textboxes the start and the end of the fixation that needs to be manipulated, and press *Execute.* This will execute the action currently selected using the specified start and end points.

- Non-active segment. The portion of the screen in grey corresponds to the tasks in the experiment that do not need to be coded.

- Active segment. The portion of the screen in white corresponds to the tasks in the experiment that need to be coded.

3.4.1.2 Input files

The *raw input file* is a csv file separated by commas with the following columns:

[(1) Time in microseconds; (2) zeros; (3) relative gaze-point X left eye; (4) relative gaze-point Y left eye; (5) relative gaze-point X right eye; (6) relative gaze-point Y right eye; (7) pupil diameter left eye (optional); and (8) pupil diameter right eye (optional)]

If the eye-tracking system is monocular all the columns corresponding to the second eye should be substituted with *-1s.*

The *segments input file* indicates which parts of the experiment need to be coded. If the whole data file is the subject of interest, this file does not need to be included.

The segments input file is a csv file separated by commas with the following columns:

[(1) Segment id, which is an unique number for each row; (2) Row number in the raw file where the segment starts; (3) Row number in the raw file where the segment ends.]

3.4.1.3 Output files

File smooth_[subject number].csv This file is created when the data is smoothed. Each row corresponds to a data point from the raw file, thus both files have the same length. This file is a csv file separated by commas with the following columns:

[(1) Time in microseconds; (2) zeros; (3) smoothed X coordinate; (4) smoothed Y coordinate; (5) velocity; (6) Is saccade flag: 0/1; (7) Is interpolated flag: 0/1; (8) post-hoc merge adjacent fixations flag: 0/1; (9) post-hoc RMS flag: 0/1; (10) post-hoc minimum fixation flag: 0/1]

File fix_auto_[subject number].csv This is created when fixations are automatically estimated. Each row contains the information for one fixation. This file is a csv file separated by commas with the following columns:

[(1) Row from the raw file where the fixation starts; (2) row from the raw file where the fixation ends; (3) duration in seconds; (4) Average X coordinate; (5) Average Y coordinate; (6) RMS; (7) Is smooth pursuit flag: 0/1; (8) Average pupil dilation]

File fix_all_[subject number].csv This file is generated when the first fixation is created and it is updated every time a fixation is manipulated. It is a csv file separated by commas with the following columns:

[(1) Row from the raw file where the fixation starts; (2) row from the raw file where the fixation ends; (3) duration in seconds; (4) Average X coordinate; (5) Average Y coordinate; (6) RMS; (7) Is smooth pursuit flag: 0/1; (8) Average pupil dilation]

3.5 Automatic detection of fixations

The first action for the two-step approach is to parse the eye-tracking data using adaptive velocity-based algorithms. The present automatic detection algorithms (1) smooth the raw data, (2) interpolate missing data points, (3) calculate fixations using a velocity-based algorithm, and (4) apply a number of post-hoc validation criteria to evaluate and remove artifactual fixations (see Figure 3-4).
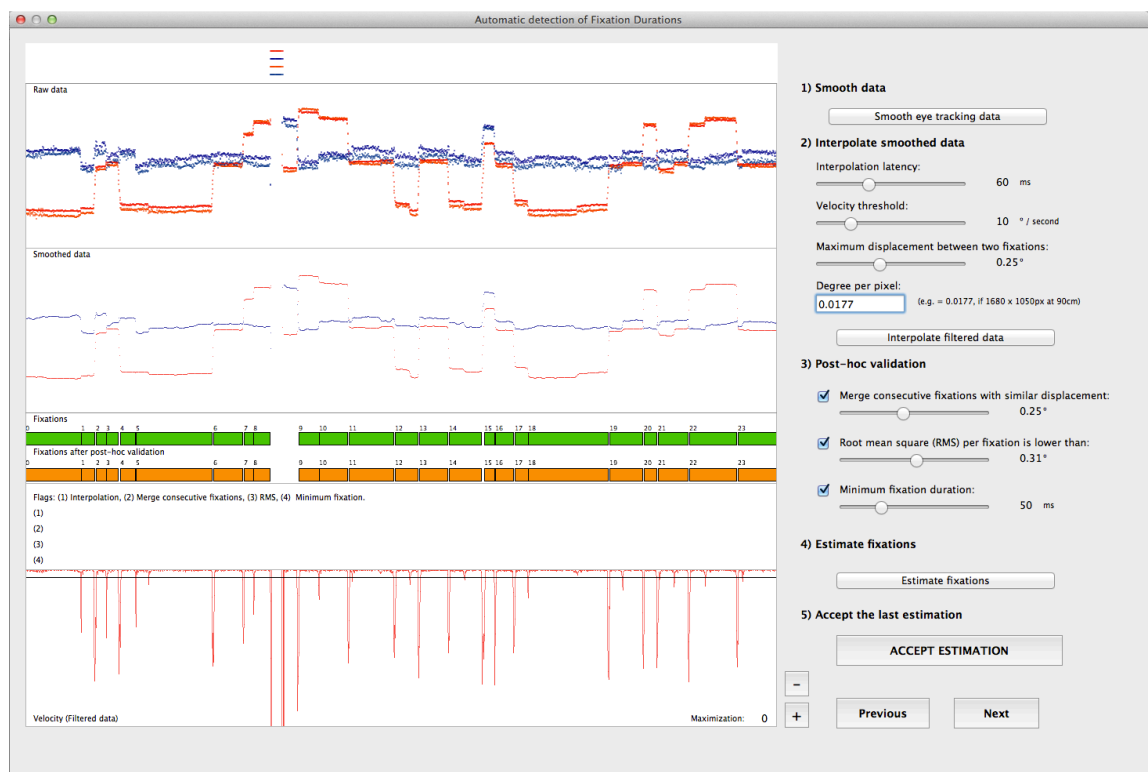
Figure 3-4 GraFIX Automatic detection of fixations. This screen displays the input parameters for the automatic detection. It is possible to adapt the parameters by simply changing their values from the sliders. By pressing *Estimate fixations* GraFIX executes the detection algorithms and displays the results on the orange rectangles. Flags indicating which post-hoc validation criterion was executed are also displayed. This process is relatively fast and thus allows multiple, and easy adjustments of the parameters. Once the user is satisfied with the results, the detection can be accepted by pressing *Accept estimation*. This will copy the automatically detected fixations (orange) on the hand-modulated fixations area (green).

The input parameters (e.g., velocity threshold, interpolation latency) can easily be manually adapted to fit the data from different participants that present different levels of data quality .The objective of these algorithms is to obtain the most accurate fixation detection for each participant and thus reduce the amount of time spent manually correcting fixations in the subsequent step.

3.5.1 Smoothing the data

GraFIX uses a bilateral filtering algorithm in order to decrease the noise levels from the raw data. The current version of the algorithm is based on previous implementations (Durand & Dorsey, 2002; Frank et al., 2009) that average the data for both eyes and eliminate the jitter while preserving saccades.

If only one of the eyes is detected, GraFIX allows the user to decide if the detected eye will still be smoothed or the sample should be excluded. Previous researches have argued that when one eye is not detected, the data from the other eye may be unreliable (Wass et al., 2013). However, when the eye-tracking data comes from special populations such as infants, the fact that one of the eyes is not detected does not necessarily mean that the sample from the other eye is inaccurate. For instance, it can be the case that the infant is simply occluding one of his/her eyes with his/her hand causing difficulties for the accurate detection of both eyes. Occasionally these missing points could lead to inaccurate results regardless the inclusion or exclusion of the data (e.g., if one eye was occluded during a fixation). In the interest of obtaining the highest number of fixations, for this thesis I also include the fragments in the data where only one eye was detected.

3.5.2 Interpolating smoothed data

Occasionally a data set will present a number of short gaps of missing data where the eyes are not accurately detected. These gaps can range from 1 to even 150 milliseconds and may severely affect the detection of fixations. To address this problem I included an algorithm that interpolates short segments of missing data (see Figure 3-5).

Clearly it is not necessary to interpolate every single segment of missing data: the algorithm will only fill the gaps that are shorter than a given threshold and that belong to a fixation, and not a saccade. The *interpolation latency* is the longest period of missing data that will be interpolated. This value may change depending on our data quality or the experimental design. For instance,

for the studies included in this thesis I used a interpolation latency of 60 ms since the shortest fixations that could be manually coded in the data were never longer than this value. Nevertheless, other researchers have also used interpolation latencies as long as 150 ms, arguing that the minimum time taken to program a saccade is 100-130 ms and hence this way it is possible to avoid interpolating through a complete saccade-fixation-saccade sequence (Wass et al., 2013).
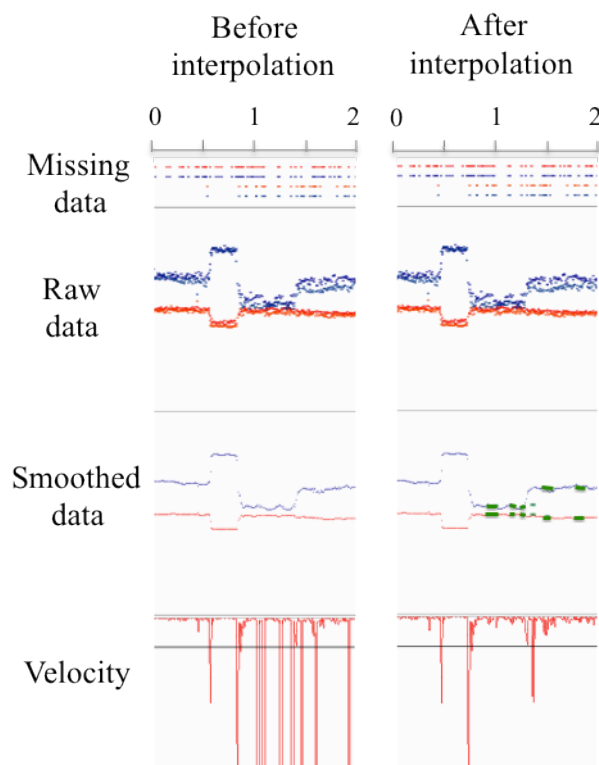
Figure 3-5 Interpolation. The sample from the left shows a fixation that has missing points. As a consequence, there are velocity peaks in the middle of the fixation. On the other hand the picture from the right shows the same sample after interpolating the missing points (the green dots on the smoothed data are the interpolated points). In this case the velocity calculation looks significantly cleaner.

First, the interpolation algorithm flags all the samples in the data whose velocities lie over the velocity threshold as saccades, and the data segments between two saccades are targeted as fixations. The *velocity threshold*, which is also used and described in the forthcoming steps, can be set to meet the requirements of different data sets.

Secondly, when a gap longer than the interpolation latency is found, the algorithm finds the subsequent and previous fixations and calculates the *mean Euclidean distances* from a central point for each of them. When the difference between both Euclidean distances is smaller than the "*Maximum displacement between the two ends of a fixation*" value, the gap is interpolated. It is important to determine the correct "*degree per pixel*" parameter (in visual angle) in order to convert the degrees to pixels properly (the pseudocode for this algorithm is described in Appendix A).

3.5.3 Velocity threshold and fixation detection

As in previous velocity-based detection algorithms (e.g., Nyström & Holmqvist, 2010; Smeets & Hooge, 2003; Wass et al., 2013), all the samples whose velocities lie over a certain threshold are flagged as saccades and the data segments between two saccades are targeted as fixations. Choosing the right velocity threshold highly depends on the characteristics of the data that is being analysed, or on how short are the saccades that need to be detected (Holmqvist et al., 2011). For instance, low sampling rates will present some limitations when detecting very fast eye-movements such as microsaccades. Previous research has shown that saccades smaller than 10° cannot be detected with systems with a sampling rate of 60Hz and lower (Enright, 1998). This is because the peak velocity calculation may not be accurate enough if only very few samples of a saccade were recorded. The lower the sampling frequency, the lower the calculated velocities for short saccades will be. In these cases a fixation detection algorithm would merge the fixations before and after an undetected saccade and this would result in longer artifactual fixations. In order to reliably detect small saccades and reduce noise it is recommended to use high sampling frequencies and lower velocity thresholds.

Nevertheless, data from special populations such as infants can still represent a challenge due to its low quality. Thus when the noise levels are higher, the velocity threshold should be increased accordingly in order to decrease the number of "false positive" saccades. This leads to the question of whether participants featuring different levels of data quality could be analysed together using the same input parameters. As suggested in previous research (Wass et al., 2013), different levels of noise can seriously affect the outcome from fixation detection algorithms. There are two opposing approaches that have been traditionally used to minimize this issue. Some researchers prefer to use exactly the same input parameters for all the participants (such as the velocity threshold) regardless the level of noise each participant

presents (Wass et al., 2013). Usually these parameters are set to fit the requirements for participants with high levels of noise. Consequently, the velocity threshold can be too high to detect relatively fast saccades, which could ultimately lead to the detection of long artifactual fixations. Moreover, these saccades would still remain undetected in very low precision data sets even after lowering the velocity threshold. On the other hand, it is possible to adapt the input parameters according to the level of noise on a participant-by-participant basis (e.g., Nyström & Holmqvist, 2010). Although the use of different velocity thresholds can lead to different outcomes from the detection algorithms, it is possible to remove a number of artifactual fixations (this is especially important in noisy data-sets) by also adapting the input parameters for the post-hoc validation criteria (explained below) and manually fixing the fixations that were not detected correctly. With current systems it will never be possible to remove all the noise, but researchers can minimize its effects and obtain the cleanest result for both low and high quality data.

3.5.4 Post-hoc validation

Once the data is smoothed and interpolated, fixations can be automatically calculated by executing the velocity-based algorithms described below. Additionally it is possible to apply a number of post-hoc validation criteria in order to manipulate the algorithm's outcome and obtain the most accurate results. All the input parameters (including the velocity threshold described in the previous section) can be adapted on a participant-by-participant basis. This way it is possible to personalise the detection process to different subjects with different levels of data quality and reduce, to a certain degree, the insertion of noise by the algorithms.

3.5.4.1 Merging adjacent fixations with similar location

Occasionally, the detection algorithms can break down fixations with low precision into a number of smaller fixations. This is because one or more samples from a fixation may present a velocity peak that is higher than the defined velocity threshold and hence be mistakenly flagged as a saccade (see Figure 3-6, first column). As such, even fixations with high spatial precision can have few samples that were not accurately recorded and consequently will generate the same peaks in velocity as saccades. Therefore, this issue is of special concern in low spatial precision data sets, even though it can affect any eye-tracking recording.
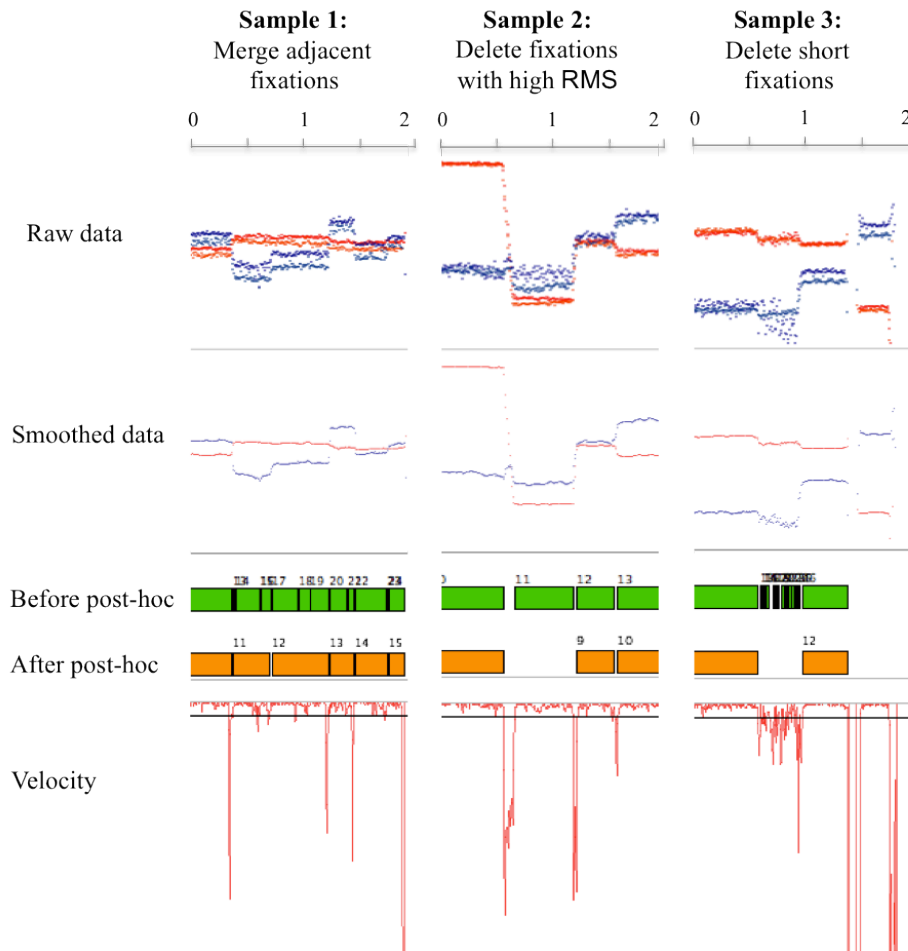
Figure 3-6 Post-hoc validation examples. We present different segments of data from one infant collected with a Tobii TX300 system at 120 Hz. The velocity threshold was set to 9°/sec. Sample 1 shows how as a consequence of the low precision in the data there are velocity peaks that pass the velocity threshold and hence are erroneously flagged as saccades. Merging adjacent fixations with similar location helps overcoming this problem. Sample 2 shows fixations with low precision that are deleted if they overcome a given RMS threshold. Sample 3 shows a number of very short artifactual fixations that are detected as a consequence of the poor precision in the data. The minimum fixation post-hoc validation criterion deletes all the fixations shorter than a given threshold.

The present post-hoc validation algorithm merges adjacent fixations that are close in time and space (see Figure 3-6, first column). First the algorithm will select fixations that have a gap in between of less than 50 ms, and then it will evaluate if these fixations are also close enough in space. For this purpose the user is able to set up a threshold and define the maximum distance in degrees between two fixations. Thresholds for merging should depend on the sampling frequency, data quality, stimulus spacing, density of visual information, or even on the research question (e.g., studies investigating microsaccades may require very low thresholds). Participants that present low spatial precision, for instance, will require higher distance-

thresholds than those with high precision samples. In high-precision data recorded with a high-sampling rate such small displacements between fixations may actually represent microsaccades and therefore post-hoc merging of fixations should be avoided.

3.5.4.2 The root mean square (RMS) of inter-sample distances per fixations does not exceed a threshold

Spatial precision is crucial for a correct detection of fixations (Holmqvist et al., 2011, 2012), in particular if the research interests rely on the study of very small fixational eye-movements such as microsaccades. Further, as explained in the sections above, in populations such as infants spatial precision can also be a major issue as a result of the many practical and technical difficulties of testing them (see Chapter 2).

The current post-hoc validation criteria calculates the RMS for each fixation for both the vertical and horizontal axis together, and deletes all the fixations with a value above a given threshold (in degrees of visual angle; see Figure 3-6, second column). In high precision data sets the RMS can be smaller than 0.10º, while in low precision data sets it can be higher than 0.20º.

3.5.4.3 Minimum fixation duration

Especially in data sets with low spatial precision, the detection algorithms may mistakenly generate very short fixations. This is because the dispersion for some of the samples belonging to a fixation may be high enough to generate velocity peaks over the given threshold that will be flagged as saccades (see Figure 3-6, third column).

In order to avoid this problem GraFIX can delete all the fixations with a duration under a given threshold (e.g., 100 ms). In the event that a data set is very noisy the minimum fixation threshold should be set higher. This does not necessarily mean that it is not possible to find short fixations in our dataset: in case there are clean short fixations, they can also be manually coded after the automatic detection of fixations.

3.5.5 Defining the right parameters for detecting fixations

GraFIX is able to run the fixation detection algorithms and visualize the results very quickly. For instance, 30 minutes of data at a sampling frequency of 300 Hz can be parsed in less than a second (the processing speed also depends on the capabilities of the machine where the application is being executed). This permits the user to adjust the input parameters and evaluate their effect on the data in a rapid manner. In fact, it can be immensely helpful (for

novice users in particular) to be able to visualize each estimation and see how changing the input values described above will affect the detection of fixations.

When the user acknowledges that the fixation detection is accurate enough, the estimation can be accepted and the manual inspection of the algorithms' outcomes will start. At this point, the user can review all fixations and manipulate them in order to meet a chosen fixation detection criteria and reduce the noise in the data.

The researcher can decide whether the input parameters for the event detection algorithms should remain the same for the entire data set or if they should change as a function of data quality. In case different parameters are used for different participants at different levels of data quality, it can be argued that as the results for these participants were calculated by using different criteria they should not be grouped together. We know, however, that the selection of certain input parameters will affect low and high quality data differently (see section 3.5.3) Therefore, even when the parameters remain the same for the entire data set they may affect the results from participants presenting high and low data quality differently. For this reason it can also be argued that adapting the input parameters is a step to reduce the levels of noise for each participant's results. Nonetheless, the algorithms alone, even after adapting the input parameters on a participant basis, are always subject to errors, particularly when processing low quality data. In order to fix some of these errors and achieve the cleanest results possible I also propose the manual adjustment of fixations.

3.6 Manual adjustments of fixations

Once the fixations have been automatically calculated, the user can examine and manipulate them in order to fix the algorithms' outcome. Even the most accurate algorithms may generate a number of artifactual fixations that can corrupt the validity of the experimental results. This is because most of the times the data is assumed to have high spatial precision or at least to present similar levels of noise across the whole duration of the experiment, and this is not always the case. In fact, when working with populations such as infants it will rarely be the case.

Fixations can be created, deleted or merged by simply clicking and dragging the mouse on the main screen. For instance, in order to create a fixation the user just needs to click the point on the screen where the fixation starts and drag the cursor until the point where the fixation ends. The tags *From* and *To* located at the upper right of the screen indicate the exact onset and

offset of the current fixation (see Figure 3-3). Once the fixation is created it will appear on the fixations list at the upper left of the screen. If a fixation A starts on the current fragment and ends on the next, we need to (1) create a first fixation whose onset fits with fixation A's onset and drag the cursor a bit further than the end of the fixations box, (2) create a second fixation on the next fragment whose offset fits with fixation A's offset, and (3) merge both fixations. Additionally, fixations can be coded as smooth pursuit once they are created.

In general high quality data sets will not need as much manual adjustment, whereas low quality sets will require significantly more. I refer to the process of first parse fixations applying detection algorithms and then fix the outcome with the hand-coding tool as the two-step approach. The only difference between this method and hand-coding is that for the two-step approach the detection algorithms are first executed and their output is used as a starting point for doing the manual-coding. Thus, the results from the two-step approach and from a purely hand-coding approach should be approximately the same, while the coding-time will be considerably reduced with the proposed method (for more details go to the section *3.9.2*). To demonstrate the time difference I coded a randomly selected subject both manually (using GraFIX hand-coding tools) and applying the two-step approach. The total length of the experiment was 18.5 minutes. I invested considerably more time hand-coding the data (51 minutes) compared to applying the two-step approach (35 minutes). Still, these values are both considerably lower than the time that was required by previous hand-coding approaches (e.g., coding the same videos frame by frame could easily take several hours).

Evidently the amount of time that the researcher needs to expend coding depends on his/her expertise and on the characteristics of the data (e.g., data quality). Further, accurate detections will require less coding than inaccurate ones, and hence the coding time in these cases will be shorter.

3.7 Visualizations

Most of the time it is relatively easy to identify fixations by looking at the 2D representation of the x and y coordinates; however, when the coder is not entirely sure about coding a particular fixation it is very helpful to visualize the data in other formats.

GraFIX allows the 2D visualization in real-time of the raw and smoothed data together with the IDs of the fixations that are being coded. Additionally, it is possible to include the stimuli in the

background for all the different tasks of the experiment. This permits a further evaluation of the fixations and facilitates the coding process, especially for novice coders.

3.8 Pupil dilation

GraFIX will also process the pupil dilation data in case it is provided. Once the pupil dilation data is included in the raw input file, it is automatically displayed on the main window. Furthermore, the visualization dialogs include the option to play the eye-tracking data together with pupil dilation. Each fixation that is created or modified by GraFIX includes the pupil dilation means.

3.9 Software Validation

GraFIX has been evaluated from four different perspectives. First, the agreement between two different raters was assessed using the intra-class correlation coefficient (ICC; Hallgren, 2012) in two groups of infants featuring low and high quality data. Secondly, hand-coding results were compared to the two-step approach (automatic detection + hand-coding), demonstrating that both techniques were generating exactly the same results. Thirdly, the outcome from GraFIX automatic algorithms was compared to the two-step approach. Finally, I compared hand-coding results with GraFIX automatic algorithms and previous automatic detection algorithms (the velocity-based algorithms from Wass, et. al, 2013; the adaptive velocity-based algorithms from Nyström and Holmqvist, 2010; and the I-VT filter velocity based algorithm as implemented in Tobii-studio 3.0.0).

Additionally, GraFIX has been successfully used to code data from various monocular and binocular eye-tracking systems such as Tobii, Eye Link or SMI systems, at different sampling rates. Even though using different eye-trackers does not affect the performance, high sampling rates may slow down the execution of the algorithms.

3.9.1 Validation 1: Inter-coder reliability

Manual coding always involves an evaluation of the degree of agreement between different raters. Data from a group of three infants with low spatial precision (RMS > 0.30° per infant; 20-25 minutes of data each) and another group of three infants with high spatial precision data (RMS <0.13° per infant; 20-25 minutes of data each) was recorded using a Tobii TX300 eye-

tracker at a sampling rate of 120Hz and Matlab (with Psychophysics Toolbox[1] Version 2 and T2T[2]). Note that even though the spatial precision for the second group was relatively high, it was still data coming from infants and thus there was a high degree of head motion and frequent missing data points.

An external coder with no eye-tracking experience and naive to expected outcomes was trained to code fixations from both groups. The second coder was the author for this thesis. The coders had to (1) run the automatic detection algorithms using the parameters from Table 3-1, and then (2) manipulate the resulting outcome in order to remove artifactual fixations or add those undetected by following the predefined guidelines. The input values for automatic detection were chosen after executing the algorithms with a wide range of values and evaluating the outcomes. The values from Table 3-1 may not necessarily be optimal in data sets with other characteristics, and may change for different participants, experiments and/or groups. I decided to have two sets of parameters for the two different groups in order to facilitate the process for the novice coder and to define some standards for the execution of the automatic detection algorithms.

In order to keep the same standards across participants and coders it is essential to define strict guidelines about how to code the data. A fixation was coded when both the x and y coordinates were stable at one point or in other words, when the 2D representation of both x and y coordinates were displaying horizontal lines. If the detection of one eye was imprecise the data from the other eye was used. If the coder was not entirely sure about coding a particular fixation he/she was advised to leave it out. Saccades that were too short to be detected by the algorithms were also coded. Fixations that were cut by blinks and smooth pursuit eye-movements (diagonal movement of the x/y trace) were deleted. These guidelines may change depending on the experimental design. For instance, if the researcher is particularly interested in smooth pursuit eye-movements, those would not be deleted.

The inter-rater reliability between the means and the number of detected fixations was evaluated using the ICC (Hallgren, 2012). A strong agreement between the mean fixation durations was found for both the low quality data group (with an ICC of .967, $p = 0.016$) and the high quality data group (with an ICC of .887, $p = 0.038$). Additionally, I also found strong

---

[1] See Psychophysics Toolbox documentation: http://psychtoolbox.org/
[2] See T2T documentation: http://psy.ck.sissa.it/t2t/

agreements in the number of fixations detected for low (with an ICC of .938, *p* = 0.037) and high

quality data (with an ICC of .971, *p* = 0.009). Interestingly, the agreement in the low quality

group is slightly higher than in the high quality group. This may be because fixations that were

not clear enough were not coded, and this can appear to be slightly more subjective in high

quality data sets, where the data quality is a bit more variable across the time course of the

experiment (due to head motion and/or data loss). Possibly, one coder was a bit more strict

than the other removing a higher number of automatically detected fixations in the parts were

the data was not optimal.

This demonstrates that the manual coding can be highly reliable, even in low quality data sets.

Table 3-1 Input parameters for high and low spatial precision data for the inter-coder reliability data.

	High spatial precision	**Low spatial precision**
Interpolation latency (ms)	60	60
Velocity threshold (°/sec)	9	20
Maximum interpolation displacement (°)	0.25	0.25
Degree per pixel (°/pix)	0.0177	0.0177
Maximum distance for merging adjacent fixations (°)	0.24	0.35
Maximum time for merging adjacent fixations (ms)	50	50
Maximum RMS per fixation (°)	0.24	0.21
Minimum fixation duration (ms)	99	120

3.9.2 Validation 2: Comparing hand-coding with the two-step approach

In this section I demonstrate how the results generated by hand-coding are the same as those

obtained applying the two-step approach, where the data is pre-processed using event-

detection algorithms before it is hand-coded.

As mentioned in previous sections, the main purpose of pre-processing the data before it is hand-coded is to speed up the process of detecting fixations: the more fixations the algorithms are able to detect accurately, the less time the coder will expend manually adjusting fixations afterwards. It can be argued, however, than having the results from the event detection algorithms as a basis to hand-code fixations can influence the coder's decisions for accepting or deleting fixations. To demonstrate that this is not the case I compared results from hand-coding (using GraFIX coding tool, but without pre-processing the data beforehand) with results from the two-step approach.

I used exactly the same data as in the previous section where two groups of infants featuring low and high quality data were analysed. One of the coders re-coded all the data using a purely hand-coding approach in order to compare it with results from the previous section coded with the two-step approach.

The inter-rater reliability between the means and the number of detected fixations was evaluated using the ICC (Hallgren, 2012). All infants were included in the same analysis regardless their data quality. A strong agreement between mean fixation durations was found (with an ICC of .993, $p < 0.001$). Additionally, I also found a strong agreement in the number of detected fixations (with an ICC of .994, $p < 0.001$).

This analysis demonstrates that the results from a purely hand-coding approach and the two-step approach are the same, thus both can be considered as close to a "ground-truth" identification of fixations as is possible.

3.9.3 Validation 3: Comparing the automatic detection with the two-step approach

In this section I compare GraFIX algorithms for the automatic detection of fixations with the two-step approach where the algorithm's outcome was also hand coded.

I used exactly the same data as in the previous sections. In particular, I took exactly the same fixations calculated by one of the raters, which were coded using the two-step approach, and use it to compare it with the outcome from the algorithms alone. As it was demonstrated in the previous section, in terms of results the only difference between hand-coding and the two-step approach is that the second one is faster. In both cases the data is manipulated to reach the same criteria, thus the two-step approach could be considered a method to hand-code the data. The input values for the automatic detection algorithms were the same specified in Table 3-1.

For the high quality data group I found a strong agreement between the automatic and the hand coding both for mean fixation durations (with an ICC of .973, *p* = 0.019) and number of fixations (with an ICC of .966, p = 0.008). On the other hand, no significant agreements were found for the low quality data group for the means (with an ICC of 14.969, *p* = 0.849), even though there was an agreement in the number of detected fixations (with an ICC of .898, *p* = 0.073). This can also be seen in the means and standard deviations from Table 3-2: the values resulted from automatic algorithms and hand-coding in high precision data look quite similar, whereas it is not the case for low precision data.

Figure 3-7 shows how the algorithms are able to accurately detect fixations in high spatial precision data (Figure 3-7, left), although even then few manual adjustments are advisable. On the contrary, low spatial precision data (Figure 3-7, right) needs major adjustments even though these algorithms alone can still capture the trend in the fixations duration distribution.

Table 3-2 Automatic vs. Hand-coding: Fixation durations means and standard deviations in low and high spatial precision data.

	High spatial precision	Low spatial precision
Automatic algorithms FDs	625.1 ± 847.8, (N = 2410)	552.2 ± 536.3, (N = 863)
Two-step approach FDs	627.9 ± 866.2, (N = 2268)	489.9 ± 445.9, (N = 858)
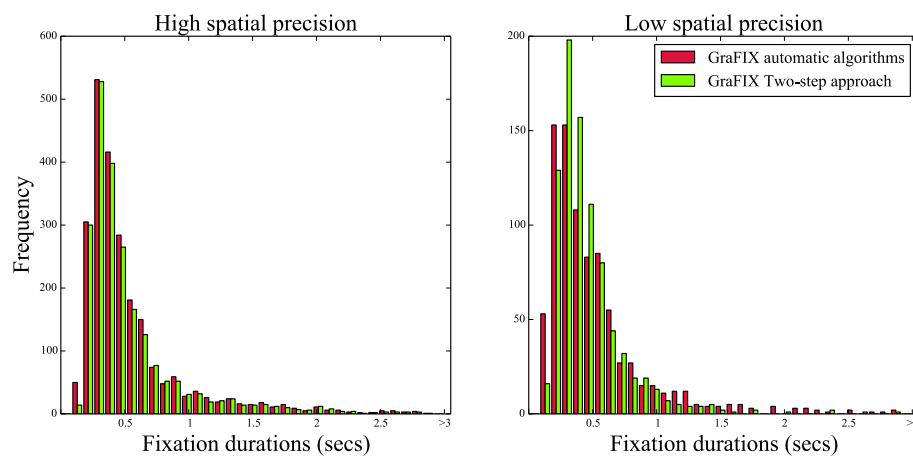
Figure 3-7 GraFIX automatic algorithms vs. Two-step approach (automatic detection + hand-coding) in high (left) and low (right) quality data.

3.9.4 Validation 4: Comparing GraFIX to previous approaches

In this section I compare the detection results from GraFIX (both the automatic algorithms and the hand-coding) with previous algorithms. In particular, I tested the fixation-parsing algorithms for low quality data described in Wass et al. (2013), the adaptive velocity-based algorithms from Nyström and Holmqvist (2010), and the I-VT filter (as implemented in Tobii Studio 3.0.0). The last two algorithms are not designed to deal with particularly low quality data, such as data recorded from infants. In fact, even though Nyström and Holmqvist (2010) adapt the velocity threshold according to the level of noise, they still maintain that the algorithm is suitable only for data collected from viewers with relatively stable heads while watching static stimuli[3]. This is obviously not the case for most data coming from infants and other special populations, which is likely to be much noisier than any of the recordings previously tested with these algorithms. However, given that these algorithms are considered a well-established method for event detection, I decided to include them in our comparison.

I selected three infants that presented high precision data (RMS <0.13º per infant; 5-6 minutes of data each) and another three infants that presented low precision data (RMS > 0.25º per infant; 5-6 minutes of data each) from an experiment that was recorded using a Tobii TX300 eye-tracker and Tobii Studio 3.0.0 at a sampling rate of 120Hz. Once again it is important to bear in mind that this is data from infants and thus it still presents a high degree of movement, variability in the levels of noise across the experiment and frequent missing data points, even in the high precision group.

When possible, I used the same input parameters for the four algorithms (see Table 3-3). Nevertheless I still kept two sets of parameters for GraFIX automatic algorithms (for high precision and low precision data), as adapting the input values according to the data quality is still one of the main advantages of the present approach.

The parameters that I used for Wass et al. (2013) algorithms and the I-VT Filter were –if applicable- the same as for GraFIX algorithms in low quality data. This is because when there are subjects with various levels of noise it is a more common practice to use thresholds that rather fit the subjects with higher noise. The algorithms from Nyström and Holmqvist (2010) include an adaptive velocity threshold that is re-calculated every 10 seconds (I divided the data

---

[3] See the README attached to the code the authors provide.

for each participant in 10 seconds chunks). The input parameters that are not reported in Table 3-3, such as the blink acceleration threshold or the post-hoc validation inputs, were set to the values that were recommended in the original articles from Nyström and Holmqvist (2010) and Wass et al. (2013) respectively.

Table 3-4 shows the means and the standard deviations obtained from the different algorithms and hand-coding and Figure 3-8 displays the graphs with all the fixation duration distributions from the four algorithms paired with the hand-coding distribution, which was coded by using the two-step approach. I assume that the algorithm that gets closer to the hand-coding distribution will be the one able to detect fixations more accurately. The differences between algorithms in both high and low precision groups are striking. Results for the high spatial precision group revealed differences in the means and also in the number of detected fixations. The I-VT filter in particular, presented an especially high number of detected fixations (N = 1199) compared to hand-coding (N = 973) that can be the result of mistakenly flagging very small fixations in segments that were slightly noisier (see Figure 3-8; First row, fourth column). Because of this reason the mean duration (M = 554.0 ms) are lower than the means for the other algorithms or for hand-coding. On the other hand, GraFIX and Wass et al. (2013) algorithms present means that are a bit above the hand-coding mean (M = 674.5 ms). An explanation for this can be related to the selection of the velocity thresholds and the sampling rate. As it has been previously mentioned, saccades with very small amplitudes may go undetected by velocity-based algorithms, especially when the data is recorded at low sampling rates (<200). As a consequence, the fixations before and after these saccades will be merged together in a longer fixation. Obviously, at higher velocity thresholds it is more likely that fixations will be merged together. Since the velocity threshold for Wass et al. (2013) was higher (20°/sec) than for GraFIX algorithms (9°/sec) it is also not surprising that the mean for the first algorithms was still higher than for the proposed algorithms. The algorithms from Nyström and Holmqvist (2010) avoided this problem by adapting the velocity threshold according to the level of noise in the data. Even though they still capture the trend in the distribution, the data quality for the samples that were analysed (even for the high precision group) is probably too low for obtaining more accurate results from these algorithms. Eyeballing the graphs from Figure 3-8 (first row) it is possible to see how the fixation duration distribution produced by GraFIX algorithms had the closest resemblance to the hand-coding distribution in high precision data.

Results from the low spatial precision group revealed even higher differences between the four algorithms. As it can be seen in Table 3-4 and in the Figure 3-8 (Second row, fourth column), the problem that the I-VT filter presented in the high precision data was even more obvious here. Likewise, Nyström and Holmqvist (2010) algorithms did not manage to deal with such a high degree of noise (see Figure 3-8; Second row, third column). Looking at the graphs and the means it seems that due to the low precision in the data the velocity threshold that was calculated may have been too high. It is also interesting to see that even though the length of the recordings was approximately the same for the low and high quality groups, the number of detected fixations in low precision data was almost half the number of fixations detected in high precision data for GraFIX algorithms, Wass et al. (2013) algorithms, and hand-coding. Wass et al. (2013) algorithms exclude a high number of fixations with their post-hoc validation criteria, and this is probably why their algorithms still detect many less fixations than the algorithms that I propose. Once again GraFIX algorithms seem to resemble the hand-coding fixation duration distribution more accurately, even though they would still need manual adjustments (the two-step approach) to be perfect.

Overall, even though all the algorithms are far from perfect, results from GraFIX algorithms were the ones that more closely matched the hand-coding results. I believe that this is not only because the particularities of these algorithms but also because I am adapting the input parameters to different levels of noise. I would not recommend, however, the exclusive use of automatic detection algorithms unless the data quality is very high. Evidently, when the algorithms' outcome is accurate, the time that needs to be invested in correcting artifactual fixations will be considerably lower.

In sum, GraFIX seems to be an effective alternative to previous methods that will improve the quality of our results and the time invested coding eye-tracking data.
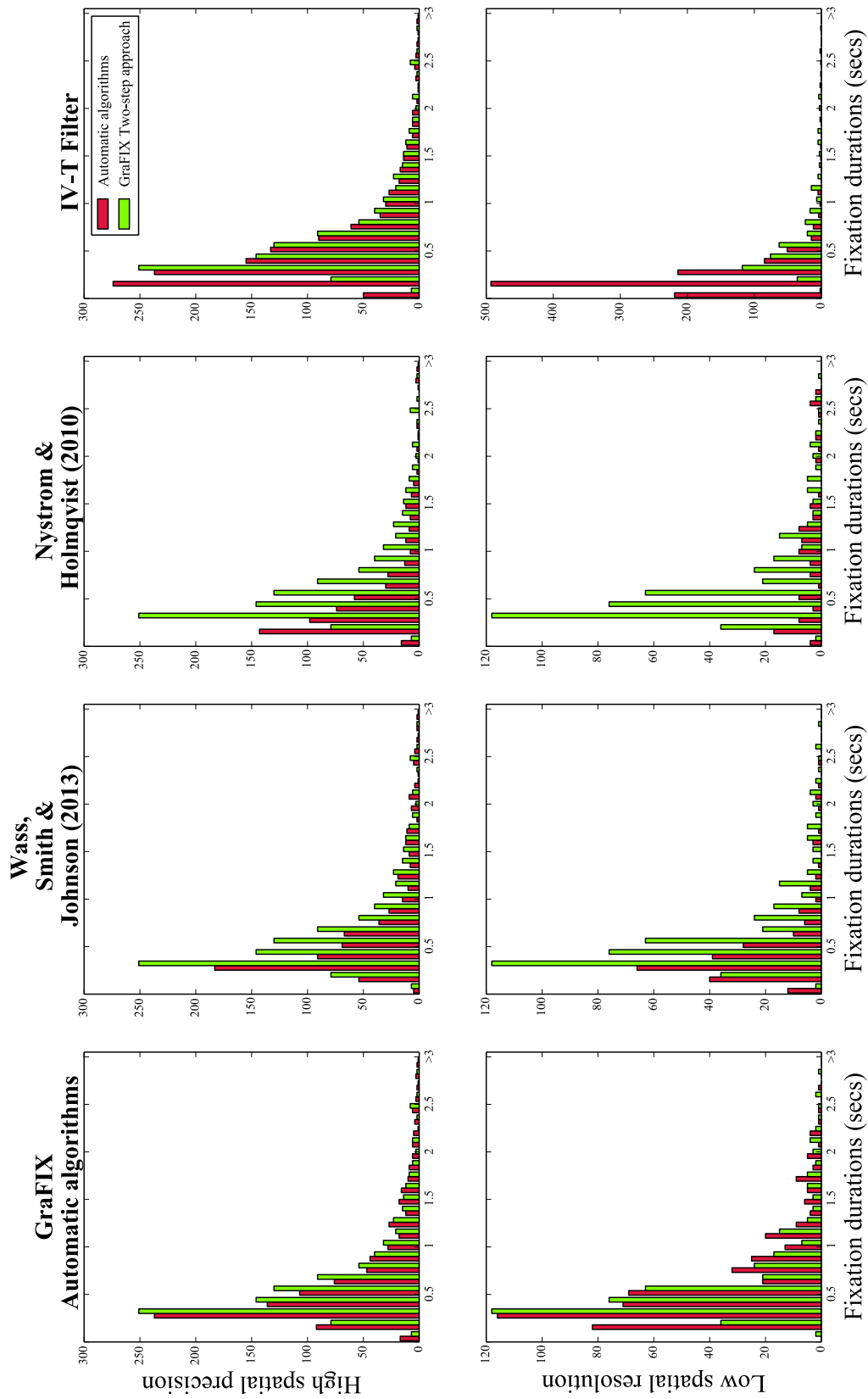
Figure 3-8 These graphs display the hand-coding (green; GraFIX two-step approach) fixations distribution paired with the distributions for the four different algorithms (red). I assumed that the algorithm that is closer to hand-coding is the one able to detect fixations more accurately.

Table 3-3 Input parameters for the automatic detection algorithms

	GraFIX (high quality)	GraFIX (low quality)	Wass, Smith & Johnson (2013)	Nyström & Holmqvist (2010)	I-VT Filter
Interpolation latency (ms)	60	60	60	n.a.	60
Velocity threshold (°/sec)	9	20	20	Adaptive	20
Maximum interpolation displacement (°)	0.25	0.25	n.a.	n.a.	n.a.
Degree per pixel (°/pix)	0.0177	0.0177	0.0177	n.a.	n.a.
Maximum distance for merging adjacent fixations (°)	0.24	0.35	n.a.	n.a.	0.35
Maximum time for merging adjacent fixations (ms)	50	50	n.a.	n.a.	50
Maximum RMS per fixation (°)	0.24	0.35	n.a.	n.a.	n.a.
Minimum fixation duration (ms)	99	120	100	100	100

Table 3-4 Comparing detection algorithms with hand-coding: Fixation durations means and standard deviations in low and high spatial precision data

	High spatial precision	Low spatial precision
Hand coding	674.5 ± 621.9, (N = 973)	657.3 ± 642.0, (N = 424)
GraFIX Automatic algorithms	719.9 ± 696.4, (N = 954)	640.4 ± 589.3, (N = 505)
Wass, Smith & Johnson (2013)	779.3 ± 826.5, (N = 676)	491.2 ± 490.2, (N = 229)
Nyström & Holmqvist (2010)	571.2 ± 588.7, (N = 540)	1337.9 ± 1435.1, (N =103)
I-VT Filter	554.0 ± 527.4, (N = 1199)	240.7 ± 169.0, (N = 1102)

3.10 Discussion

In this chapter, I described a new method and software to parse fixations in low and high quality data. This method was used in all the studies described within this thesis. Previous fixation detection methods are based on either purely automatic approaches or manual coding of the eye-tracking data. The high variability in data quality across participants and even during the experiment can seriously affect the automatic detection algorithms and as a consequence their results can be remarkably unreliable. On the other hand, current hand-coding methods can be extremely time consuming and imprecise. The present method implements a two-step approach for detecting fixations, where the data is first parsed by using a new adaptive velocity-based algorithm specially designed to deal with low quality data, and second the algorithm's outcome is manipulated with a view to fix the errors that the automatic process may have generated.

GraFIX fixation detection algorithms go through a number of steps in order to parse fixations accurately. First, the raw data is smoothed by using a bilateral filtering algorithm (based on previous implementations from Durand & Dorsey, 2002; Frank et al., 2009). This algorithm averages the data for both eyes and eliminates the jitter while preserving saccades. Secondly, missing data points are interpolated in order to avoid the detection of artifactual fixations in the subsequent steps. I showed how these missing points generate peaks in velocity that can be mistakenly flagged as saccades. Thirdly, a velocity-based algorithm gives us an initial parsing of the data. However, the results from this algorithm may still include artifactual fixations. Finally, GraFIX executes three post-hoc validation algorithms aiming to fix and/or remove the artifactual fixations detected in the previous step. In particular, the post-hoc algorithms (1) merge adjacent fixations that are close in space and time, (2) remove all the fixations whose RMS is over a given threshold, and (3) delete all the fixations with shorter duration than the minimum fixation (in this order).

GraFIX detection algorithms aim to obtain the most accurate fixation detection for each participant and thus reduce the amount of time the researcher has to spend correcting fixations in the subsequent step. Further, the hand-coding graphical tool alone - where the user simply needs to click on the screen to manipulate fixations- is already much faster than previous hand-coding approaches (e.g., coding fixations analysing videos frame by frame).

I evaluated GraFIX from four different perspectives: (1) I used the ICC (Hallgren, 2012) in order to evaluate the agreement between two different researchers when coding two groups of infants featuring low and high quality data, (2) hand-coding was compared to the two-step approach demonstrating that both methods generate near identical results, (3) GraFIX automatic algorithms were compared to the two-step approach, and (4) GraFIX automatic algorithms were compared to previous automatic detection methods (the velocity-based algorithms from Wass et al., 2013, the adaptive velocity-based algorithms from Nyström and Holmqvist, 2010, and the I-VT filter). Additionally, GraFIX was tested with data from different eye-tracking systems. Results from these analysis revealed that GraFIX automatic algorithms was the method that more closely matched hand-coding results, and that these algorithms alone can be a more reliable technique than other methods, overcoming some of the previous issues detecting fixation in low and high quality data. However, as demonstrated in this chapter, I strongly believe that given the nature of the infant data presented in this thesis any automatic algorithm should be used in combination with a later hand-coding approach.

Many of the current detection algorithms, especially those commercially available, obscure the quality of the data and the fixation detection process. Consequently evaluating the reliability of results or checking how different parameters affect the fixation detection may become an arduous task. GraFIX allows the adaptation of the input parameters for the automatic algorithms (e.g., velocity threshold, interpolation latency) to fit the data from different participants that present different levels of data quality. As it was previously explained, choosing the right velocity threshold highly depends on the quality of the data, and the experimental design. Ideally the user will select low thresholds in order to be able to detect small fixational eye-movements. However, the lower the data quality is, the higher the velocity threshold needs to be. If all the subjects are at very different levels of data quality, it can be worthwhile to adapt this value on a participant basis. Likewise, the rest of the parameters (such as the post-hoc validation parameters) should be adapted according to the subject's data quality in the interest of obtaining the most accurate results. Additionally, the execution of the algorithms is fast and it displays all the information that is needed to precisely evaluate what the algorithms calculated (interpolation and post-hoc validation flags for each point, visualization of the results paired with the velocity and the raw and smoothed data.). This can enormously facilitate the correct selection of the input parameters for the detection algorithm.

Traditionally, all participants are grouped together and a single velocity threshold that is usually selected to fit low quality sets is chosen to parse all the data. I have shown how applying high velocity thresholds can be the reason why the algorithms detect artifactual long fixations. I believe that adapting the input parameters on a participant basis can avoid the detection of artifactual fixations and will lead to more accurate and reliable experimental results. However, it is still up to the researcher to decide whether they prefer to use the same parameters for all the participants, adapt the parameters on a participant basis, or rather have different input parameters for different groups of subjects featuring different levels of data quality (as I did in section 3.9). It is also a topic of debate for the field of eye-tracking as to how the parameters used to parse each subject's data should be reported in publications, factored in to statistical analysis or standardised within populations and across labs.

But, to which extent is it acceptable to group together participants with very different data quality? For instance, this can be an important issue in clinical group comparisons where one group may present considerably lower quality data than the other (e.g., ADHD children vs. control groups), when analysing different age groups (e.g., 3 month-olds vs. 14 month-olds), or even when comparing "long fixators" with "short fixators". In low quality data there is less probability of finding clean long fixations that can be reliably detected, also when the data is hand-coded. This can lead to correlations between data quality and fixation durations where low quality data sets are more likely to present shorter fixations on average. To at least acknowledge these limitations in current eye-tracking studies, it would be advantageous to consistently report data quality measures and detailed descriptions of the detection methods, together with a data quality correlational analysis. Nevertheless, this still does not solve the problem.

Another limitation is related to the way fixations are hand-coded. In the interest of improving the reproducibility of the experimental outcomes, it is essential to include precise guidelines to define how the fixations are being coded. Without these guidelines the inter-coder reliability loses its value.

In sum, the proposed method and software proves to be a more reliable and accurate technique to parse fixations in low and high quality data, and overcomes many of the issues that previous methods presented. More accurate outcomes and reporting data quality measures and

descriptions of the detection methods in scientific papers can considerably improve the viability of research results and hence facilitate the replication of previous studies. This can have a big impact not only in research from populations that are particularly difficult to test and that typically present higher degrees of noise (such as infants, people in the Autism spectrum or ADHD patients), but also in subjects that simply do not reach certain data quality standards. In fact, nowadays we are experiencing an increase in the number of new low-cost eye-tracking systems that inherently suffer from data quality issues even with compliant participants. Additionally, GraFIX could also be adapted to code data from head-mounted eye-tracking systems by including the head-position and the eye and scene images.

3.11 Conclusions

In the present chapter I have presented GraFIX, a new method and software to detect fixations in low and high quality eye-tracking data. This method consists of a two-step process in which eye-tracking data is initially parsed by using velocity-based algorithms whose input parameters are adapted by the user, and then manipulated using the graphical interface, allowing accurate and rapid adjustments of the algorithms' outcome. I assessed this method by performing an inter-coder reliability analysis in two groups of infants presenting low and high quality data, and compared it with previous methods. Results revealed that the present approach gives rise to more reliable and stable measures in low and high-quality data.

GraFIX is the method that will be used for fixation detection in the upcoming chapters of this thesis, where I will investigate the development of saccadic control and fixation durations during the first year of life.

Chapter 4: Fixation durations in 6-month-old infants and adults

4.1 Introduction

Eye-movements are the earliest fine-motor skill to develop in infancy, and their progression to maturity provides a window onto the development of cognitive functions over the first years of life. The mechanisms underlying saccadic control in adults have been widely investigated through studying saccadic responses in simple oculomotor aiming tasks such as the gap-overlap (Fischer & Boch, 1983; Fischer & Breitmeyer, 1987; Fischer & Ramsperger, 1984; M. H. Johnson et al., 1991; Saslow, 1967), the double-step (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954) or the anti-saccade paradigms (Evdokimidis et al., 2002; Walker et al., 1998). Furthermore, a number of studies have investigated saccadic control and fixation durations in dynamic visual-cognitive tasks such as scene perception (Nuthmann & Henderson, 2012; Nuthmann et al., 2010; Walshe & Nuthmann, 2013). Nevertheless, the practical and technical constraints that testing infants entails (see Chapters 2 and 3) have impeded equivalent research in infancy, leaving many questions about the development of saccadic control and fixation durations unanswered: what are the mechanisms underlying saccadic control and fixation durations in infancy? What are the factors responsible for the large variations in infant fixation duration distributions?

Previous research has highlighted two main factors influencing fixation durations in infants: visual and cognitive processing (e.g., Bronson, 1990, 1994; Harris et al., 1988), and the developmental state of the visuo-motor system ("sticky fixation"; e.g., M. H. Johnson et al., 1991). However, not knowing to which extent these factors affect fixation durations at different developmental stages poses a problem for the interpretation of results when investigating the development of cognitive functions over the first year of life. Furthermore, from previous research with adults we know that there are other mechanisms related to saccadic control that can influence fixation durations such as saccade cancellation (e.g., Nuthmann et al., 2010; Walshe & Nuthmann, 2013) or inhibition (Evdokimidis et al., 2002; McSorley, Haggard, & Walker, 2009; Walker & Benson, 2013; Walker et al., 1998).

This chapter aims to shed light on some of these questions by investigating the factors modulating fixation durations and saccadic control in 6-month-old infants and adults when viewing complex dynamic stimuli and performing simple oculomotor tasks. For this purpose I

report an empirical study in which 6-month-old infants and adults were presented with a set of customized naturalistic and non-naturalistic videos while eye-movements were recorded. All participants performed a gap-overlap task in order to measure their disengagement abilities, and a double-step task that aimed to investigate saccade cancellation in infants.

The data and results from this study will constitute the basis for exploring the mechanisms underlying saccadic control in the next chapter, where data will be modelled using the CRISP theoretical framework and computational model of fixation durations in adult scene-viewing (Nuthmann et al., 2010).

4.2 Fixation durations and saccade latencies in adults

The relationship between fixation durations and visual and cognitive processing is particularly well investigated in skilled adult reading. Corresponding research over the past decades has revealed how fixation durations are affected by both low-level non-linguistic variables such as word length as well as lexical, syntactic or discourse factors (Rayner, 1998). In recent years, the mechanisms that control the duration of fixations have also been investigated in the context of scene viewing. For instance, factors such as the viewing task (search vs. memorization; Henderson et al., 1999; Nuthmann et al., 2010; Võ & Henderson, 2009), the visual characteristics of the stimulus (e.g., luminance, image degradation; Henderson et al., 2013; Loftus, 1985), the semantics of the scene (e.g., Henderson et al., 1999; Loftus & Mackworth, 1978; Võ & Henderson, 2009), or familiarity (e.g., Althoff & Cohen, 1999) can affect gaze control and fixation durations. From all this evidence it can be concluded that at least in adults visual and cognitive processing is associated with, and to an extent modulates, fixation durations (For more detailed information about fixation durations in adults see Chapter 1).

With regard to terminology, fixation duration needs to be distinguished from saccade latency (see Chapter 1). Saccade latency describes the time needed to program an eye-movement (reaction times), whereas fixation durations are the time intervals between successive saccades. In dynamic visual-cognitive tasks like scene perception, longer saccade latencies are thus likely to be associated with longer fixation durations (Nuthmann et al., 2010). The mechanisms underlying saccadic control in adults have been greatly investigated studying saccadic responses in simple oculomotor aiming tasks, such as the double-step and the gap-overlap paradigms.

4.3 Associations between fixation durations and visual and cognitive processing in infancy

The mechanisms that modulate fixation durations in infants are much less understood. There is evidence that suggests how as early as 3- to 4-months infants' looking behaviour can be influenced by cognitive factors related to the visual input. For example, Spelke (1990) used familiarization techniques to show how 3- and 4-month-olds were able to understand and apply perceptual properties of objects such as cohesion, or spatiotemporal continuity (see also Leslie & Keeble, 1987). Even though these studies demonstrated the relationship between cognitive processing and infants' looking behaviour they did not investigate the micro-dynamics of visual and cognitive processing, such as fixation durations. This research is becoming increasingly important to investigate questions related to the assessment of attention and information processing in spontaneous unconstrained settings.

Few studies have investigated the relationship between infant fixation durations and cognitive and information processing. Harris and collegues (1988) tested infants and adults in a free-viewing task, revealing that fixation durations decreased as the stimulus size increased (see also Hainline et al., 1984). Bronson (1990) investigated the development of visual scanning (and fixation durations) in a longitudinal study with infants from 2 to 14 weeks and found more mature and accurate scanning strategies in older infants. Furthermore, he found that even the older groups showed a higher proportion of longer fixations when looking at flicker stimuli, which is a looking behaviour more common in younger infants (see also Bronson, 1994). More recent studies have investigated infant scanning abilities when presented with familiar and non-familiar complex dynamic stimuli. Hunnius and Geuze (2004) followed infants between the ages of 6 and 26 weeks and presented them with a video of their mother's face, and an abstract video. They found that infants only adapted their eye-movements according to the type of stimulus from 14 weeks on, showing longer mean fixation durations for the abstract unfamiliar condition. Results also showed that the median fixation duration did not stabilize before 18 weeks, which is slightly later than what has been reported for static stimuli (Bronson, 1990). Interestingly, Hunnius and Geuze (2004) opted for the use of more ecologically valid stimuli. The use of stimuli that are of limited ecological validity has been often addressed and criticized since the generalizability of the experimental results is uncertain (Hunnius & Geuze, 2004a; Neisser, 1976; Schmuckler, 2001).

Fixation durations in infancy have also been associated with later intellectual function in childhood. In a recent study Papageorgiou and colleagues (2014) demonstrated that individual differences in fixation durations in early infancy predict differences in measures of temperament, behaviour and attentional control in later childhood.

These findings suggest that fixation durations in infancy can also reflect the visual and cognitive processing of the visual input, even though it is still unclear whether these factors have the same influence in infants and adults. In addition, fixation durations can be an indication of different cognitive or neural processes emerging at different developmental stages (For a wider review on fixation durations in infants go to Chapter 1).

4.4 Disengagement and the neural mechanisms underlying eye-movement control in infancy

As described in Chapter 1, whilst the subcortical structures involved in saccadic generation are relatively developed at birth (e.g., superior colliculus), cortical pathways associated with the generation of more complex eye-movements (e.g., the frontal eye fields, FEF) remain underdeveloped until 3 to 4 months of age (Atkinson, 2000; Bronson, 1974; M. H. Johnson, 1990, 2011).

"Sticky fixation" is thought to occur due to problems with "disengagement", defined as the difficulty generating an eye-movement after a fixation (M. H. Johnson, 1990). According to this disengagement may not only be due to a relative immaturity of the areas in the brain involved in eye-movement generation (e.g., the cortical pathway through the frontal eye fields, FEF) but also the result of interrupting and shifting the current focal point (Kikuchi et al., 2011; see Chapter 1). The ability to disengage from a central target to shift the gaze to a peripheral one has been traditionally evaluated using the gap-overlap paradigm, described in Chapter 1 (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991). In infancy research, longer disengagement latencies have been typically associated with greater immaturity of the visual system (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), particularly during the first 6 months of life when the neurological structures involved are thought to develop rapidly and approach their adult form (Rothbart et al., 1994). Nevertheless, evidence from various neurophysiological and behavioural studies suggests that the neural mechanisms involved in

the execution of a saccade are still undergoing development or need to increase in efficiency by 6 months postnatal age, despite hundreds of thousands of practice saccades. Butcher and colleagues (2000) evaluated the ability to shift gaze from a central to a peripheral stimulus from a group of infants that were tested from 6 to 26 weeks. They reported that even though infants were generally able to reliably disengage from a central stimulus from 16 weeks on, their efficiency was still increasing at 26 weeks, which suggests that at this age the attentional and oculomotor processes specific to shifting gaze are still developing (see also Hood & Atkinson, 1993; Matsuzawa & Shimojo, 1997). In adults, it has been reported that the superior parietal lobe is involved in disengagement and saccade execution (Posner, 1988). Furthermore, a number of characteristic saccade-related potentials (SRPs) associated with the eye-movement related areas of the parietal cortex have been found to precede the onset of a saccade (for more details see Kurtzberg & Vaughan, 1982). Infant studies, however, have reported that at 6 months these SRP components are weak, less synchronized to the saccade execution, or completely absent (Csibra, Tucker, & Johnson, 1998). These results suggest that while by 6 months the frontal cortex (and hence the frontal eye fields, FEF) is already involved in eye-movement control, other eye-movement related areas in the parietal cortex might still be in a "calibration phase".

4.5 Individual differences: Long vs. short lookers

Disengagement has also been associated with aspects of visual processing such as looking times in a familiarization context (see Chapter 1), but not fixation durations (Frick et al., 1999; see also Blaga & Colombo, 2006). In these studies, the younger infants are the ones showing longer looking times and greater disengagement difficulties, which could explain the relationship between look duration and the developmental state of the visuo-motor system (e.g.,Colombo, 1995; Johnson, 1990).

Another hypothesis states that shorter looking times reflect more rapid encoding of visual information within and across ages (Colombo et al., 1987; Colombo, 1995), with a number of studies suggesting how short lookers are more efficient processors compared to long lookers (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997) or present different strategies of visual intake (e.g., Colombo, Freeseman, Coldren, & Frick, 1995; Colombo et al., 1991). A recent study by Domsch, Lohaus, and Thomas (2010) investigated the joint influences of information processing and disengagement in look duration

within a habituation paradigm at 6 months. Using a computational modelling approach, they concluded that both factors are involved in infants' looking behaviour in visual habituation.

Look duration has also been associated with later intellectual function and cognitive performance in childhood and adolescence (Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Kavšek, 2004). Interestingly, whilst the correlation between look duration and cognitive performance during the first year of life has been reported to be negative—with short lookers showing better later cognitive performance (Colombo et al., 1995; Jankowski & Rose, 1997; Sigman et al., 1986; A Slater, 1995)—from 1 year on the correlation is reversed and longer lookers are associated with better cognitive function later on (Choi & Vaswani, 2014; Lawson & Ruff, 2004). This shift could be due to a change in the processes underlying look duration: during the first year of life—and particularly during the first 6 months—look duration may be predominantly affected by the developmental state of the visuo-motor system as well as by the speed of visual and cognitive processing. At ages beyond one year, however, look duration may reflect the ability to voluntarily sustain or maintain attention and the appearance of executive functions (Colombo, 2006).

4.6 The two-step programming of saccades: The double-step paradigm

The double-step paradigm has been widely used in adult research to investigate saccade programming (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954), but so far it has only been used in infancy research to identify a shift from retinotopic to spatiotopic coordinate systems (e.g. Gilmore & Johnson, 1997; Kaufman, Gilmore, & Johnson, 2006; Senju et al., 2011; for more details see Chapter 1).

In the original double-step paradigm, adult participants are instructed to follow a target while it makes two quick and successive movements or steps that are separated by a varying temporal gap (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954; see Chapter 1). Findings from these studies constituted the first evidence about the parallel programming of saccades in which saccade programming occurs in two stages: an initial labile stage during which the saccade is subject to cancellation, and a subsequent non-labile stage in which the program cannot be cancelled. Becker and Jurgens (1979) showed that participants' performance is best predicted by the time D elapsing between the second target

step and the onset of the response. The delay D itself also depends on two variables: $D = R_1 - \tau$. While R_1 represents the participant's reaction time, τ factors the interstep time between the first and the second target's appearances, and is the only variable under the control of the experimenter. If τ is *short*,[4] participants tend to make only one eye-movement to the second, *final* location of the target. In such cases, the oculomotor system begins programming a saccade to the second target location, while the saccade program related to the first target location is still in its labile stage of development. In this situation the first program is cancelled and only the second program is executed, prolonging the duration of the saccade latency and hence the fixation. On the other hand, if τ is *long* (e.g., > 70-80 ms for adult subjects) participants tend to make two saccades. The first saccade is directed to the first shifted target location and the second to the second shifted location. This is because the second saccade program is initiated while the first program is already fully specified or in other words, is in its non-labile stage of development, when no longer can be altered. The general finding that saccades are programmed in two stages has been adopted in computational models of fixation behaviour in reading (e.g., Engbert, Longtin, & Kliegl, 2002; Reichle, Pollatsek, Fisher, & Rayner, 1998) and scene viewing (Nuthmann et al., 2010; Trukenbrod & Engbert, 2014).

A more recent study by Walshe and Nuthmann (2013) adopted the double-step paradigm to a scene-viewing context and showed that saccade cancellation processes generalize to scene viewing and that cancelling a saccade prolongs fixation durations. Notably, the duration of the non-labile stage, which can be estimated from such data, was found to be longer in the scene-viewing condition compared to a classic "static" double-step condition. In addition, there were considerable individual differences in the duration of the non-labile stage (Walshe & Nuthmann, 2013).

Nevertheless, even though there are studies suggesting that infants as young as 4 months are able to inhibit a saccade to a target location (Johnson, 1995; see Chapter 1), to date there is no empirical evidence demonstrating if infants program saccades in two-steps (labile and non-labile stages), and that hence, are able to cancel a saccade program.

---

[4] The second target appeared in close temporal proximity to the onset of the response saccade.

4.7 The current study

For the current study I recorded the eye-movements of groups of 6-month-olds and adults who were presented with two different complex dynamic viewing conditions (naturalistic and semi-naturalistic videos) and performed the gap-overlap (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991) and double-step (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954) paradigms. The free viewing tasks were used to analyse the micro-dynamics of visual and cognitive processing during spontaneous orienting by measuring fixation durations. The gap-overlap task was used to measure participants' disengagement abilities. Finally, the double-step paradigm was used to investigate saccade cancellation in infants.

In particular, the present study aims to

(a) explore age differences in fixation durations between 6-month-olds and adults. I predict differences in mean fixation durations across age groups (infants vs. adults), with infants showing longer mean fixation durations;

(b) investigate the influence of the viewing condition—and hence of visual and cognitive processing—in fixation duration distributions. I predict differences in mean fixation durations across viewing conditions (naturalistic vs. semi-naturalistic videos);

(c) examine the relationship between disengagement abilities—which can be indicative of the general development of the visuo-motor system—and fixation durations. I predict a positive relationship between fixation durations and disengagement latencies in both age groups, particularly in infants due to the underdeveloped state of their visuo-motor system;

(d) explore individual differences in fixation durations in infants and adults. I predict a positive relationship between mean fixation durations in the naturalistic and non-naturalistic conditions for both infants and adults;

(e) demonstrate that 6-month-old infants are able to cancel a saccade program. I predict longer reaction times in the double-step task when infants make a single saccade—assuming in these cases a previous saccade program is being cancelled—compared to when they make two consecutive saccades.

4.8 Methods

4.8.1 Participants

In total, 18 typically developing 6-month-olds and 18 adults (10 female) with no known visual impairments participated in this experiment. Seven infants had to be excluded from the analysis due to data quality issues (N = 5) or fussiness (N = 2). Thus the analysis included 11 infants (mean age = 171.45 days, range = 150 to 191 days, 6 girls). Most infants (N = 10) were of Caucasian middle socioeconomic status. The infants were recruited via magazine advertisements, social networking media and flyers. Families were given baby t-shirts or bags as gifts for their participation. Adults were reimbursed for their participation. The study protocol was approved by the Birkbeck, Psychological Sciences Ethics Committee.

4.8.2 Apparatus

Participants' gaze was monitored using a Tobii TX300 eye-tracker running in a sampling rate of 120 Hz. This particular eye-tracker model tolerates large head movements allowing the infants to move naturally in front of the stimuli presentation screen. The stimuli were presented on a 23'' wide screen TFT monitor attached to the eye tracker unit by using a MATLAB program (MATLAB version R2010a 32-bit) specially developed for the purposes of the experiment. This program made use of the T2T (Talk to Tobii) package and the Psychophysics Toolbox Version 3 (PTB-3) (Shukla et al., 2011). The sounds were played through stereo external speakers located at both sides of the screen. The participant was monitored and recorded through an external video camera located under the Tobii screen by using the ScreenFlow screen-casting software. The objects and figures for the gap-overlap and double-step paradigms were created using Adobe Photoshop CS6.

4.8.3 Stimuli

The spontaneous looking task included two different viewing conditions: the naturalistic condition whereby three people performed baby-friendly actions simultaneously or at different times; and the semi-naturalistic condition whereby the people were substituted with simple geometrical shapes (e.g., triangles, circles, rectangles) rotating simultaneously or at different times. The two types of stimuli were presented in colour and had the same size.

4.8.3.1 Naturalistic videos

Each naturalistic video lasted between 20 and 25 seconds and comprised three regions of interest (right, left and centre of the screen) in which three different volunteers performed a random baby-friendly action, e.g., waving a balloon or tossing a ball. The position of the right and the left elements corresponded to a visual angle of 18.26°. This design aims to uniformly distribute the semantic and motion information across the screen. Further, to have three large and well-defined areas of interest facilitates the areas of interest (AOI) analysis in low-quality data sets (see Figure 4-1).

Figure 4-1 A frame from a randomly selected naturalistic video.

The videos were recorded using five female volunteers wearing simple and colourful clothes. The set included 10 different videos where three of the volunteers were instructed to perform 10 different actions (e.g., playing with a balloon, dancing). The volunteers could be either in an active state (they move and perform the correspondent action), or in an inactive state (they do not perform any action or movement). During the first second of each video all volunteers were in the inactive state. Subsequently the three volunteers enter the active state and start performing the correspondent action for about 3 to 5 seconds. After this period only one

volunteer at a time is active while the other two are in the inactive state. The videos were always presented in the same order. For a precise description of each video go to Appendix B.

The naturalistic videos were recorded with a high-resolution camera and assembled and manipulated using Final Cut Pro X Version 10.0.5. Each volunteer was recorded individually in front of a green cloth (see Figure 4-2).

Figure 4-2 Recording naturalistic videos using a green cloth as a background.

The green background was subtracted by using the *Mask* and *Keyer* video filters. Further, the volunteers were positioned in their location (centre, right or left) and a naturalistic photo was added as a background. Each volunteer's inactive states were controlled by using the *hold* Final Cut function, which freezes a selected frame for a concrete period of time. The music accompanying the naturalistic videos was comprised of fragments of instrumental songs from a group of Basque folk music called Oskorri.

4.8.3.2 Semi-naturalistic videos

In the semi-naturalistic condition (see Figure 4-3) the volunteers were substituted with simple geometrical shapes (coloured triangles, rectangles, circles, squares, pentagons, hexagons or ellipses) that rotate clockwise simultaneously or at different times. During the presentation of each video (1) the objects remain still for the first second, (2) the three objects rotate together for 5 seconds, (3) the three objects rotate subsequently for 3 seconds each, and (4) the three objects rotate together for another 4 seconds. The background music was the same as for the naturalistic stimuli. In total 12 semi-naturalistic videos were presented. The two extra videos

were added in order to compensate a possible decrease of attention to this condition towards the end of the experiment and to obtain the largest number of fixations possible.

Figure 4-3 A frame from a randomly selected semi-naturalistic video.

4.8.4 Gaze-contingent paradigms

In a gaze-contingent paradigm the experimental events respond to the participant's online gaze permitting the study of eye-movements in a truly interactive manner (McConkie & Rayner, 1975; Wass et al., 2011; Wilms et al., 2010). The current experimental design includes two gaze-contingent tasks: the gap-overlap and the double-step paradigms. The contingencies were created by using the T2T (Talk to Tobii) functions (Shukla et al., 2011).

4.8.4.1 Implementation of the gap-overlap paradigm

As it was explained in Chapter 1, the gap-overlap paradigm investigates disengagement abilities and facilitation by measuring the time it takes for participants to move their eyes away from a central stimulus after a peripheral one is presented (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991). For the present implementation, the peripheral stimulus was always the same (a cloud), and the central stimulus as well as the background colour changed after every 12 trials (see

Figure 4-4). Each trial started with the presentation of a central stimulus. When the participant fixated the central stimulus, the peripheral target randomly appeared either on the right or left side of the screen (at 18° of visual angle).

Figure 4-4 The central and peripheral stimuli used for the gap-overlap paradigm.

Depending on the time when the central stimulus disappeared and the peripheral stimulus appeared the trials were flagged as overlap, baseline or gap. In the overlap condition the central stimulus stays visible during the presentation of the peripheral target until the end of each trial. In the baseline condition, the central stimulus disappears simultaneously after the peripheral target comes into view. And in the gap condition, the central stimulus disappears immediately after fixation, followed by a gap of 200 ms before presentation of the peripheral stimulus. When the participant looked at the peripheral stimulus a short animation was presented. If the participant did not look at the peripheral stimulus after 4 seconds, the next trial started (see Figure 4-5). The disengagement latencies were calculated by subtracting the baseline condition from the overlap condition. For the analysis presented in this thesis the gap trials were excluded.

A total number of 72 trials were presented. Forty per cent of these trials were gap trials, 30% overlap trials and 30% baseline trials. They were presented in groups of 12 trials (iteration) and alternated with other tasks of the study.
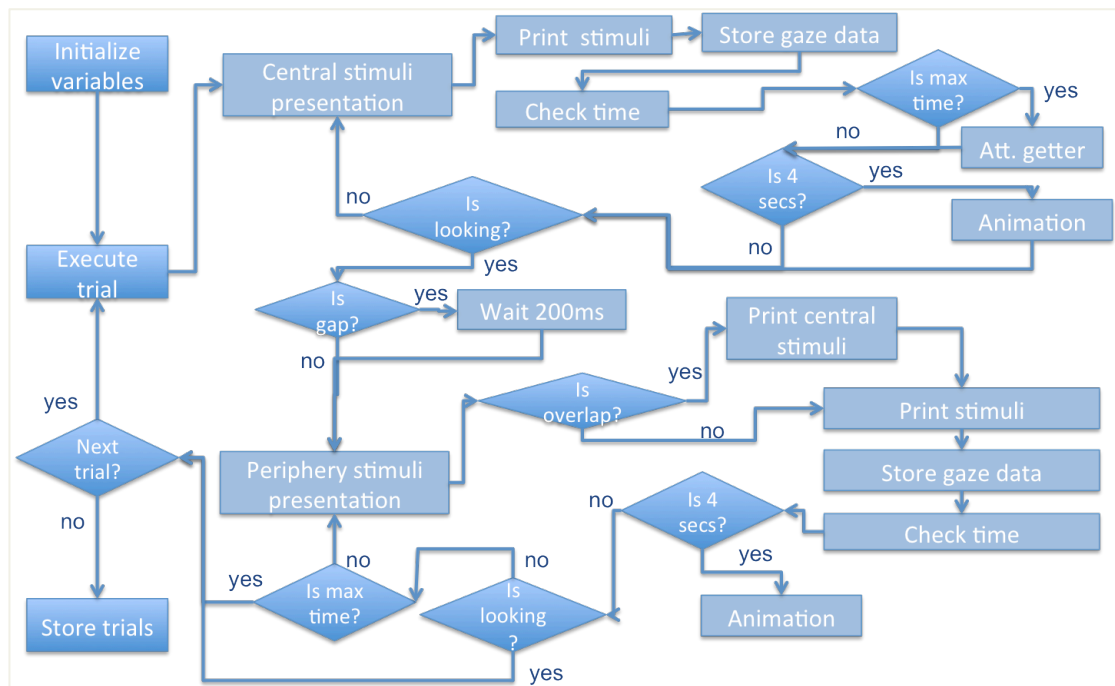
Figure 4-5 Flow diagram for the gap-overlap paradigm.

4.8.4.2 Implementation of the double-step paradigm

The current study included a version of the double-step paradigm (see Chapter 1 and the introduction of the present chapter for more details) that aimed to investigate saccade cancellation in infants by comparing the reaction times for the trials in which the infant executed a single saccade to the second target –for which the saccade program to the first target was assumed to be cancelled, prolonging reaction times-, or two consecutive saccades to the first and the second targets.

In previous adult studies participants were instructed to look at the second target if it was present. Further, the studies included around 576 trials per participant, which means that in total they obtained more than 10.000 saccadic responses (Becker & Jurgens, 1979). This allowed researchers to test different variations in the presentation of the stimuli and obtain distributions of reaction times with sufficient number of trials for different conditions. As a result of the difficulties of testing infant populations (e.g., they do not attend to experimental instructions; reduced attentive periods; for more information see Chapter 2), it is not possible to run such a high number of trials. For this reason the double-step task implemented for this chapter included a series of modifications to deal with the specificities of testing infants with this particular paradigm.

As explained earlier in this chapter (see also Chapter 1), in the double-step paradigm the presentation of a central stimulus is followed by the appearance of two sequential targets (see Figure 4-6). When the time during which the first target is present (τ) is long enough the participant is expected to execute a first saccade to the first target followed by a second one to the second target. On the other hand, when the presentation time of the first target (τ) is too short, it is hypothesized that the participant will cancel the current saccade program and execute a single saccade to the second target, prolonging the reaction time. Therefore, as for adults, in infants it is expected that saccade cancelation will prolong reaction times (e.g., Nuthmann et al., 2010; Walshe & Nuthmann, 2013).
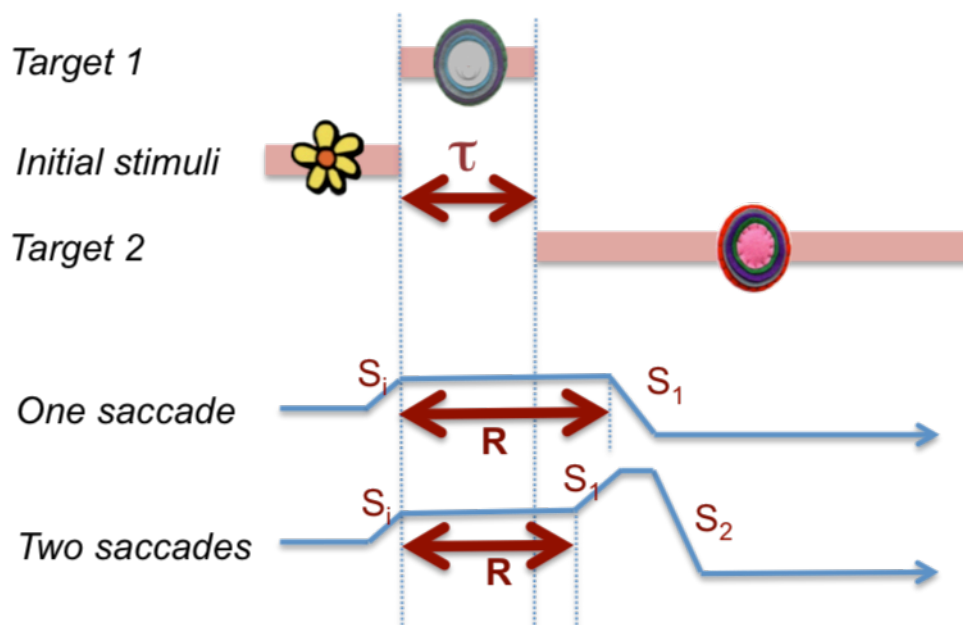
Figure 4-6 A trial for the double-step paradigm and its two possible outcomes. When the initial stimulus is presented, the participant makes an eye-movement to look at it (S_i). In this moment, the first target is displayed for a т period of time followed by the presentation of the second target. For the present implementation two possible outcomes are considered: (1) The participant executes a single saccade to the second target (S_1). In this case it is hypothesized that they may be cancelling an initial saccade program directed to the first target, prolonging the reaction time (R). (2) The participant executes an initial saccade to the first target (S_1) followed by a second saccade to the second target (S_2).

In the current version of the paradigm, the time during which the first target is present (τ) is calculated for each trial by using an adaptation of the Modified Binary Search (MOBS) thresholding algorithm (Anderson & Johnson, 2006; Tyrrell & Owens, 1988). This method, which is an adaptive and non-parametric technique used for identifying thresholds such as eyes resting states, allows to test the τ values that are most relevant for each participant. In

particular, the MOBS algorithm will try to identify the τ values where each participant will go from executing one to two saccades. For this purpose the MOBS algorithm calculates the next τ depending on the participant's actions during the previous trials (e.g., the participant makes one saccade vs. two saccades for a particular τ value). For instance, if the infant tends to execute two saccades for a certain τ value, the next trials will test shorter τ values.

The rules of the original MOBS algorithm do not allow a high number of errors. For instance, if a participant performs two saccades when the first target is presented for a period of 100 ms, the algorithm assumes that the threshold will be lower than this value and the program will rarely test the same τ value again. Certainly, these rules are too strict for testing infant populations, as a participant's response can also be the consequence of other factors unrelated to oculomotor control, such as the infant not being interested in the task and systematically making a single saccade to the second target in every trial. Thus, a tested τ value could not simply be diminished, and the original MOBS algorithm had to be adapted. The adaptation for the MOBS algorithm implemented for the current study is described below:

1. The minimum boundary is represented by a three elements "stack" called "low-stack". The initial minimum values are [0,0,0].
2. The maximum boundary is represented by a three elements "stack" called "high-stack". The initial minimum values are [200, 200, 200].
3. The current presentation time (τ) is the middle point between the top values at the "low-stack" and the "high-stack", unless 5 occurs.
4. After each trial, one "stack" is updated.
   4.1. If the participant makes two saccades, each element from the "high-stack" is decreased, having the τ just presented on the highest position and losing the bottom element.
   4.2. If the participant makes one saccade, each element from the "low-stack" is decreased, having the τ just presented on the highest position and loosing the bottom element.
5. If the participant presents the same response twice in a row:
   5.1. If the participant makes two consequent two saccades responses, the current presentation time (τ) is the top of the "low-stack". If the response to this is inconsistent

with the previous response to the top time of the "low-stack", the stack goes through a process of regression (6)

5.2. If the participant makes two consequent one saccade responses, the current presentation time (τ) is the top of the "high-stack". If the response to this is inconsistent with the previous response to the top time of the "high-stack", the stack goes through a process of regression (6)

6. Regression is the process whereby all the elements of one "stack" are moved up by one position. The bottom element of the "low-stack" would be reset to the minimum value, 0, and the bottom element of the "high- stack" would be resettled to the maximum value, 200. Regression occurs when:

6.1. "Low-stack" regression:

6.1.1. The conditions specified in the step 5.1 appear to be present.

6.1.2. The participant makes two consecutive two-saccade responses and the previous two presentation times (τ_{-1} and τ_{-2}) are equal to the "low-stack's" top value.

6.1.3. The participant makes two consecutive two-saccade responses and the "low-stack's" top value is the initial maximum (200).

6.1.4. The participant makes two consecutive two-saccade responses and the difference between the previous presentation time (τ_{-1}) and the current one (τ) is less than 10.

6.2. "High-stack" regression:

6.2.1. The conditions specified in the step 5.2 appear to be present.

6.2.2. The participant makes two consecutive one-saccade responses and the previous two presentation times (τ_{-1} and τ_{-2}) are equal to the "high-stack's" top value.

6.2.3. The participant makes two consecutive one-saccade responses and the "high-stack's" top value is the initial minimum (200).

6.2.4. The participant makes two consecutive one-saccade responses and the difference between the previous presentation time (τ_{-1}) and the current one (τ) is less than 10.

Aiming to encourage infants to look at the second target when it appeared, the second target was visually more attractive than the first one (see Figure 4-7). Additionally it changed after every 12 trials in order to keep infants interested for a longer period of time over a session of 72

trials. To prevent infants from predicting the targets' next locations 25% of the trials were one-step trials and the targets could randomly appear at the locations described in Table 4-1.

Figure 4-7 The stimuli used for the double-step paradigm.

Table 4-1 Locations were the double-step targets appeared.

	Central stimulus (Visual angle)	Target 1 (Visual angle)	Target 2 (Visual angle)
Location 1	11.31° top	14.57° right	14.57° left
		14.57° left	14.57° right
Location 2	11.31° right	14.57° top	14.57° bottom
		14.57° bottom	14.57° top
Location 3	11.31° left	14.57° top	14.57° bottom
		14.57° bottom	14.57° top
Location 4	11.31° bottom	14.57° right	14.57° left
		14.57° left	14.57° right

The double step paradigm started with the presentation of an initial central stimulus that was presented randomly at one of the locations described in Table 4-1. When a participant looked at the initial stimulus, the first and the second targets were displayed (see Figure 4-7). When participant looked at the second target a short animation was presented. After each trial the MOBS algorithm calculated the next time lapse between the first and the second targets (τ)

based on a participant's previous actions. Figure 4-8 displays the flow diagram showing the current implementation for the double-step paradigm.

As in the gap-overlap paradigm, 72 trials were presented in groups of 12 trials that were alternated with other tasks of the study. The initial stimulus, the second target and the background image (a noise mask) changed after every 12 trials.
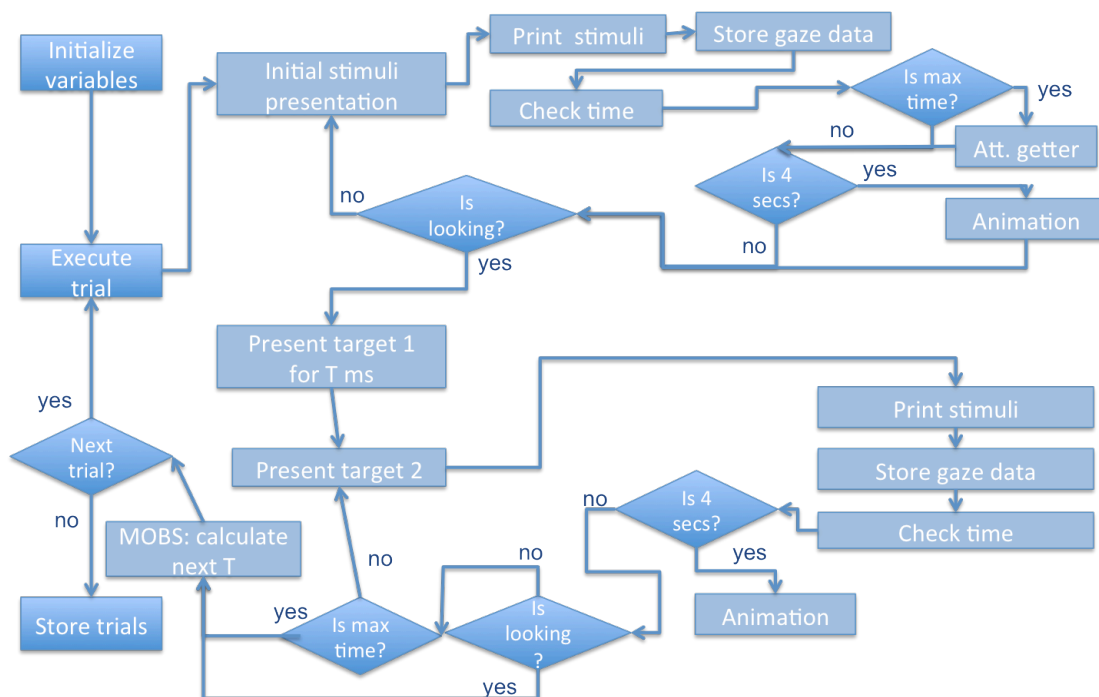
Figure 4-8 Flow diagram for the double-step paradigm.

4.8.5 Design and Procedure

Participants were welcomed in a lab waiting room where infants acclimatised to the experimenter and the lab. Next, they were tested individually in a darkened room while sitting on a baby-chair located 60 cm away from the monitor. Prior to starting the experiment the infants were calibrated with a 5 point calibration. Infants were shown a moving puppet accompanied by a sound on each point of the screen and on the centre until the target was fixated. This process was repeated until the infant looked at the five points accurately (for more details see Chapter 2). Adults were seated on a chair located 60 cm away from the monitor and also performed an infant 5 point calibration procedure.

The different tasks of the study were alternated in the following order: (1) twelve gap-overlap trials, (2) one naturalistic video presentation (20-25 seconds), (3) twelve double-step trials, and (4) one semi-naturalistic video presentation (20-25 seconds). This process was repeated 6 times. Next, 6 additional iterations of naturalistic and semi-naturalistic videos were presented. In total, participants viewed 10 naturalistic videos, 12 semi-naturalistic videos, and performed 72 gap-overlap trials, and 72 double-step trials. In case the infant became upset during the study the program was stopped (see Figure 4-9). The whole protocol took about 20 minutes to complete.
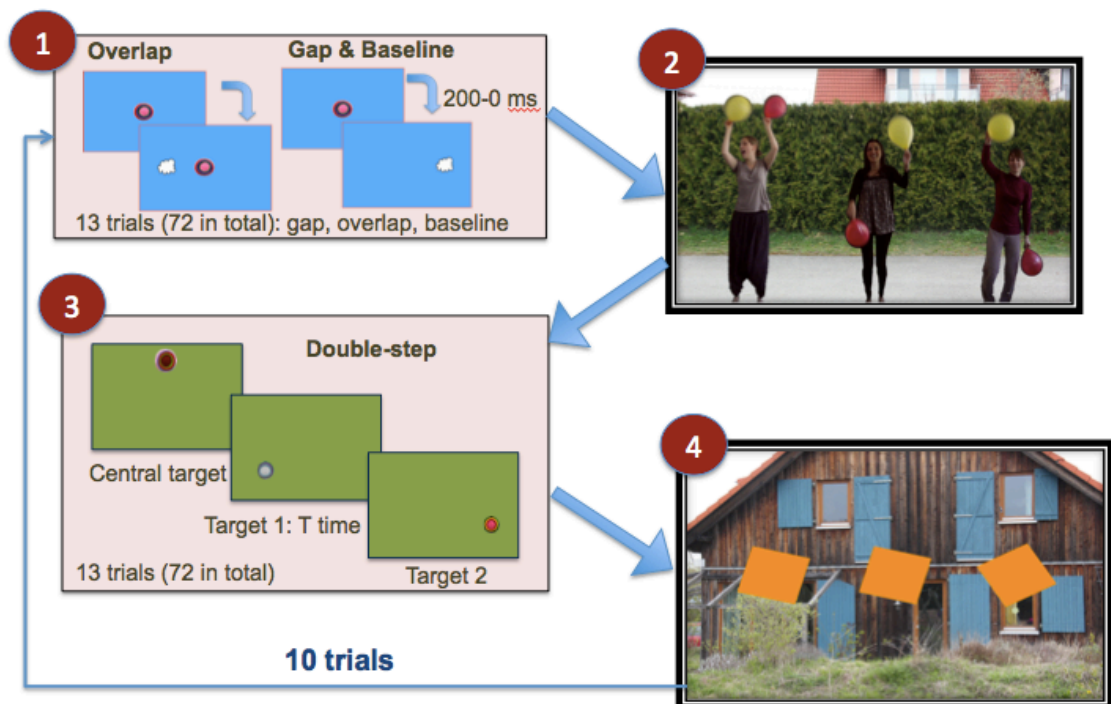
Figure 4-9 Experimental protocol.

4.9 Analysis

4.9.1 Fixation detection and coding

As described in Chapter 3, the quality of the raw eye-tracking data may vary depending on many different factors such as the eye-tracker model and manufacturer, or the eye physiology (Holmqvist et al., 2011, 2012; Saez de Urabain et al., 2014). Eye-tracking data coming from populations such as infants may contain considerably higher levels of noise compared to data from more compliant participants.

Fixations were detected using GraFIX, a semi-automatic approach for parsing low- and high-quality eye-tracking data. The input is initially parsed by using velocity-based algorithms whose input parameters are adapted by the user, and then manipulated using the graphical interface, allowing accurate and rapid adjustments of the algorithms' outcome (Saez de Urabain et al., 2014).

The procedure to code fixations was the same described in Chapter 3: the input parameters for the automatic detection algorithms were adapted according to the data quality for each participant and visit, having two different sets of parameters for participants featuring lower (>0.25° root mean square of inter-sample distances, RMS) or higher (<0.25° RMS) data quality (see Table 3-3). These parameters, chosen after executing the algorithms with a wide range of values and evaluating the outcomes, may not necessarily be optimal in data sets with other characteristics. Having only two sets of parameters rather than multiple options can facilitate the coding process (specially for novice users; (Saez de Urabain et al., 2014). In order to keep the same standards across participants and coders strict coding guidelines were defined. A fixation was coded when both the x and y coordinates were stable at one point, i.e. when the plots of both x and y coordinates against time were displaying horizontal lines. If the detection of one eye was imprecise the data from the other eye was used. If the coder was not entirely sure about coding a particular fixation they were advised to leave it out. Saccades that were too short to be detected by the algorithms were also coded. Fixations that were cut by blinks and smooth pursuit eye-movements (diagonal movement of the X/Y trace) were excluded.

4.9.2 Data quality analysis

As described in Chapter 3, spatial precision is defined as the consistency in detecting and calculating gaze points accurately. Low spatial precision can seriously affect the detection of fixations (Holmqvist et al., 2011, 2012) and hence it is essential to measure and report it.

There are a number of methods to calculate spatial precision, such as the RMS (commonly used by manufacturers) or the standard deviation, which measures the dispersion of each sample from a mean value (see Chapter 3). For the present study, the RMS was calculated for each participant.

In low quality data there is less probability of finding clean long fixations that can be reliably detected even if the data is hand-coded. For this reason one could expect to find negative

correlations between fixation durations and precision measures, where participants producing lower data quality would also feature proportionally shorter fixations.

Correlational analysis between RMS values and fixation durations were run in order to ensure data quality was not interfering with the fixation detection method. There were no negative correlations between data quality and fixation durations for the 6-month-olds nor for the adults confirming that data quality did not affect the experimental results of the present study.

4.9.3 Cross-validation

4.9.3.1 Fixation durations

Manual coding always involves an evaluation of the degree of agreement between different raters. An external coder naïve to the expected outcomes was trained to code fixations from infants featuring low and high quality data. The coder had to (1) run GraFIX automatic detection algorithms using the parameters described in Chapter 3, and (2) manipulate the resulting outcome in order to remove artifactual fixations or add those undetected by following the predefined guidelines described earlier. In total, the external coder coded 10 % of the data analysed for this study.

The inter-rater reliability between the mean fixation durations was evaluated using the intra-class correlation coefficient (ICC; Hallgren, 2012). An excellent agreement was found (with an ICC of .898, p = .05).

4.9.3.2 Gap-overlap

All trials from the gap-overlap paradigm in the group of infants were reviewed in order to manually exclude those where (1) the accuracy was not good enough to trigger the gaze-contingencies, or (2) the infant looked away during the presentation of the peripheral stimulus. An external coder naïve to the expected outcomes evaluated the validity of the gap-overlap trials for 10% of the data from this study.

Inter-rater reliability was measured using the Kappa statistic with a view to determine consistency among raters. This analysis showed a strong agreement between the two coders, Kappa = .878 (p < .001), 95% CI (.833, .923).

4.9.4 Statistical analysis

The mean fixation duration for each participant and viewing condition was calculated after excluding all fixations with a duration greater than two standard deviations from the initial mean. Hence, fixations included in the analysis accounted for about 95% of the set for each participant. This procedure was necessary to exclude very long or very short fixations on a participant basis rather than establishing a minimum and maximum fixation duration threshold for all participants together regardless of their age. Please note that this criterion was established in order to exclude those very long fixations (e.g., 10 seconds) that were rare and a consequence of the participant's sleepiness rather than a result of a cognitive or neural process. This did not exclude the long fixations that are typical in infant populations.

For the group of infants, the gap-overlap and double-step trials with latencies shorter than 200 ms were excluded from the analysis because such fast eye-movements are considered to be eye-tracking errors or anticipatory saccades begun prior to stimulus onset (e.g., Canfield, Wilken, Schmerl, & Smith, 1995; Frick et al., 1999; Rose, Feldman, Jankowski, & Caro, 2002). On the other hand, for the group of adults the minimum latency was lowered to 80 ms since latencies below this value are considered to be anticipatory or express saccades (Fischer & Weber, 1993).

Some gap-overlap studies also define a maximum latency limit (e.g., Elsabbagh et al., 2009; Johnson et al., 1991; Matsuzawa & Shimojo, 1997). However, given that our two groups are very apart in age, one can expect to find group and inter-individual differences in their disengagement abilities and reaction times. For this reason I decided to establish a conservative criterion excluding all the gap-overlap trials that were one standard deviation above each participant's mean latency.

To avoid misleading results, only the double-step trials in which the saccade was executed shortly after the second target appeared were included.

4.10 Results

4.10.1 Age differences

An independent samples t-test showed how the group of adults (naturalistic: M = 327 ms, SD = .045; semi-naturalistic: M = 377 ms, SD = .069) showed significantly shorter mean fixation durations than the group of 6-month-olds (naturalistic: M = 644 ms, SD = .142; semi-naturalistic:

M = 724 ms, *SD* = .120) for both naturalistic (*t*(11.237) = 7.205, *p* < .001) and semi-naturalistic conditions (*t*(27) = 9.928, *p* < .001).

4.10.2 Changes in fixation durations across viewing conditions

Paired samples *t*-tests showed longer mean fixation durations for the semi-naturalistic condition (6-month-olds: *M* = 724 ms, *SD* = .120; adults: *M* = 377 ms, *SD* = .069) compared to the naturalistic condition (6-month-olds: *M* = 644 ms, *SD* = .142; adults: *M* = 327 ms, *SD* = .045), for both 6-month-olds (*t*(10) = -3.980, *p* = .003) and adults (*t*(17) = -3.912, *p* = .001; see Figure 4-10).

These results provide evidence that even though at 6 months infants' oculomotor control may not have reached adult levels, infants' fixation durations are already influenced by the viewing condition. Differences in fixation durations between the two age groups aside, infants responded in the same manner as adults, showing shorter mean fixation durations for the naturalistic stimuli compared to the semi-naturalistic ones.
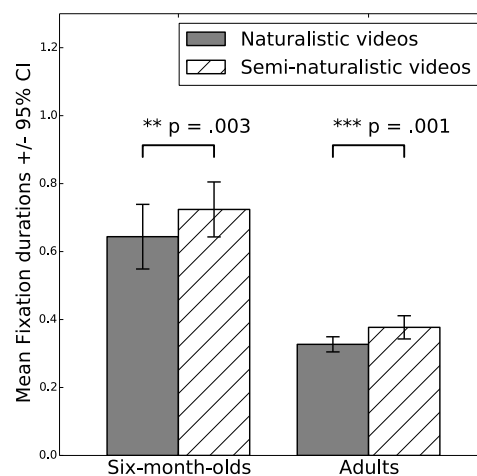
Figure 4-10 Mean fixation durations for the naturalistic and semi-naturalistic viewing conditions for infants and adults

4.10.3 Fixation durations and disengagement

Both age groups showed a very clear gap-effect (6-month-olds: *t*(10) = -4.795, *p* = 001; adults: *t*(17) = -6.004, *p* < .001). Disengagement was calculated for each participant by subtracting the baseline mean latency to the overlap mean latency.
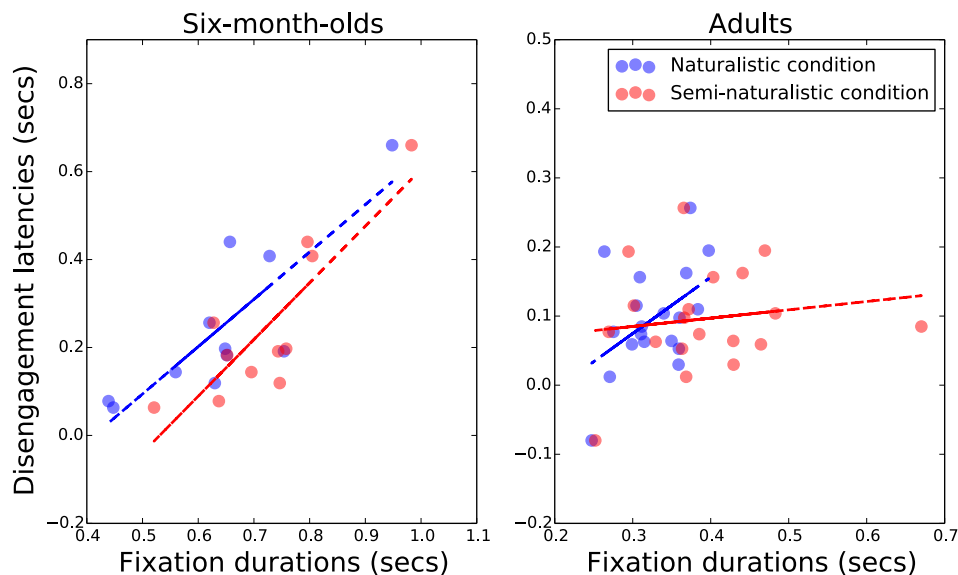
Figure 4-11 Correlations between disengagement latencies and fixation durations in infants and adults

High correlations were found for 6-month-olds between disengagement latencies and mean fixation durations in both naturalistic (r(11) = .836, p = .001) and semi-naturalistic conditions (r(11) = .851, p = .001). Moreover, there was a relationship between disengagement latencies and mean fixation durations in the naturalistic condition (r(18) = .471, p = .049) in the group of adults, but not between disengagement latencies and mean fixation durations in the semi-naturalistic condition (r(18) = .251, p = .315; see Figure 4-11). These correlations were also calculated after applying log transformations for mean fixation durations in naturalistic and semi-naturalistic stimuli in order to correct for the non-normal (ex-Gaussian) fixation duration distribution, and the same pattern was found. These results indicate the influence that the ability to disengage from a focal point may have on fixation durations. This influence was not only present at 6 months, but could also be seen to a lesser extent in adults.

4.10.4 Individual differences in fixation durations

Participants showing short or long mean fixation durations in one viewing condition also showed proportionally short or long mean fixation durations in the other condition, indicating stable individual differences in both 6-month-olds (r(11) = .882, p < .001) and adults (r(18) = .615, p = .007).

137

4.10.5 Double-step paradigm: Saccadic cancelation

Results from the double-step paradigm were analysed by comparing the reaction times for the trials where the infant executed a single saccade to the second target, or two consecutive saccades to the first and the second targets. It was hypothesized that when infants performed a single saccade to the second target they could be cancelling the initial saccade program to the first target before starting the new saccade program to the second target, consequently prolonging the reaction time when moving their eyes away from the central target (see Figure 4-12). According to this hypothesis, the reaction times from the trials in which the infant performed a single saccade should be longer than those found when two consecutive saccades were executed. To avoid misleading results, only the trials in which the saccade was executed shortly after the second target appeared were included.
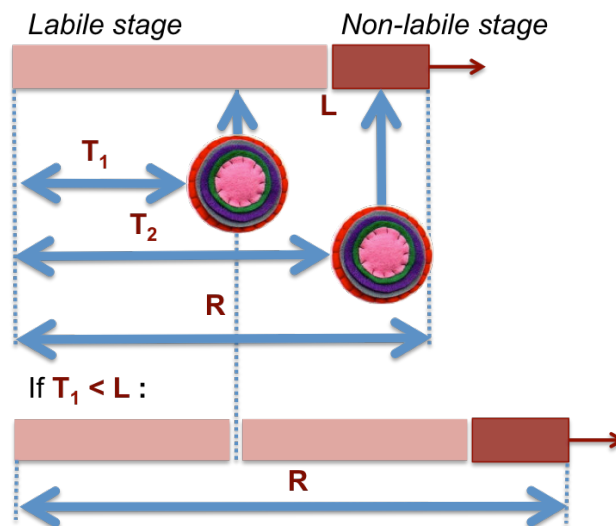
Figure 4-12 A saccade program is cancelled if the second target appears during its labile stage. When this happens, the reaction time is prolonged. If the second target appears during the non-labile stage, the initial saccade program is executed followed by a second saccade to the second target.

As predicted, reaction times when infants performed a single saccade to the second target ($M = 636$ ms, $SD = .136$) were longer than when they executed two consecutive saccades to the first and the second targets ($M = 548$ ms, $SD = .060$; $t(10) = 3.225$, $p = .009$), suggesting that they were able to cancel a saccade program.

4.11 Discussion

The study of fixation durations is becoming an increasingly important technique to investigate attentional and cognitive processes in unconstrained ecologically valid settings. Though fixation durations and scene perception in infancy are still largely unexplored, a few studies have highlighted how fixation durations can be influenced by perceptual and cognitive activity (Bronson, 1990, 1994; Harris et al., 1988) or by the developmental state of the visuo-motor system ("sticky fixation"; e.g., M. H. Johnson et al., 1991). It is still unclear, however, what the mechanisms underlying saccadic control in infancy are and the extent to which large variations in fixation duration distributions found during the first months of life are due to an underdeveloped visuo-motor system or the speed of visual and cognitive processing. The present study aimed to shed light on these questions and contribute to a better understanding of the underlying mechanisms of eye-movement control at 6 months of age.

For the present study, I recorded eye-movements from groups of 6-month-old infants and adults while they were presented with customized naturalistic and semi-naturalistic videos. All participants performed a gap-overlap task (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991) in order to measure their disengagement abilities, and a double-step task (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954) aiming to investigate saccade cancellation in 6-month-old infants. The goals for this study were (a) to explore age differences in fixation durations between 6-month-old infants and adults, (b) to investigate the influence of the viewing condition—and hence visual and cognitive processing—in fixation duration distributions, (c) to examine the relationship between disengagement abilities—which can be indicative of the general development of the visuo-motor system—and fixation durations, (d) to explore individual differences in fixation durations in 6-month-olds and adults, and (e) to demonstrate that 6-month-olds are able to cancel a saccade program. These results will form the basis for modelling the data in the next chapter.

As predicted, results showed age differences in mean fixation durations between 6-month-olds and adults, where the group of adults displayed significantly shorter mean fixation durations compared to the group of infants (prediction a). These differences could be due to differences in the developmental state of the visuo-motor system and/or a consequence of age variations in visual and cognitive processing speed. In line with previous findings I found stable individual

differences across viewing conditions in both infants and adults (prediction d). In other words, infants showing long or short mean fixation durations in one viewing condition also showed proportionally long or short fixations in the other condition. These findings suggest that 6-month-old infants already have a well developed top-down control that modulates gaze allocation.

Additionally, results revealed significant differences in mean fixation durations between naturalistic and semi-naturalistic viewing conditions for both age groups (prediction b). These results fit with previous findings (e.g., Hunnius & Geuze, 2004) demonstrating that also at 6 months the characteristics of the visual stimuli—and hence visual and cognitive processing—influence fixation durations. Interestingly, infants and adults showed shorter mean fixation durations for the naturalistic condition, which was visually and semantically more complex as compared to the semi-naturalistic condition, in which very simple geometrical shapes were presented. This raises the question of whether longer fixation durations are reflecting the higher cognitive requirements of processing a more complex visual stimulus, or if on the other hand they are not necessarily reflecting any cognitive activity related to the stimulus being viewed. To explain the shorter fixation durations in the naturalistic condition one could argue that the visual competition in the naturalistic condition is much higher than in the semi-naturalistic condition, thus the viewer needs to execute more and faster eye-movements in order to capture all the competing semantic and perceptual aspects of the scene. According to this hypothesis fixation durations would not only be affected by the processing of the current focal point, but also by the online requirements of the scene or the visual task (e.g., Henderson et al., 1999; Nuthmann et al., 2010; Võ & Henderson, 2009). On the other hand, according to a second hypothesis very long fixations (e.g., > 2 secs), would not necessarily reflect perceptual activity or cognitive processing and consequently these results—for which such long fixations are included—would not be definitive. In fact, some infant researchers have argued that measures of look duration do not reflect information processing directly (Courage et al., 2006; Reynolds & Richards, 2007). According to Reynolds and Richards (2007), who measured look duration together with heart-rate, it is only in the portion of looking that happens during the heart-rate defined phase of sustained attention when the infants engage in actively processing the information. Even though these studies measured look duration, it could be the case that the same is true for fixation durations. Based on this, studies investigating adults' oculomotor control have traditionally excluded the long fixations that can sometimes be found in the data (Inhoff & Radach, 1998).

Nevetherless, these long fixations can be very common in infants, presumably due to difficulties disengaging from a visual stimulus, or less experience processing perceptual and cognitive information. For the present analysis the exclusion criteria for long fixations were applied on a participant basis rather than selecting a random cut-off threshold for all participants, regardless of their individual performance. More precisely, the fixations included in the analysis accounted for about 95% of the set for each participant, allowing the subjects with longer (or shorter) fixations on average to maintain most of their long (or short) fixations. The fact that results also showed individual differences between short and long lookers in both age groups make the first hypothesis—infant fixation durations reflect the processing of the current focal point as well as the online requirements of the scene —plausible. Nevertheless, at this point the second hypothesis – fixation durations do not necessarily reflect information processing directly – cannot be diminished.

Strong correlations were found between disengagement latencies and fixation durations in infants for both naturalistic and semi-naturalistic viewing conditions (prediction c). In adults, a weaker though significant correlation was found for the naturalistic condition. The strong correlations in the infant group might be due to an underdeveloped visuo-motor system. In infants, saccade planning and execution is greatly affected by the development of the frontal eye fields (FEF) and other neural structures such as some areas of the parietal cortex. Thus, infants showing more difficulties disengaging are believed to be those with a less developed visuo-motor system and with longer mean fixation durations. This would not explain, however, the correlations that, though reduced, still exist in the adult group for the naturalistic condition. These results suggest that disengagement is not only a consequence of an underdeveloped visuo-motor system, but that it also affects saccade execution as a function of the characteristics of the visual stimuli being processed. Consequently, disengagement could be used to quantify the stimulus dependency of fixation durations. These results fit with the findings from a study by Kikuchi and colleagues (2011) who found larger disengagement latencies and saccade-related event-related potentials (ERPs) when children disengaged from faces, suggesting that the encoding and processing of the foveated stimulus plays a role in the ability to shift the gaze from a central target.

Whilst the double-step paradigm has been widely used in adult research to investigate saccade programming (see Chapter 1; Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker &

McSorley, 2006; Westheimer, 1954), in infancy research it has only been applied to identify the shift from retinotopic to spatiotopic coordinate systems (Gilmore & Johnson, 1997; Kaufman et al., 2006; Senju et al., 2011). Technical and practical constraints in infant testing constitute a significant difficulty for investigating the development of saccade programming. For instance, in previous studies using the double-step task with adults each participant performed around 576 trials, which means that, in total, more than 10.000 saccadic responses could be recorded (Becker & Jurgens, 1979). This allowed researchers to test different variations in the presentation of the stimuli and obtain distributions of reaction times for different conditions in a single study. Furthermore, since adult participants were instructed to look to the second target when it appeared, the researcher could be certain that the task was being accomplished. For obvious reasons, the studies with infants are comprised of a small number of trials that are not always valid, and consequently, to date it has not been possible to reproduce the adult findings on saccade programing in infant populations. The present study included a version of the double-step paradigm that aimed to overcome some of the issues found in infant studies. The main variation compared with previous implementations of the paradigm was the use of an adaptation of the MOBS thresholding algorithm (Anderson & Johnson, 2006; Tyrrell & Owens, 1988) in order to calculate the time lapse during which the first target is present (τ) for each trial. This method allows focusing on testing the τ values that are most relevant for each individual participant, reducing the total number of trials that need to be executed. Infants were also encouraged to look at the second target by displaying a visually more attractive stimulus and presenting a short animation when it was gazed. Finally, predictive saccades were avoided by presenting the targets at various random. Based on previous findings from adult studies arguing that saccade cancellation prolongs reaction times (e.g., Nuthmann et al., 2010; Walshe & Nuthmann, 2013), I hypothesized that reaction times from the trials in which a single saccade was executed would be larger than those in which two consecutive saccades were performed. This is because in the occasions infants performed a single saccade to the second target it was assumed they were cancelling the initial saccade program to the first target before starting a new saccade program to the second target, thus the ultimate reaction time would be comprised of the duration for the initial, cancelled saccade program and the duration for the second saccade program to the second target. Results supported this hypothesis, suggesting that 6-month-old infants are able to cancel a saccade program (prediction e). Nevertheless, not enough data was collected in order to further investigate—as it was done in the studies with

adults—the notion of the two-step saccade programming based on the effects of the time τ elapsing between the initial saccade and the second target step. Even though these results are encouraging, more research is needed in order to disentangle the mechanisms behind saccade programming in infancy.

4.12 Conclusions

In sum, in this chapter I presented an empirical study that highlighted the developmental differences in eye-movement control between infants and adults, and helped shedding light on some of the factors affecting fixation durations and saccadic control in infancy. The specific mechanisms underlying saccadic control in 6-month-olds, or the extent to which different factors affect fixation durations and saccadic control at this age remains undetermined.

In the next chapter I will model the present empirical data in order to investigate the mechanisms underlying fixation durations and saccadic control in infancy. More specifically, in the next chapter I present three simulation studies that will test whether the data from infants can be explained by a single model architecture (CRISP: Nuthmann & Henderson, 2012; Nuthmann et al., 2010), with age-specific and task-specific influences realized by differences in parameter settings. The model simulations will also be used to test specific developmental hypotheses on oculomotor control at 6 months of age. In particular, I will examine if fixation durations at 6-months are affected by developmental aspects of the visuo-motor system and/or by visual and cognitive processing.

Chapter 5: Investigating the mechanisms underlying fixation durations in 6-month-olds: A computational account

5.1 Introduction

In the previous chapter I investigated the factors modulating fixation durations and saccadic control in 6-month-old infants and adults when viewing complex dynamic stimuli and performing simple oculomotor tasks such as the gap-overlap paradigm (e.g., Fischer & Boch, 1983; Fischer & Breitmeyer, 1987; Fischer & Ramsperger, 1984; Saslow, 1967) or the double-step paradigm (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954). Results revealed differences in mean fixation durations across age groups (6-month-olds vs. adults) and viewing conditions (naturalistic vs. semi-naturalistic videos), and a relationship— particularly evident in the group of infants—between the ability to disengage from a focal point and mean fixation durations. As shown in the differences across viewing conditions, these findings suggest that infant fixation durations reflect on-line perceptual and cognitive activity in a similar way to adults, but that the individual developmental state of the visuo-motor system still affects this relationship at 6 months, as implied by the high correlations between disengagement latencies and mean fixation durations in infants. Moreover, results from the double-step paradigm suggest that by 6 months of age infants are able to reliably cancel a saccade program. Even though these empirical findings disclosed some of the factors affecting fixation durations and saccadic control in 6-month-olds, the behavioural nature of the study did not allow investigating the precise mechanisms underlying oculomotor control in infancy.

The present chapter aims to utilize a theoretical framework and computational model of fixation durations in scenes and determine its generalizability to infants (CRISP: Nuthmann & Henderson, 2012; Nuthmann et. al, 2010), in order to investigate previously unexplored mechanisms underlying oculomotor control in 6-month-olds. Notably, the model assumes that saccades are programmed in two stages, an initial labile phase during which saccade programs can be altered or cancelled and a subsequent non-labile phase in which programs cannot be cancelled. Whether infant saccadic programming operates via these two phases is not known. In addition, the CRISP model will be used to test specific developmental hypotheses on oculomotor control at 6 months of age. In the present chapter I will examine to which extent fixation durations at this age are affected by developmental aspects of the visuo-motor system

and/or by visual and cognitive processing. I also analyse the CRISP model's limitations when applied to infant data. For this purpose I describe three simulation studies where I will model the fixation duration data from 6-month-old infants and adults presented in Chapter 4, in which participants viewed naturalistic and semi-naturalistic dynamic scenes. These studies will test whether the data from both infants and adults can be explained by a single model architecture, with age-specific and task-specific influences realized by differences in parameter settings (Nuthmann et al., 2010). In the following sections I will introduce the background literature on modelling fixation durations with CRISP.

5.2 Modelling fixation durations with CRISP

Empirical studies alone may be insufficient to answer questions related to the extent to which longer fixation durations reflect a less developed visuo-motor system, or differences in the speed of visual and cognitive processing. By explaining the empirical data and predicting possible outcomes computational models of visual attention provide us with another perspective that can help disentangle questions about the mechanisms underlying eye-movement control in adults and infants (Schlesinger & McMurray, 2012).

Most current models of visual attention have focused on explaining where viewers fixate (for a review see Borji & Itti, 2013). Many of these models have implemented various versions of the visual saliency hypothesis, where bottom-up factors such as the colour or the luminance of the scene, drive gaze allocation (Itti & Koch, 2000, 2001). More recent theories, however, have highlighted the fact that visual attention may be primarily driven by top-down factors such as the viewers task (Henderson et al., 2007).

In comparison, only a few models provide control mechanisms for the duration of fixations to shed light on the mechanisms underlying saccade generation and the encoding of visual information. Most of these models were developed for the task of reading. Here, the two most advanced models are the E-Z Reader model (Reichle et al., 1998, 2003) and the SWIFT model of saccade generation during reading (Engbert et al., 2002, 2005). Both models implement the notion of saccade programming being completed in two stages, as suggested by results from double-step experiments (Becker & Jurgens, 1979; see also Chapters 1 and 4). However, the two models differ with regard to the mechanisms that control fixation durations. What triggers a new saccade program is conceptualized to be lexical processing (E-Z Reader) or a random

timer (SWIFT). In the SWIFT model, lexical processing difficulty of the currently fixated word modulates fixation durations by inhibiting the timer so that it delays the initiation of the next saccade program.

Building upon the idea of a random saccade timer (Engbert et al., 2002) the CRISP model is a computational model and a theoretical framework that accounts for fixation durations in adult scene viewing (Nuthmann et al., 2010). CRISP is a timer (C)ontrolled (R)andom-walk with (I)nhibition for (S)accade (P)lanning model (see Figure 5-1 for a model overview).
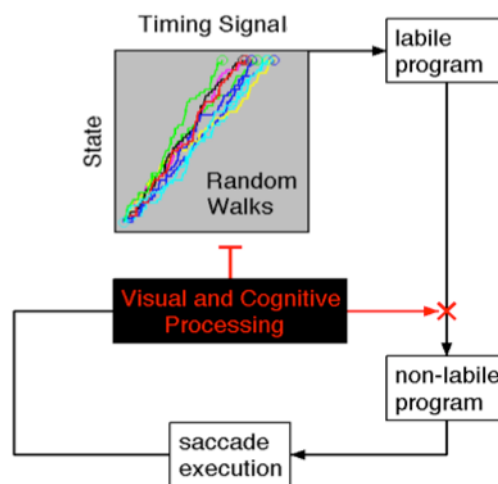
Figure 5-1 The CRISP Model overview. From Nuthmann et al. (2010).

The model architecture can be summarized with the following three main assumptions on saccade timing and programming:

(1) The accumulation of activity to a saccade threshold is implemented via a random walk process (Boccignone & Ferraro, 2004) and is responsible for generating inter-saccadic intervals and hence variations in fixation durations.

(2) Saccade programming occurs in two stages: an initial, labile stage that is subject to cancellation, and a later, non-labile stage (see Chapters 1 and 4).

(3) Processing difficulty can inhibit saccade timing and programming in a moment-to-moment fashion. The latter can happen in two ways. First, in case of increased processing demands the

random walk saccade timer slows down, which delays the initiation of a new saccade program and eventually leads to longer fixation durations. Secondly, processing difficulties can even cancel an ongoing labile saccade program, which extends the duration of the current fixation.

Figure 5-2 From Nuthmann et al., 2010, summarizes the CRISP model's architecture by displaying its temporal scheme on saccade timing and saccade programming.
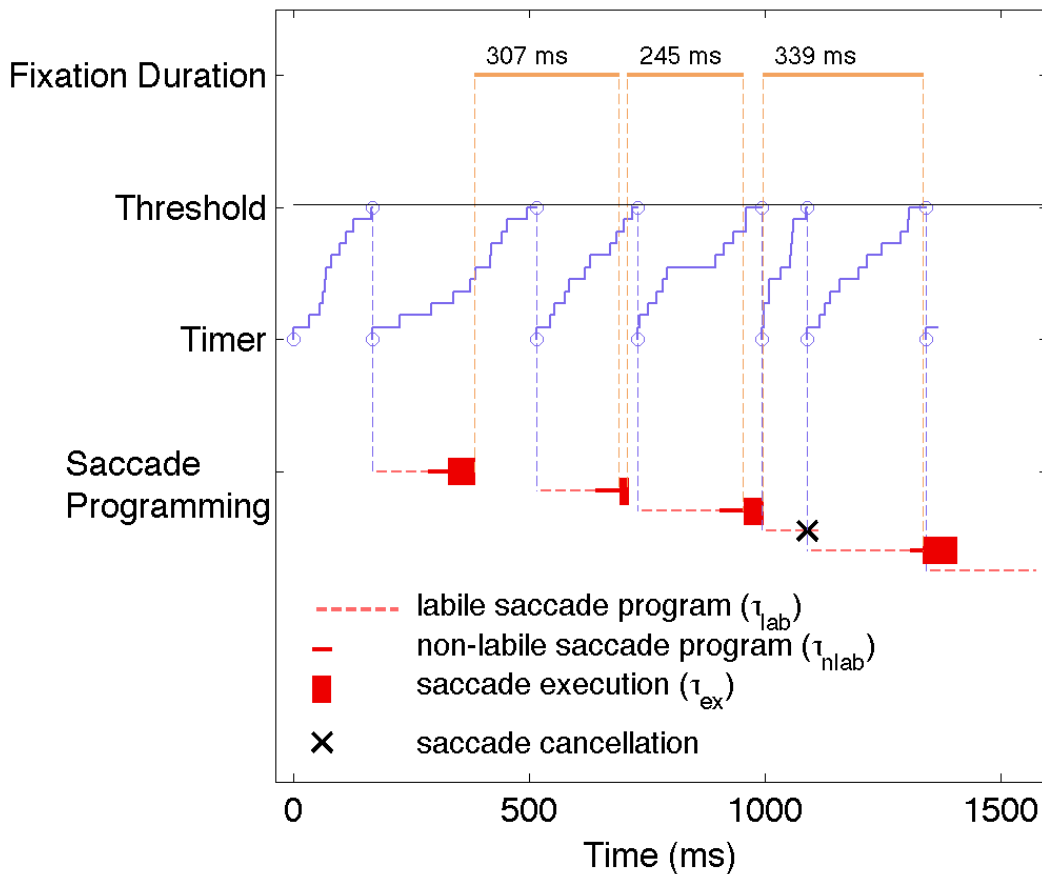
Figure 5-2 From Nuthmann et al. (2010). Temporal scheme of saccade timing and saccade programming in CRISP. The random walk timing signal accumulates until it reaches a threshold, where a new saccade program is initiated. Each saccade program is composed by an initial labile stage (τ_{lab}), during which it is subject to cancelation, and a subsequent non-labile stage (τ_{lab}), where the saccade program can no longer be cancelled. After the non-labile stage, the saccade is executed (τ_{ex}). Fixations start after a saccade is executed and end with the execution of the subsequent saccade. The random walk timing signal for the current saccade may not have reached its threshold before the saccade was executed (see the fixation that lasts 307 ms). This will prolong the duration of the subsequent fixation.

The random walk timing signal accumulates towards a threshold and when it is reached a new saccade program is initiated. Each saccade program starts with its labile stage (τ_{lab}), during

which it can be probabilistically cancelled, followed by its non-labile stage (τ_{nlab}), where the saccade program can no longer be cancelled. Once this point is reached, the saccade is executed (τ_{ex}). Fixations occur during the time lapse between two consecutive saccade executions. Occasionally, when the random walk timing signal does not reach its threshold before the saccade is executed (see Figure 5-2, the fixation that lasts 307 ms), the duration for the subsequent saccade is prolonged. Another source for increasing fixation durations is saccade cancelation (see Figure 5-2, the fixation that lasts 339 ms). In these cases the ultimate duration of the fixation includes the time lapse for the cancelled saccade program and the program that was executed (Nuthmann et al., 2010; Walshe & Nuthmann, 2013; see also Chapter 4).

Mathematically, in its basic form the CRIPS model comprises five parameters related to saccade timing and programming (t_{sac}, N, τ_{lab}, τ_{nlab}, τ_{ex}). This model still includes fewer parameters than other models that also investigate eye-movement control such as the E-Z Reader model (Reichle et al., 1998, 2003), the SWIFT model (Engbert et al., 2005) or the ICAT model (Trukenbrod & Engbert, 2014). The random walk timing signal accumulates toward a threshold, and once the threshold is reached, a new saccade program is initiated. The transition rate for the random walk's increments determines how fast the process of saccade timing operates. The transition rate r_1 is defined as

Equation 1

$$r_1 = \frac{N}{tsac}$$

where t_{sac} is the mean duration of the timing signal (i.e., the mean time interval between two commands to initiate a saccade program). N denotes the number of states the random walk process can adopt and determines the variance of the timing signal; a low N leads to high variance and thus broad fixation duration distributions. The random walk is implemented as a discrete-state continuous-time Markov process with exponentially distributed waiting times between elementary transitions (for further details, see Nuthmann et al., 2010).

Furthermore, saccade latency l_{sac} is derived as the sum of the implemented saccade programming stages, i.e.,

Equation 2

$$l_{sac} = \tau_{lab} + \tau_{nlab}$$

where τ_{lab} and τ_{nlab} denote the average duration of the labile and non-labile phases, respectively. At the end of the non-labile saccade programming stage, the saccade is executed (with average duration τ_{ex}). Saccade programming processes are assumed to be stochastic in nature. Therefore, for each realization of the model simulation, parameter values for the different saccade programming phases (τ_{lab}, τ_{nlab}, and τ_{ex}) are drawn from gamma distributions. This introduces unsystematic variability in the duration of saccade programming stages. The mean parameter value for the saccade execution stage (τ_{ex}) could be estimated from the mean saccade duration in the empirical data. For the remaining parameters, best-fitting values can be determined using an optimization technique such as genetic algorithms. It is important to ensure that the range of possible values for each parameter is psychologically and/or neurophysiologically plausible. In its implementation, the model generates sequences of saccades and fixations, whereby fixation durations are the time intervals between successive saccades (excluding saccade execution).

5.3 Modelling infant fixation durations with CRISP: Theoretical assumptions

In the past decades few computational models have attempted to explain cognitive development in infancy (Mareschal, 2010; Schlesinger & McMurray, 2012). Even though some of these models have investigated different aspects of visual orienting in infancy such as gaze following (Bugajska et al., 2009; Carlson & Triesch, 2004; Triesch et al., 2006), visual selective attention (Schlesinger et al., 2007) or familiarization (Domsch et al., 2010; Mareschal, 2010; Sirois & Mareschal, 2002), to date no computational model has attempted to explain the specific mechanisms underlying saccadic control in infancy (see Chapter 1).

As outlined in Chapter 4, it has been reported how infant fixation durations can be affected by developmental aspects of the visuo-motor system, or by the speed of visual and cognitive processing. As a result one can expect to find more variability in fixation durations in infants compared to adults, or in other words, more positively skewed distributions of fixation durations

(Harris et al., 1988). The CRISP model (Nuthmann et al., 2010) represents a theoretical framework that may help to disentangle some of the questions introduced in Chapter 4 and at the beginning of this chapter: What are the mechanisms underlying saccadic control in early infancy? To which extent is the variability in fixation durations a consequence of developmental aspects of the visuo-motor system and/or the speed of visual-cognitive processing?

In this section, I revisit CRISP's theoretical assumptions and input parameters with the specific aim to explore whether they are, in principle, applicable to infant data and in what way age differences (infants vs. adults) could be represented by certain differences in parameter settings. From the perspective of age-specific influences on fixation durations, the following CRISP parameters contribute toward systematic differences in fixation durations: (1) the mean (t_{sac}) and variance (i.e., number of states N) of the random walk timing signal, (2) the mean duration of the labile saccade program (τ_{lab}), and (3) the mean duration of the non-labile saccade program (τ_{nlab}) (see Nuthmann et al., 2010).

a)   Random walk timing signal

The CRISP model assumes that saccade programs are initiated according to some preferred idiosyncratic mean rate (Engbert et al., 2002) and based on this it implements a rhythmic saccade timer that is responsible for generating variations in fixation durations (Nuthmann et al., 2010). All living things are subject to spontaneous fluctuations of physiological and behavioural processes that maintain complex organisms in a fairly stable state (Wolff, 1991). Further, evidence suggests that the central nervous system as well as the oculomotor system are rhythmic in nature (McAuley et al., 1999). The temporal organization of spontaneous movements has also been studied in newborn infants, where it was found that the fluctuations in movement over time were not random but rather rhythmic (Robertson, 1982). The same can be interpreted when investigating saccadic eye-movements during the first month of life, which are believed to be mainly under subcortical control and hence are rapid and input-driven (Atkinson, 2000; M. H. Johnson, 1995b, 2011).

b)   *Random walk timing and speed of processing*

In CRISP, decreases in processing speed and increases in processing difficulty will slow down the random walk saccade timer. This delays the initiation of a new saccade program, which in turn leads to longer fixation durations. For the present simulations, it is reasonable to assume

that the mean and variance of the random walk process underlying the initiation of a new saccade program are different for infants compared to adults. Specifically, CRISP would capture a slower speed of visual-cognitive processing in infants by a higher mean value t_{sac} for the random walk timing signal. In addition, following previous viewing-task simulations with the CRISP model (scene memorization vs. search: Nuthmann et al., 2010; scene viewing vs. reading: Nuthmann & Henderson, 2012), I allow both parameters of saccade timing (t_{sac} and N) to vary across viewing conditions. In sum, I predict that these values will vary across viewing conditions and age groups and hence will be indicative of the speed of visual and cognitive processing.

c)   Two-stage saccade programming.

In CRISP, once the random walk process reaches threshold, a new saccade program is initiated. Saccade programming is completed in two stages: an initial, labile stage that is subject to cancellation and an ensuing, non-labile stage in which the program can no longer be cancelled (see Chapters 1 and 4). The notion of the two-stage saccade programming is motivated by findings from double-step experiments (e.g., Becker & Jurgens, 1979), which used much simpler situations than scene viewing.

To date, no study has specifically investigated the two-step notion of saccade programming in infants (but see Chapter 4). However, the ability to inhibit a saccade has been investigated by using the anti-saccade paradigm (e.g., M. H. Johnson, 1995; see Chapter 1). In this task, after the appearance of a central stimulus a cue appears either on the left or right side. The infant is encouraged to look in the opposite direction by subsequently presenting a more attractive object at this location. Johnson (1995) found that by 4 months, infants were able to, after a training period, reliably inhibit the saccade to the cue and move their eyes to the second more attractive target instead. Interestingly, the ability to inhibit a saccade programming concurs with the major development of the premotor areas of the frontal lobes -which contain the frontal eye fields (FEF)- that occurs from 3 to 4 months (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011). These findings suggest that saccade cancellation and hence the two stages of saccade programming may be present already by 4 months. Thus, assuming that infants may not be able to systematically cancel a saccade before 4 months, one could expect the non-labile program to be longer than the labile program before this age. From 4 months onwards, I predict (1) a relative increase of the labile stage with respect to the non-labile stage

program, and (2) a gradual decrease in the absolute times for both labile and non-labile programs as the infant's saccadic control increases in efficiency. Furthermore, research has shown that even by 6 months infant saccadic control is not as efficient as in adult participants (Butcher et al., 2000; Csibra, Tucker, & Johnson, 2001; see Chapter 4), thus both labile and non-labile stages are likely to be longer at this age compared to adults. Moreover, results from the double-step paradigm described in Chapter 4 suggested that 6-month-old infants are able to reliably cancel a saccade program.

Under the assumption that both the existence and duration of the labile (τ_{lab}) and non-labile (τ_{nlab}) stages of saccade programing are subject to developmental changes in the visuo-motor system, I will treat them as an indicator of the system's developmental state. Longer labile and non-labile stages will, on average, generate longer fixation durations.

As outlined earlier in this chapter, CRISP introduces stochastic (i.e., unsystematic) variability in the durations of the labile and non-labile stages in that, for each simulated saccade program, individual durations are drawn from gamma distributions with means τ_{lab} and τ_{nlab}, respectively. The question arises whether there are additional systematic age-dependent and/or task/viewing condition-dependent differences in the duration of saccade programming stages. Moreover, there is evidence suggesting that saccade programming and hence saccade latencies could be affected by various exogenous factors such as those observed in the gap-overlap paradigm (M. H. Johnson et al., 1991; Kopecz, 1995; Matsuzawa & Shimojo, 1997), or endogenous factors such as those emerging from certain task instructions (Castelhano, Mack, & Henderson, 2009; Nuthmann et al., 2010). However, for reasons of model parsimony I assumed that saccade programming parameters do not vary across viewing conditions within a given age group.

5.4 The current studies

The current chapter includes three simulation studies with the CRISP model (Nuthmann et al., 2010) where the fixation durations data from Chapter 4 were modelled. The overarching goal was to establish a theoretical framework able to explain the unexplored mechanisms underlying oculomotor control in 6-month-olds, by demonstrating the generalizability of the CRISP theoretical framework to data from 6-month-olds, and more particularly, for the first time introducing the two-step notion of saccade programming in infants. In addition, simulation

studies 2 and 3 were used to test specific developmental theories on saccadic control and explore the limitations of the CRISP adult theoretical framework when applied to infant data.

In simulation study 1 the empirical data from adults and infants were modelled and compared. In simulation study 2, data from two groups of infants with long and short disengagement latencies were modelled. Finally, using this data simulation study 3 included a series of studies, in which certain input parameters were held constant across the two infant groups while others were allowed to vary in order to test different theoretical assumptions.

The goal of simulation study 1 was to: (a) demonstrate the CRISP model and theoretical framework's generalizability –such as the two-step notion of saccade programming– to data from 6-month-old infants and (b) analyse the differences between infant and adult simulations by means of differences in the model's input parameters. Simulation studies 2 and 3 tested particular developmental theories on saccadic control. More specifically, these studies explored whether (c) fixation durations in infancy are influenced by the developmental state of the visuo-motor system and/or by visual and cognitive processing speed. Given that difficulty in disengagement is commonly associated with poorer developmental states of the visual system, I hypothesized that the absolute and relative values for the labile and especially non-labile stages of saccade programming would be prolonged in infants with longer disengagement latencies.

5.5 Simulation study 1: Baseline

The goal of the first simulation was to test the CRISP model's generalizability to fixation-duration data from 6-month-old infants, which includes the two-step notion of saccade programming. Furthermore, I analysed the age differences and the influence of the material viewed on fixation durations during scene viewing in both infants and adults.

5.5.1 Behavioural data

For the current simulation study the infant and adult data from Chapter 4 for both naturalistic and semi-naturalistic viewing conditions were modelled. In this subsection I will summarize the results that are relevant for the present simulation study. Significant differences between mean fixation durations were found across viewing conditions and age groups. In particular, the group of adults presented significantly shorter mean fixation durations than the group of 6-month-olds for both naturalistic and semi-naturalistic conditions. The semi-naturalistic condition displayed

longer mean fixation durations compared to the naturalistic condition, for both 6-month-olds and adults. Participants showing short or long mean fixation durations in one viewing condition also showed short or long mean fixation durations in the other condition, indicating stable individual differences in both age groups. These results provided empirical evidence that even though at 6 months infants' oculomotor control may not have reached adult levels, infants' fixation durations are already influenced by the viewing condition. Differences in fixation durations between the two age groups aside, infants responded in the same manner as adults, showing shorter mean fixation durations for the naturalistic stimuli compared to the semi-naturalistic ones. The empirical data set was comprised of 12315 adult fixations ($N_{naturalistic}$ = 5949; $N_{semi\text{-}naturalistic}$ = 6366) and 2675 infant fixations ($N_{naturalistic}$ = 1447; $N_{semi\text{-}naturalistic}$ = 1228).

5.5.2 Model adjustments

The goal of these simulations was to test the CRISP model's generalizability to infant data. To this end, the same model architecture was applied to both the infant and adult data. Thus, model generalizability was analysed in the restricted sense of parameter changes, which is a more stringent test than adding new parameters to the model (Nuthmann & Engbert, 2009). Based on the theoretical and empirical considerations outlined at the introduction of this chapter, it was decided which model parameters were allowed to vary across age groups and/or viewing conditions.

As explained earlier, the mean (t_{sac}) and variance (N, number of states) of the random walk timing signal define the transition rate (r_1, see Equation 1) for the random walk process, which determines how fast the process of saccade timing operates. For the present simulations I allowed the parameters t_{sac} and N to vary across different viewing conditions and age groups. Parameter estimates for the different conditions serve as an index for the speed of visual and cognitive processing.

When the model simulates a given saccade program, the durations of the labile, non-labile and execution stages are drawn from gamma distributions (see also Engbert et al., 2005; Reichle et al., 1998). For example, the actual time required to complete the labile saccadic programming stage is sampled from a gamma distribution with $\mu = \tau_{lab}$ and $\sigma = \sigma_\gamma \times \mu$. The means of the gamma distributions (μ) are either free or fixed parameters (Nuthmann et al., 2010), whereas the relation between the standard deviation and mean (σ_γ) is fixed (e.g., at 0.33 or 0.25,

Nuthmann et al., 2010). As outlined above, the means for the labile and non-labile stages are free parameters, which I allow to vary across age groups while they are constant for viewing conditions within a given age group. The parameter for the standard deviation of the gamma distributions (σ_γ) is a fixed parameter. To accommodate the higher variability generally observed in infant data compared to adult data, it was set to 0.33 for the infant data and 0.25 for the adult data. These values were adopted from previous model simulations (Engbert et al., 2005; Nuthmann et al., 2010; Reichle et al., 1998, 2003).

Finally, I treat the mean duration of saccade execution (t_{ex}) as a fixed parameter. This is a parameter that directly depends on the saccadic amplitude (Hainline et al., 1984). For adults I adopted a parameter value used in previous CRISP simulations (t_{ex} = 37 ms, simulation study 2 in Nuthmann et al., 2010). For the infant simulations I slightly increased this parameter value (t_{ex} = 50 ms). Hainline and colleagues (1984) investigated the relationship between saccadic amplitude and velocity in infants and adults. They reported that the average peak velocity for a 10° saccade for infants or adults was around 250-300 °/sec, which leads to mean saccade durations of 33-40 ms. Most of the saccade amplitudes from our study were less than 5°. According the Hainline and colleagues' (1984) main-sequence plots, it can be seen that at around 5°, peak velocities can reach 100-150 °/sec which would lead to saccade durations of around 33-50 ms. These results fit with the values we used for the saccade execution parameter. Each model simulation replicates the results from 10 fictive participants.

5.5.3 Modelling results

Free parameters were estimated using a genetic algorithm optimization technique that was used in previous simulation studies with the CRISP model (simulation study 2 in Nuthmann & Henderson, 2012; Nuthmann et al., 2010). For each age group, a total of six free parameters were fitted: the mean durations for the labile and non-labile stages of programming; the mean and variance of the random walk timing signal for the naturalistic condition; and the mean and variance of the random walk timing signal for the semi-naturalistic condition. The ranges for estimating these values were based on previous oculomotor research from infants and adults, even though I used relatively wide ranges (especially for the group of adults). The genetic algorithm estimated the input parameters by minimizing a goodness-of-fit measure and by evaluating how much the simulated data differed from the empirical data. The parameter values

generating the simulations that best resembled the empirical distributions are described in Table 5-1.

Table 5-1 Best-fitting parameters for simulation study 1

	Parameter	Symbol	Range	Naturalistic	Semi-naturalistic
	Saccade programming				
	Labile stage (ms)	τ_{lab}	200-400	385	
	Non-labile stage (ms)	τ_{nlab}	30-120	62	
	Standard deviation (ms)	σ	---	0.33	
	Saccade timing				
Infants	Mean (ms)	t_{sac}	200-650	458	496
	Variance	N	5-20	20	5
	Saccade execution	t_{ex}	---	50	
	Error	E	---	4.275	3.718
	Error cumulative distribution	E_c	---	4.093	9.821
	Saccade programming				
	Labile stage (ms)	τ_{lab}	150-350	224	
	Non-labile stage (ms)	τ_{nlab}	30-50	37	
	Standard deviation (ms)	σ	---	0.25	
	Saccade timing				
Adults	Mean (ms)	t_{sac}	150-650	253	271
	Variance	N	5-20	20	10
	Saccade execution	t_{ex}	---	37	
	Error	E	---	4.513	6.042
	Error cumulative distribution	E_c	---	6.758	4.248

The goodness-of-fit measure (E) was calculated by assessing how much the simulated data diverged from the empirical data. For this purpose fixation duration distributions and mean fixation durations were compared and the errors were added together (see Nuthmann et al., 2010, Appendix B). The parameters displaying the best fit were selected following the next procedure. First, the deviation between the two distributions was measured by calculating the

mean-root-square error. For each viewing condition, fixation duration distributions were calculated from 46 bins with bin centres ranging from 80 to 6000 milliseconds, in steps of 130 milliseconds. Secondly, the difference between the means was evaluated by the mean-squared normalized error for average fixation durations (Reichle et al., 1998). As a result of the long tail and the high skewedness of the fixation duration distributions presented in this chapter, the simulation with the lower E did not necessarily indicate the best fit. Results were considered a good fit when the simulation captured the peak and the tail of the distribution. The fixation duration distributions disclosed in this chapter included particularly long fixations (>2 secs) that are typically excluded from the analysis when studying adults (Nuthmann et al., 2010). These long fixations, however, are not uncommon in young infants and can in fact be revealing when studying the development of oculomotor control. I decided to include these fixations for both age groups in order to maintain consistency in the analysis despite the additional difficulties this entails when fitting the data with the genetic algorithm. For instance, when the empirical distribution has a long tail the simulation that resembles the peak and the tail of the distribution best is not necessarily the one that shows the lower E: in these cases, mean fixation durations for the empirical and simulated distributions can still be far apart, which increases the E value. Thus when the distribution includes long fixations this will affect the parameter estimation from the genetic algorithm. To overcome this issue I selected the set of parameters that (1) fitted the peak and the tail of the empirical distribution, and (2) presented the best goodness-of-fit measure.

Additionally, I report the mean-root-square error of the cumulative distributions (E_c). This function, which gives more weight to the peak of the distribution when calculating the error, constitutes an additional source for evaluating the disparity between distributions. The goodness-of-fit measure for the genetic algorithm was also evaluated using this formula, but results did not vary significantly.

As for the empirical data, the model simulations showed significant differences in fixation durations across viewing conditions and age groups. In particular, the simulations for the group of adults presented significantly shorter mean fixation durations than the group of 6-month-olds for both naturalistic ($t(13.177) = 53.781$, $p < .001$) and semi-naturalistic conditions ($t(18) = 40.460$, $p < .001$). The semi-naturalistic condition displayed longer mean fixation durations compared to the naturalistic condition, for both 6-month-olds ($t(9) = -8.131$, $p < .001$) and adults

(t(9) = -5.858, p < .001). Furthermore, the group of adults displayed significantly shorter mean labile stages (t(18) = 100.050, p < .001) and non-labile stages (t(18) = 160.543, p < .001) as compared to the group of infants. The results of the simulations are summarized in Figure 5-3.
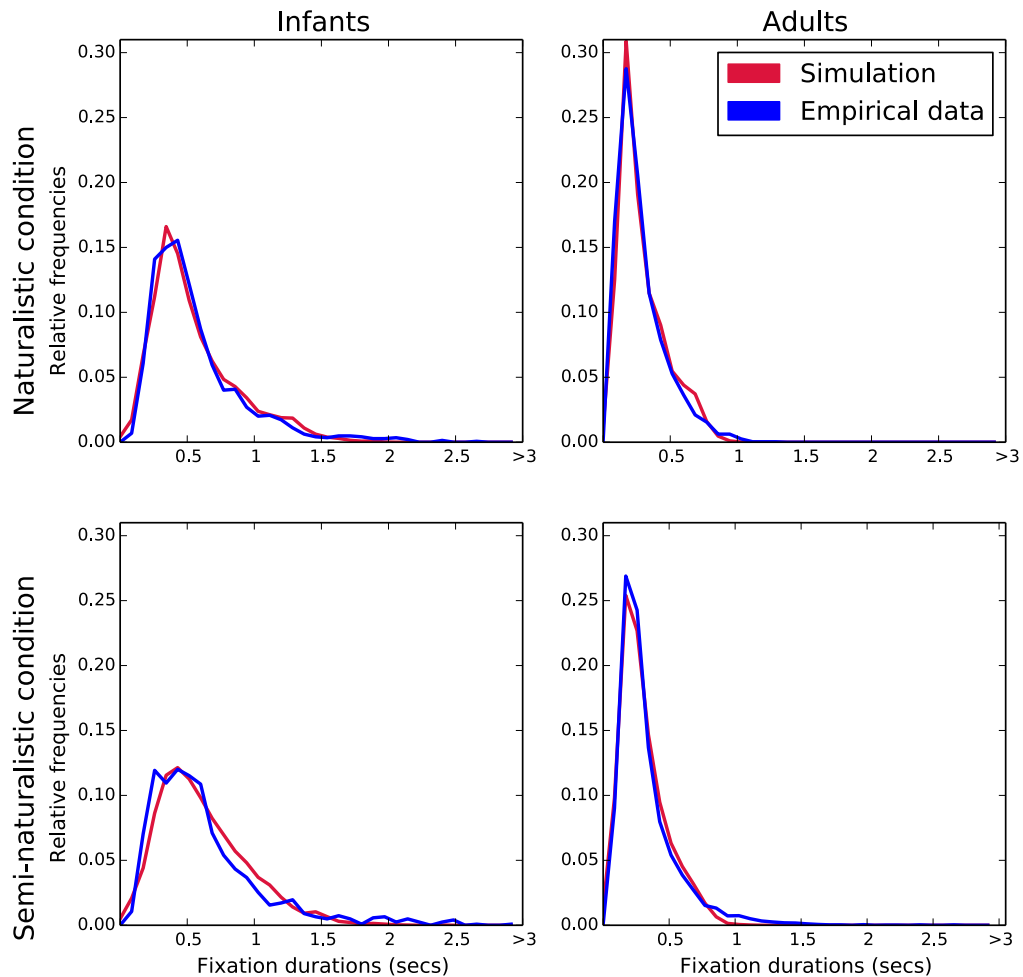
Figure 5-3 Simulation study 1: Simulations paired with the empirical data.

The two right panels show the data from adult observers, which are captured very well by the model simulations. Specifically, the simulated data reproduced the characteristic positive skew in fixation duration distributions. Compared to the adult data, the fixation duration distributions for the 6 month-olds are much noisier, partly due to the fewer number of fixations. For the infant data, the modal portion of the distribution is clearly shifted towards longer fixation durations, and the tail of the distribution is increased. These overall trends are captured well by the simulated data. Table 5-1 displays the best-fitting parameters for adults and infants.

Results evidence differences between age groups in both saccade timing and saccade programming parameters, with the infant group displaying considerably higher values (see Table 5-1). With respect to saccade timing parameters, the mean values for the random walk timer from the infants group (naturalistic = 458 ms; semi-naturalistic = 496 ms) almost doubles the values for adults (naturalistic = 253 ms; semi-naturalistic = 271 ms). The saccade timing variance (number of steps for the random walk, N), on the other hand, do not seem to vary that much across age groups, especially for the naturalistic condition where N is 20 for both infants and adults. This can be a side effect of the model trying to accommodate the very long fixations (> 2 secs) that were included in the analysis, since higher N values will lead to wider distributions (see Appendix A in Nuthmann et al., 2010). The striking differences in saccade timing parameters between infants and adults evidence how visual and cognitive processing speed is considerably slower for 6-month-olds, which prolongs the duration of fixations. Additionally, the model shows that the differences between the naturalistic and semi-naturalistic distributions are mainly due to a lower mean value for the random walk timer and the difference in the saccade timing variance (number of steps for the random walk; see Table 5-1).

Pertaining to saccade programming parameters, the simulation results can be summarized as follows. First, there was no relative increase in the duration of the labile stage with respect to the non-labile stage when comparing infants and adults (the labile stage takes up 86.13% of saccade programming time in infants and 85.82% in adults). Second, there was a gradual decrease with age in the absolute durations of both labile and non-labile saccade programming stages (see Table 5-1). These results suggest that even though two-stage saccade programming is fully functioning at 6 months, it must develop further in order to reach adult-like levels.

From this simulation study I can conclude that the CRISP model's architecture can be applied to infant data, even though the model in its original form has difficulty in dealing with the very long fixations common in infants (> 2 secs). The simulation results suggest that at 6 months saccade programming occurs in two stages (the labile and non-labile stages), which constitutes the first evidence for the two-step notion of saccade programming in infants. Furthermore, the CRISP model was able to capture the differences in fixation durations that resulted from the presentation of different viewing conditions in both age groups, which shows that the influence

of cognitive processing in fixation durations can be reproduced by adapting the random walk timer of the saccade generator.

Regarding age differences, the discrepancies in both saccade programming values (labile and non-labile stages) and saccade timing values (mean of the random walk timing signal) between infants and adults are compelling, suggesting that at 6 months, fixation durations can still require further development of the visuo-motor system to reach adult levels.

5.6 Simulation study 2: Modelling participants showing long and short disengagement latencies

The previous simulation study demonstrated the CRISP model's generalizability to infants. This allows testing in subsequent studies particular developmental theories by applying the model. This simulation study aims to compare the differences in fixation durations between infants showing increased disengagement difficulties and those whose latencies were closer to adults. As discussed in Chapters 1 and 4, long disengagement latencies in early infancy are associated with a poorer development of the visual system. I also assume that infants with less developed visuo-motor systems, i.e. infants with longer disengagement latencies, will present longer labile and/or non-labile stages. This assumption is based on past research using the anti-saccade paradigm, which suggested that infants younger than 3-4 months were not able to readily inhibit a saccade program to the wrong target -and hence, cancel a saccade (e.g., M. H. Johnson, 1995). Comparing the modelling results and parameter estimates for infants with varying disengagement abilities allows for exploring whether the differences in mean fixation durations between the two groups are due to developmental aspects of the visuo-motor system (saccade programming) or rather variations in the speed of visual and cognitive processing (saccade timing).

5.6.1 Behavioural data

For the present simulation study I modelled the infant data presented in Chapter 4 for the naturalistic and semi-naturalistic viewing conditions. In this subsection I will summarize the results that are relevant for the present simulation study. High correlations were found for 6-month-olds between disengagement latencies and mean fixation durations in both naturalistic and semi-naturalistic conditions. Moreover, there was a relationship between disengagement latencies and fixation durations in the naturalistic condition in the group of adults, but not

between disengagement latencies and the semi-naturalistic condition. These results highlighted the influence that the ability to disengage from a focal point may have on fixation durations, which was not only present at 6 months, but could also be seen to a lesser extent in adults.

The model simulations were restricted to the infant data and aimed at comparing data from infants with larger or smaller disengagement difficulties as these infants should demonstrate the most interesting differences in model parameters. Gathering a large amount of data from the same infant entails some technical and practical difficulties (see Chapters 2 and 3). Due to the low number of fixations gathered from each infant it was not possible to model participants individually. For this reason I decided to select two groups of three infants. This was the minimum number of participants needed to obtain an acceptable number of fixations in order to model the data efficiently (> 200 fixations per condition). The long disengagement (LongD) group was comprised of the three infants with the longest disengagement latencies (all of them above the mean). In contrast, the short disengagement (ShortD) group included the three infants with the shortest disengagement latencies (all of them below the mean). As a result of the correlations between disengagement latencies and fixation durations, the infants with long (or short) disengagement latencies also showed long (or short) fixation durations.

The data from the LongD group was comprised of 486 fixations ($N_{naturalistic}$ = 274; $N_{semi-naturalistic}$ = 212). The mean fixation duration was 819.5 ms ($M_{naturalistic}$ = .778, $SD_{naturalistic}$ = .152; $M_{semi-naturalistic}$ = .862, $SD_{semi-naturalistic}$ = .105) and the mean disengagement latency 502 ms (SD = .137).

The ShortD group included 1094 fixations ($N_{naturalistic}$ = 614; $N_{semi-naturalistic}$ = 480). The mean fixation duration was 570 ms ($M_{naturalistic}$ = .505, $SD_{naturalistic}$ = .108; $M_{semi-naturalistic}$ = .635, $SD_{semi-naturalistic}$ = .113) and the mean disengagement latency 87 ms (SD = .029). An independent samples t-test showed significant differences in disengagement latencies when comparing LongD and ShortD groups ($t(4)$ = 5.139, p = .007).

5.6.2 Model adjustments

The model adjustments and assumptions were exactly the same as those described in simulation study 1.

5.6.3 Modelling results

The best-fitting values for free model parameters were again determined with the same genetic algorithm optimization technique described for simulation study 1. Figure 5-4 displays the empirical and simulated fixation duration distributions for the LongD (left panels) and ShortD (right panels) groups and the two viewing conditions (top: naturalistic video, bottom: semi-naturalistic videos).
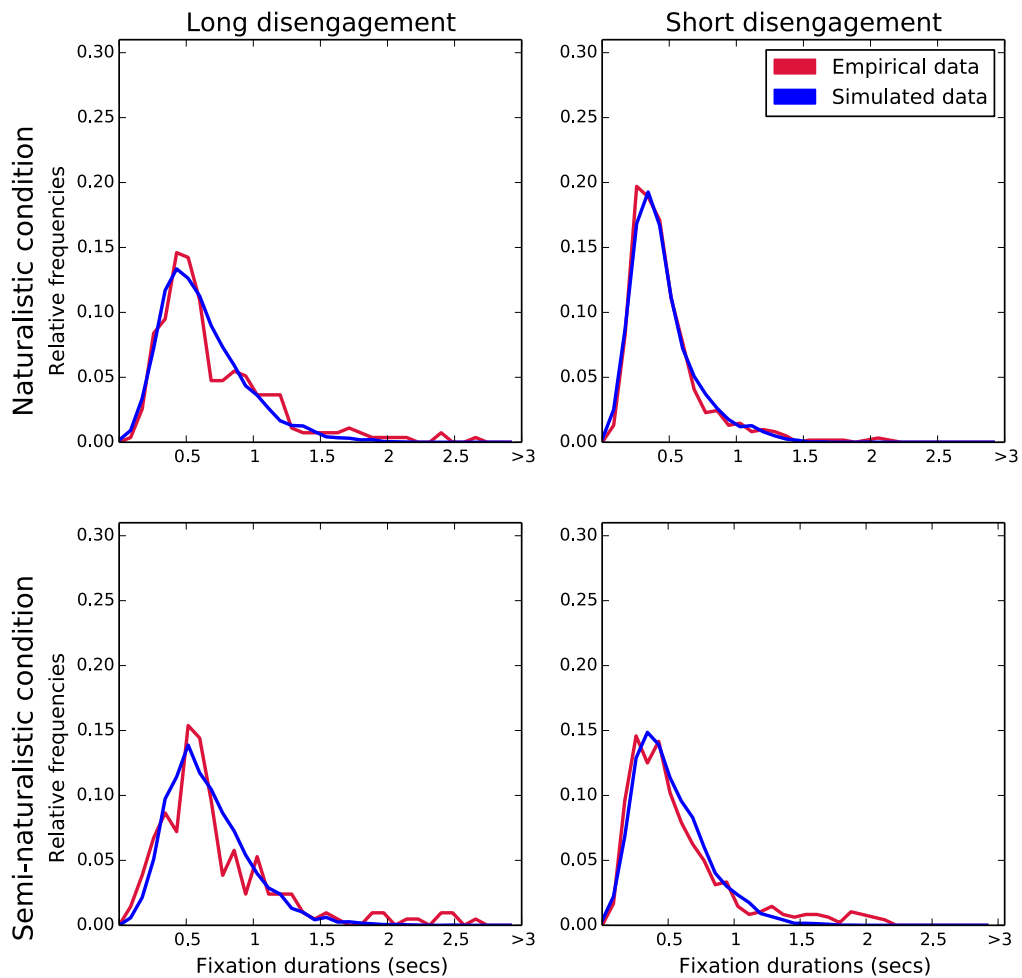
Figure 5-4 Simulation study 2: Simulations paired with the empirical data

The reduced number of fixations for the LongD group led to rather noisy empirical fixation duration distributions (red lines). Yet it is evident from the data that, for the LongD group, the modal portion of the distribution was shifted toward longer fixation durations, and the tail of the distribution was somewhat increased, with a larger proportion of the data lying in the tail; this

was more pronounced for the naturalistic videos than for the semi-naturalistic ones. Overall, the computational model simulations were able to capture the group differences in fixation duration distributions for the LongD and ShortD groups reasonably well. However, the CRISP model performed worse on the more irregular LongD data as compared to the ShortD data, as is evident from the larger deviation measures (see "Error" in Table 5-2).

Table 5-2 Best-fitting parameters for simulation study 2

Parameter	Symbol	Range	LongD		ShortD	
			Nat	Non-nat	Nat	Non-nat
Saccade Programming						
Labile stage (ms)	τ_{lab}	200-400	400		322	
Non-labile stage (ms)	τ_{nlab}	30-120	103		40	
Std (ms)	σ	---		0.33		
Saccade timing						
Mean (ms)	t_{sac}	200-650	560	648	437	466
Variance	N	5-20	10	10	20	6
Saccade execution	t_{ex}	---		50		
Error	E	---	14.641	23.684	11.643	12.747
Error cumulative distribution	E_c	---	4.244	8.680	4.950	3.715

Reducing the number of participants by grouping them according to their disengagement abilities revealed some limitations of the CRISP model in simulating infants' fixations. Even though the peak and the tail of the distributions were efficiently captured by CRISP, there were a number of long fixations (> 2 seconds), particularly evident in the LongD group that were not captured by the model. These long fixations have traditionally been associated with disengagement difficulties (see Chapters 1 and 4). The fact that most of these very long fixations were found in the LongD group supports this hypothesis.

Analysis of model parameters (Table 5-2) offers the intriguing possibility of exploring at what level the differences in fixation durations between the two groups primarily originate – at the

level of saccade timing, saccade programming, or both? As outlined before, I consider saccade timing parameters as an indicator for the speed of visual-cognitive processing, whereas saccade programming parameters serve as an index for the development of the visuo-motor system.

The model simulations suggested a higher mean value for the random walk timing signal for the LongD group compared to the ShortD group regardless of the viewing condition. This suggests that infants with LongD are slower at processing the visual information required for triggering a saccade program, contributing to the greater prevalence of long fixations in this group.

On the other hand, Table 5-2 also shows differences in saccade programming parameters between both groups. In the CRISP model, at the transition from a labile to a non-labile stage of saccade programming a "point of no return" is passed. Interestingly, for the ShortD group the labile stage covers 88.95% of the total saccade programming time, while for the LongD group it covers 79.52% only. In relative terms at least, this leaves more time for the modification or cancellation of saccade programs for infants in the ShortD group, which may be indicative of a more developed visuo-motor system. Furthermore, the model simulations suggested that while the non-labile stage program is already as short as for adults in the ShortD group (40 ms), it is still prolonged for the LongD group (103 ms). These results suggest that even though the neural structures involved in saccade programming (e.g., the frontal eye fields, FEF) are thought to be in place at 6 months, at least in some infants they may still not be developed enough to perform adult-like on-line control of eye-movements. Furthermore, both groups show relatively long labile stages, which means that even the ShortD group may still be going through a "calibration phase" that will lead to more efficient adult-like eye-movements later on. The differences between the two groups are striking, providing evidence for large variability in saccadic control across infants from the same age group.

In sum, according to these simulations with the CRISP model, fixation durations at 6 months are influenced by both the speed of visual and cognitive processing, and by developmental aspects of visual and cognitive processing. This second factor is particularly evident for the group that showed higher disengagement difficulties and could be responsible for the programming of particularly long fixations. Furthermore, the ability to disengage from a focal point seems to greatly influence fixation duration distributions and hence the parameters of the model. These

findings are in agreement with past research reporting that looking behaviour within a habituation paradigm was affected by both information processing and disengagement (Domsch et al., 2010).

5.7 Simulation study 3: The individual contributions of saccade timing and saccade programming

In the previous simulation study I compared the empirical results and simulations from infants showing large difficulties disengaging with those who did not. Results suggested that long fixation durations at 6 months are a consequence of both the developmental state of the visuo-motor system and the speed of visual and cognitive processing. In particular, infants with more difficulties disengaging showed increased saccade programming and saccade timing parameters and hence a lower developmental state of the visuo-motor system and slower processing. In this section I present four simulations that aim to explain the individual contributions of both sets of parameters and the relationship between the developmental state of the visuo-motor system and visual and cognitive processing speed in both LongD and ShortD groups.

5.7.1 Behavioural data

The behavioural data was exactly the same as described in simulation study 2, though I only analysed the data from the naturalistic condition.

5.7.2 Model adjustments

To determine the individual contribution of saccade programming and saccade timing model parameters in fixation duration distributions from LongD and ShortD groups, I ran two types of simulations estimating different sets of parameters: I simulated the data for the naturalistic condition for the LongD and ShortD groups by (1) estimating the saccade programming parameters (mean durations of labile and non-labile stages) and maintaining saccade timing parameters fixed and by (2) estimating saccade timing parameters and maintaining saccade programming parameters fixed. Simulations were analysed and only considered valid if they fitted the empirical data well (capturing the mode and the tail of the distribution) and if the estimated parameters were neurophysiologically plausible.

5.7.3 Modelling results

As in previous simulations, parameters were estimated with a genetic algorithm. The values for the fixed parameters were taken from the results obtained in simulation study 1 (see Table 5-1), in which all infants were modelled together. The possible ranges for estimating parameters were wider than in the previous simulation studies in order to also allow physiologically implausible values. The limitations of the fitting procedure, previously encountered in simulation study 2, were more pronounced for these simulations, where only one set of parameters was allowed to vary. As a consequence, it could happen that a parameter set that had the lowest error according to the genetic algorithm[5] did not fit the peak and the tail of the distribution very accurately. In this case, the set of parameters was adjusted manually (Table 5-3).

Table 5-3 Best-fitting parameters for simulation study 3. The grey boxes indicate those simulations that fitted the empirical data and that were physiologically plausible.

Parameter	Symbol	Estimating sacc. programming			Estimating sacc. timing		
		Range	LongD	ShortD	Range	LongD	ShortD
Saccade Programming							
Labile stage (ms)	τlab	200-600	420	328	---	385	
Non-labile stage(ms)	τnlab	30-300	110	42	---	62	
Std (ms)	σ	---	0.33		---	0.33	
Saccade timing							
Mean (ms)	$tsac$	---	458		200-800	590	460
Variance	N	---	20		5-20	15	20
Saccade execution	tex	---	50		---	50	
Error	E	---	20.047	9.806	---	34.03	49.48
Error cumulative distribution	E_c	---	70.469	5.010	---	17.261	94.975

---

[5] Genetic algorithms do not always find the optimal, best-fitting solution.

Figure 5-5 depicts the empirical and simulated fixation duration distributions for the various conditions.
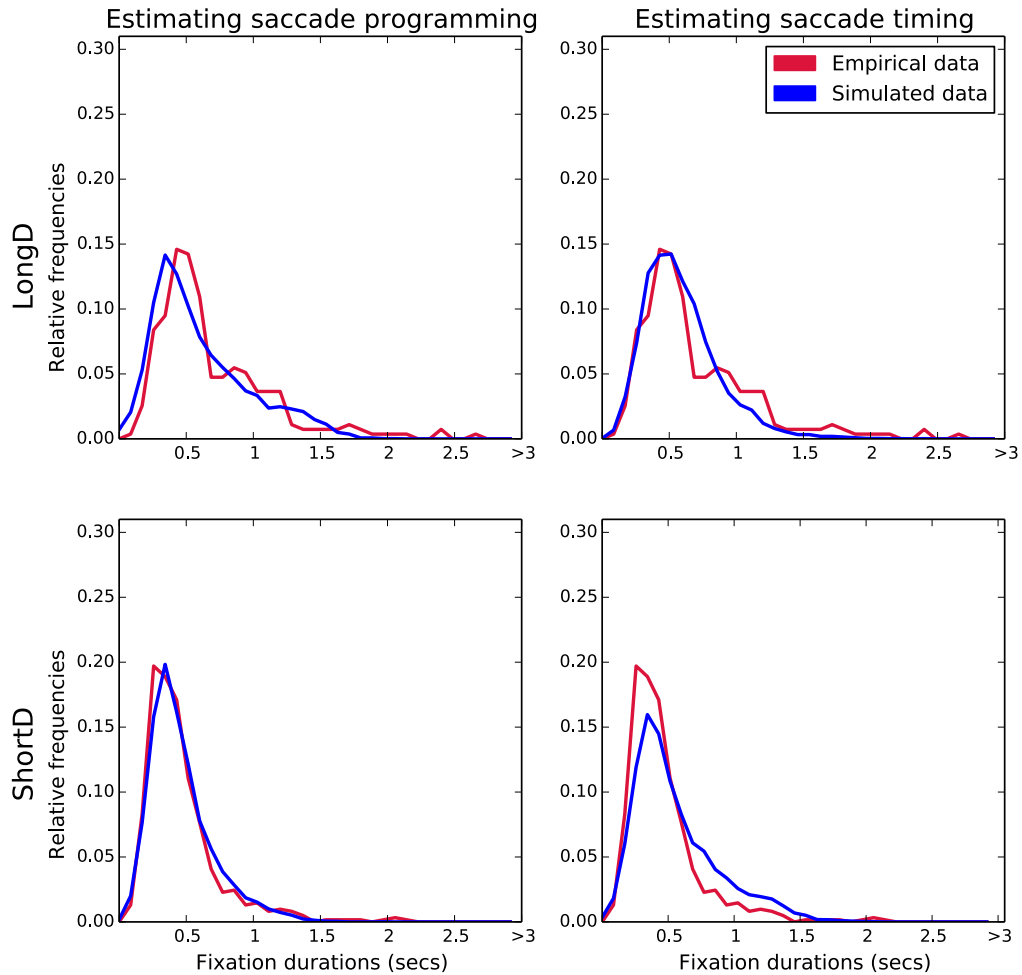
Figure 5-5 Simulation study 3: Left column: empirical data (fixations on naturalistic stimuli only) paired with the simulations after estimating the saccade programming parameters and keeping saccade timing constant. Right column: empirical data paired with the simulations after estimating saccade timing parameters and keeping saccade programming constant.

Inspection by eye suggests that the best model fit was obtained for the ShortD group when allowing the saccade programming parameters but not the saccade timing parameters to differ between the two groups (Figure 5-5, bottom-left). Consequently, allowing only the saccade timing parameters to vary (Figure 5-5, bottom-right) did not achieve a good fit for the ShortD distribution. For the LongD group, the opposite pattern was found: the fit was better when the saccade timing parameters were allowed to vary (Figure 5-5, upper-right) but not the saccade

programming parameters (Figure 5-5, upper-right). These results show that both ShortD and LongD groups need a particular combination of saccade-programming and saccade timing parameters for achieving a good fit. This suggests that the developmental state of the visuo-motor system (i.e. saccade programming) and visual and cognitive processing speed (i.e. saccade timing) are linked together, such that infants with a less developed visuo-motor system could also be slower processors.

Comparing the parameters for ShortD and LongD valid simulations (grey-shaded cells in Table 5-3) it becomes apparent that the largest differences were seen in values for saccade timing parameters (random walk parameters), while values for saccade programming parameters (labile and non-labile stages), though still larger for the LongD group, did not suffer a radical change. This suggests that even though developmental aspects of visual processing still affect fixation durations at 6 months (as seen in the increased saccade programming parameters from the LongD group), it is visual and cognitive processing speed that generates the largest differences between the two groups.

5.8 Discussion

Even though the results from the empirical study presented in Chapter 4 already highlighted the influence of both developmental factors and cognitive processing in fixation durations, it did not explain which mechanisms are responsible for the variations in saccadic control in infancy, and questions such as the extent to which the variations in fixation duration distributions in infancy are due to developmental aspects of the visuo-motor system and/or visual and cognitive processing speed. Investigating these issues using more traditional psychophysical experiments can be challenging if not unworkable due to the practical and current technological limitations that testing infants entails (see Chapter 2). On the other hand, computational modelling allows us to describe, predict and explain data that is itself unobservable (Lewandowsky & Farrell, 2011).

The goal of this chapter was to investigate the unexplored mechanisms underlying saccadic control and fixation durations in infancy by determining the generalizability and potential limitations of extending the CRISP theoretical framework and computational model of fixation durations in adult scene-viewing (Nuthmann et al., 2010) to infants. The CRISP model was used to investigate the underlying mechanisms modulating fixation durations in 6-month-olds and test

specific developmental theories of oculomotor control by applying the model to the empirical fixation-duration data described in Chapter 4.

Simulation study 1 confirmed that the CRISP model and theoretical framework is able to capture the trends observed in empirical fixation-duration data from 6 month-old infants. This also suggests that at 6 months saccade programming occurs in two stages (the labile and non-labile stages), which constitutes the first evidence for the two-step notion of saccade programming in infants. Additionally, this study explored the influence of the viewing condition on mean fixation durations during scene viewing in infants and adults. Age differences were examined by exploring differences in model parameters for the two age groups. CRISP fit both the infant and adult data well (Figure 5-3), even though the very long fixations found in infants produced additional difficulties when modelling the data. Comparisons between the parameters estimated for different age groups suggested that both saccade timing and saccade programming (labile and non-labile stages) are considerably slowed down at 6 months compared to adults. This suggests that at 6 months fixation duration may still be affected by the development of the visuo-motor system as well as by the efficiency of visual and cognitive processing. Moreover, as in previous CRISP simulations (Nuthmann et al., 2010), task differences for both age groups were captured by adjusting how fast the random walk timer operates.

Simulation study 2 aimed to apply the CRISP model to investigate specific developmental theories. In particular, this study aimed to explore the differences in fixation durations between infants with large (LongD group) and short (ShortD group) disengagement latencies. A specific goal was to test whether the differences in fixation durations were primarily due to an underdevelopment in the visuo-motor system, the speed of visual and cognitive processing, or both. I assumed that the group of infants with longer disengagement latencies would be the one presenting a less developed visuo-motor system. Based on previous research on the inhibition of saccades in infants (e.g., M. H. Johnson, 1995), I reasoned that infants with larger disengagement difficulties may also present longer saccade latencies, with both labile and non-labile phases being prolonged. Results from this study evidenced differences in saccadic control between the LongD and ShortD groups. The model simulations revealed differences between the ShortD and LongD groups in both saccade timing as well as saccade programming parameters. The LongD group presented considerably larger saccade timing parameters than the ShortD group, evidencing the influence of visual and cognitive processing on fixation

durations. At the same time, differences in saccade programming –in both labile and non-labile stages- highlighted the effect of the developmental state of the visuo-motor system on infant oculomotor control. Interestingly, while the duration of the non-labile stage was already at an adult level in the ShortD group (40 ms), it was still considerably prolonged for the LongD group (103 ms). These results indicate that even though the neural structures involved in eye-movement control (e.g., the frontal eye fields, FEF) are supposed to be in place by 6 months, they may still be going through a "calibration phase" that will lead to more efficient adult-like eye-movement behaviour later on (Csibra et al., 1998). In sum, findings from this simulation suggested that fixation durations at 6 months may be both affected by the speed of visual and cognitive processing and by developmental aspects of the visuo-motor system, with the latter being particularly prominent in infants with disengagement difficulties. These results are in line with findings from Domsch and colleagues (2010), reporting that looking behaviour within a habituation paradigm was affected by both information processing and disengagement.

Simulation study 2 also highlighted the limitations of the CRISP model when simulating fixation durations from infants. In adults, very long fixation durations (> 2 secs) are fairly uncommon and tend to be discarded from the analysis, the rationale being that such long fixations may not be indicative of any perceptual activity or information processing (Inhoff & Radach, 1998). Young infants, however, can present a number of very long fixations that are mainly assumed to be a consequence of disengagement difficulties ("sticky fixation"; Atkinson, 2000; Farroni et al., 1999; Frick et al., 1999; M. H. Johnson, 2011). For this reason I decided not to discard long fixations from data analysis for any of the age groups. The model simulations were able to capture the trends in fixation duration distributions as expressed in the mode and the slope of the distribution. Perhaps not suprisingly, they did not capture the very long fixations present in the tail of the distribution. The fact that this problem was more evident in the LongD group may suggest that these fixations were a consequence of poor disengagement abilities. To accommodate this, future work should extend the model by including a "disengagement component" able to stocastically prolong certain fixations and simulate the long tails found in infants with disengagement difficulties. This, however, raises the question of whether disengagement should prolong the labile or the non-labile stage of saccade programming. Disengagement difficulties are though to be a consequence of problems triggering an eye-movement after a fixation. According to this, the disengagement component should rather

influence the non-labile stage of saccade programming, since it is at the end of this stage when a saccade is executed. Nevertheless, as our results from Chapter 4 have shown, disengagement abilities are stimulus dependent and can also affect adult populations, which suggests that the labile-stage of saccade programming should also be affected.

Simulation study 3 presented four simulations in order to explore the individual contributions of both sets of parameters (saccade programming and saccade timing) and the relationship between the developmental state of the visuo-motor system and visual and cognitive processing speed in both LongD and ShortD groups. Results suggest that the main factor affecting infant saccadic control at this age is visual and cognitive processing. The differences in speed of processing between groups of infants that differ in their ability to dissengage from a visual target could indicate a relationship between the developmental state of the visuo-motor system and information processing, so that infants with a less developed visuo-motor system would also be the slower processors. While the model simulations allowed for an initial exploration of these theoretical possibilities, it is clear that more empirical and computational research is needed to clarify these issues.

Even though the present studies helped to shed light on the mechanisms and processes underlying eye-movements in 6-month-old infants, the causes that explain the large individual differences between short and long lookers are still unknown. For instance, from a skill learning perspective, it could be argued that infants with shorter fixation durations will generate a larger number of saccades per day, and as a consequence, they will become more skilled in this aspect of motor control compared to other infants. In order to test this hypothesis, future simulations may consider factoring a "learning component" into the model.

The present chapter was also concerned with differences in overall fixation duration distributions in infants as opposed to adults, and with the global effects of the viewing condition. A related issue concerns the degree to which fixation durations are under the direct moment-to-moment control of the current visual stimulus. Previous simulations with the CRISP model have investigated this issue by applying the scene onset delay paradigm, which selectively manipulates global scene processing difficulty (see Chapter 1). The empirical data and model simulations showed that fixation durations are, at least partly, influenced by the current visual stimulus (Nuthmann et al., 2010). Such an exhaustive investigation has still not been possible in

infants due to a number of constrains (e.g., data quality, high degree of movement, the use of systems with lower sampling frequency). Nevertheless, in the future the CRISP model could be applied to investigate the influence of endogenous and exogenous factors in eye-movement control in infancy with the expectation that direct control (i.e. endogenous factors) may increase over the first year of life, mirroring frontal cortex development.

5.9 Conclusions

In summary, the present chapter demonstrated the generalizability of the CRISP theoretical framework to data from 6-month-old infants. Perhaps one of the most critical contributions of this framework to the infant literature is the postulation of the two-step notion of saccade programming, where saccades are programmed in an initial labile stage that is subject to cancellation, and a later, non-labile stage. Additionally, simulation studies 2 and 3 were used to test the extent to which the developmental state of the visuo-motor system and visual and cognitive processing affect fixation durations at 6 months, with results suggesting that while both factors influence oculomotor control at this age it is visual and cognitive processing which causes the biggest effect.

In the present and the previous chapter I have investigated the stability in fixation durations within a single testing session. Whilst cross-sectional studies can tell us about the developmental state of in this case the oculomotor system in a particular moment, they are not the best tool to explain developmental change.

In the next two chapters I will present longitudinal data in order to explore the trajectories of the same infants at different points in time during the first year of life and test some of the hypothesis made in this chapter that were not possible to test in a cross-sectional study. How do fixation durations develop over the first year of life? How do saccade processing and saccade timing parameters develop over the first year of life?

Chapter 6: The development of fixation durations in dynamic and static complex scenes: longitudinal studies

6.1 Introduction

In the previous two chapters I investigated the factors modulating saccadic control and fixation durations in 6-month-old infants and adults when viewing complex dynamic stimuli and performing simple oculomotor tasks such as the gap-overlap paradigm (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991) or the double-step paradigm (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954). Results from the empirical study described in Chapter 4 revealed differences in mean fixation durations across age groups (6-month-olds vs. adults) and viewing conditions (naturalistic vs. semi-naturalistic videos), and a relationship—particularly evident in the group of infants—between the ability to disengage from a focal point and mean fixation durations. As shown in the differences across viewing conditions, these findings suggest that infant fixation durations reflect on-line perceptual and cognitive activity in a similar way to adults, but that the individual developmental state of the visuo-motor system still affects this relationship at 6 months, as implied by the high correlations between disengagement latencies and mean fixation durations in infants. Moreover, results from the double-step paradigm suggest that by 6 months of age infants are able to reliably cancel a saccade program. Additionally, in Chapter 5 I investigated the unexplored mechanisms underlying saccadic control and fixation durations in 6-month-olds by determining the generalizability and potential limitations of extending the CRISP theoretical framework and computational model of fixation durations in adult scene-viewing (Nuthmann et al., 2010) to infants. Perhaps one of the most critical contributions of this framework to the infant literature was the two-step notion of saccade programming, where saccades are programmed in an initial labile stage that is subject to cancellation, and a later, non-labile stage. In this chapter I also investigated the extent to which the developmental state of the visuo-motor system and visual and cognitive processing affect fixation durations at 6 months, with results suggesting that while both factors influence oculomotor control at this age it is visual and cognitive processing what causes the biggest effect. The studies described in Chapters 4 and 5 investigated the stability in fixation durations within a single testing session. Whilst cross-sectional studies can tell us about

the developmental state of in this case the oculomotor system in a particular moment, they are not the best tool to explain developmental change.

In the present chapter I present two longitudinal studies and a study with adult participants that aim to investigate the development of fixation durations and saccadic control over the first year of life when viewing a battery of dynamic and static stimuli and performing the gap-overlap paradigm. This design will also allow exploring individual differences across different viewing conditions and visits.

The data and results from these studies will constitute the basis for exploring the development of the mechanisms underlying saccadic control during the first year in the next chapter, where data will be modelled using the CRISP theoretical framework and computational model of fixation durations in adult scene-viewing (Nuthmann et al., 2010).

6.2 The developmental change in looking times, fixation durations, and reaction times

Even though the developmental course of look duration is non linear across infancy and toddlerhood (Colombo, 2001; Ruff & Rothbart, 1996), many studies have reported a decrease in look duration during the first year of life (Colombo & Mitchell, 1990; Kagan et al., 1971; M. Lewis et al., 1969; Mayes & Kessen, 1989; Shaddy & Colombo, 2004). Interestingly, some of these studies also reported an increase in looking at around 13 months (Kagan et al., 1971; M. Lewis et al., 1969), perhaps due to the emergence of executive functions. The decrease in looking times during the first year could be the result of an improvement in processing efficiency (Colombo et al., 1991), the disengagement of attention (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), and/or advances in memory capacity – in either encoding, storage, retrieval, or some combination of these processes (Colombo, 1993). On the other hand, some have also reported different trends in look duration trajectories for different stimulus types (Courage et al., 2006; Frank, Amso, & Johnson, 2014; Frank et al., 2009; Shaddy & Colombo, 2004). For instance, Courage, Reynolds, and Richards (2006) investigated the development of look duration as a function of age and stimulus type in 14- to 52-week-old infants that were presented with a battery of static and dynamic versions of faces, Sesame Street material and achromatic patterns. They found that while look duration decreased significantly from 14 to 26 weeks for all types of stimuli, from 26 weeks on the course of look

duration diverged according to the stimulus type. When the older infants viewed the achromatic static patterns, their look duration continued to decline and/or it reached a point where it did not decline nor increase. On the contrary, look duration significantly increased for the Sesame Street material and the faces, in a way that their performance from 14 to 52 weeks was better described by a quadratic function. Similar trends have been found when investigating infants looking at faces (Frank et al., 2014, 2009) or looking times during infant playing with toys. Ruff, Saltarelli, Capozzoli, and Dubiner (1992) found no change in looking over time when 5 to 11 months infants were let to play with single objects. Nevertheless, some other studies have reported an increase in looking times over the same age range when several objects were given to the infant simultaneously (Bakeman & Adamson, 1984; Ruff & Saltarelli, 1993). These opposing results where looking times decrease or increase according to the stimulus type may suggest that while some influences on attention decline, others increase, probably as a result of the development of new cognitive and motor skills (Ruff & Rothbart, 1996).

As described in Chapter 1, while few studies have specifically investigated fixation duration trajectories in infancy, some have used fixation duration measures to investigate the development of visual scanning of static and dynamic stimuli. For instance, Bronson (1994) analysed the visual scanning patterns of 6–13 week old infants when they were presented with static geometric figures. Besides the gradually decreasing fixation durations, he found that infants' scanning is initially salience-driven, and that it is progressively replaced by volitional control over the choice of saccadic targets. In another study (Bronson, 1990), he studied the scanning patterns of 2–14 week-old infants when presented with simple dynamic and static shapes. Even though as infants grew older they were increasingly more likely to fixate the different features of the static shapes, they still had difficulties disengaging their attention when the stimulus was flickering.

Similar findings have been reported in a number of more recent studies using stimuli that are more ecologically valid than static pictures or drawings. Note that the use of stimuli that are of limited ecological validity has been often addressed and criticized since the generalizability of the experimental results is uncertain (Hunnius & Geuze, 2004a; Hunter & Richards, 2011; Neisser, 1976; Schmuckler, 2001). Hunnius and Geuze (2004a) analysed in a longitudinal study the scanning patterns of 10 infants between the ages of 6 and 26 weeks that were presented with a video of their mother's face and an abstract stimulus. Their results show a clear

developmental progression with scanning patterns stabilizing at around 18 weeks, which is slightly later than what is usually reported for the scanning of static images. From 14 weeks on they also found stable differences between the median fixation durations of the two stimuli. Interestingly, the fact that they used more ecologically valid and visually complex stimuli made a difference in the results: infants spent more time looking at the mouth than with less realistic stimuli (e.g., the drawing of a face), and much less time looking at the edge of the face. Note that studies using reverse correlation methods have identified the corners of the mouth as a primary locus of emotional content in faces (Kontsevich & Tyler, 2004).

Few studies have also investigated the developmental trajectories of reaction times as an indicator of information processing. For instance, Canfield, Wilken, Schmerl, and Smith (1995) used the visual expectation paradigm in order to study continuity in reaction times, anticipation, and off-task behaviour between the ages of 4 and 6 months. Besides reporting strong stability in reaction times, their results showed a decrease from 4 to 6 months. In another study, Butcher, Kalverboer, and Geuze (2000) tested the same infants from 6 to 26 weeks in their ability to shift their gaze from a central to a peripheral stimulus during competition and non-competition trials. They found that the latency on non-competition trials decreased from 8 to 16 weeks, while for competition trials latencies continued decreasing until 26 weeks.

6.3 Individual differences in looking times, fixation durations and reaction times

Many studies have reported individual differences in looking times and have speculated about the differences between short and long lookers (Colombo et al., 1995, 1991; Colombo & Mitchell, 1990; Colombo, 1995; Cuevas & Bell, 2013; Jankowski & Rose, 1997). Some have hypothesized that shorter looking times reflect more rapid encoding of visual information within and across ages (Colombo et al., 1987; Colombo, 1995), with a number of studies suggesting how short lookers are more efficient processors compared to long lookers (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997) or present different strategies of visual intake (Colombo et al., 1995, 1991).

Individual differences have also been the focus of some of the work investigating reaction times (Butcher et al., 2000; Canfield et al., 1995) and fixation durations in infants. For instance, Bronson (1994) tested 6- to 13-week-old infants and found that decreases in visual

attentiveness were associated with longer fixation durations while looking at static figures. In another more recent study, Wass and Smith (2014) presented 11-month-old infants with a battery of complex dynamic stimuli and non-complex static stimuli. They reported that individual differences in fixation durations were stable across different stimulus types and testing sessions (but note that all testing sessions happened within 15 days). Furthermore, reaction times gathered from the gap-overlap paradigm correlated with dynamic stimuli, but not with static stimuli for which measures of cognitive control and arousal were most predictive of performance.

Look duration has been associated with later intellectual function and cognitive performance in childhood and adolescence (see Chapters 1 and 4; Choi & Vaswani, 2014; Colombo et al., 1995; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997; Kavšek, 2004; Lawson & Ruff, 2004; Sigman, Cohen, Beckwith, & Parmelee, 1986; Slater, 1995). Likewise, recently Papageorgiou and colleagues (2014) investigated the relationship between fixation durations and later intellectual function in childhood. They demonstrated that individual differences in fixation durations in early infancy can predict individual differences in temperament and behaviour in childhood

 Fagan III (1984) and Bornstein and Sigman (1986) argued that the stability in look duration is mediated by stable individual differences in speed or efficiency of information processing. This means that infants who habituate more quickly, have shorter looking times, or look longer at novel stimuli are demonstrating faster processing efficiency compared with other infants, which means that more information can be processed in a given unit of time. This gives them an advantage over their peers to perform better in ability tests during childhood (Canfield & Ceci, 1992).

A second hypothesis states that the stability in look duration is mediated by stable individual differences in memory capacity, in either encoding, storage, retrieval or a combination of these processes (Colombo, 1993). According to this hypothesis, infants that show better memory capacity will have quicker habituations and will show higher levels of recognition memory for the familiar stimulus, thus showing a preference for the novel one. To date, both theories seem to be feasible and not mutually exclusive, as proposed by studies investigating continuity in infancy and early childhood (Colombo, 1993; Jacobson et al., 1992).

6.4 Ex-Gaussian modelling

Generally, adult fixation durations when viewing pictures form a positively skewed distribution with a mode at 230 ms, a mean of 330 ms, and fixation durations that go from 50 ms to 1000 ms approximately (Henderson & Hollingworth, 1999). Infant fixation duration distributions have also been reported to be highly skewed and definitely not normally distributed (Harris et al., 1988). For this reason, means may not necessarily be representative of fixation duration distributions.

With the view to solve this issue some studies have reported means and standard deviations, or median fixation durations (Hunnius, Geuze, & van Geert, 2006; Hunnius & Geuze, 2004a, 2004b). More recent studies from reading research (Glaholt, Rayner, & Reingold, n.d.; Reingold & Rayner, 2006; Staub & Benatar, 2013) have used the ex-Gaussian function to fit fixation data. Interestingly, it has been reported that the ex-Gaussian function provides a very good fit to several empirical reaction time distributions (Hockley, 1984; Ratcliff & Murdock, 1976).

The increasing popularity of the ex-Gaussian function lies in the fact that it is theoretically justified and that it provides parameter values that are easy to interpret (Lacouture & Cousineau, 2008). The ex-Gaussian function is the convolution of two additive processes, a Gaussian (normal) function and an exponential function (Lacouture & Cousineau, 2008). It has three parameters, μ, σ and τ. The first two (μ, σ) correspond to the mean and standard deviation of the Gaussian component of the distribution, and can be interpreted as localization and variability indicators respectively. On the other hand the third parameter (τ) is the mean of the exponential component, and corresponds to the right "tail" of the distribution, with larger τ-s representing more skewed distributions. A more skewed fixation duration distribution shows more variability in the data. In this chapter distributions of fixation durations will be fitted using the ex-Gaussian distribution.

6.5 Current studies

In the present chapter I present two longitudinal studies and a study with adults. For the first longitudinal study, all infants completed three visits that were scheduled every 6 weeks at 3.5, 5 and 6.5 months calculated from the due date. For the second longitudinal study, infants were tested every 8 weeks at approximately 6, 8, 10 and 12 months.

All participants from all studies were eye-tracked while they were presented with two different complex dynamic viewing conditions (naturalistic and abstract videos), a static condition (complex static images) and performed the gap-overlap paradigm (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991). The free viewing tasks were used to analyse the micro-dynamics of visual and cognitive processing during spontaneous orienting by measuring fixation durations. The gap-overlap task was used to measure participants' disengagement abilities.

In particular, the studies described in this chapter aim to

a) describe fixation duration trajectories for different types of stimuli during the first and the second half of the first year;

b) analyse the differences in the processing of dynamic and static stimuli during the first year;

c) explore individual differences in fixation durations across visits and stimulus types;

d) explore the relationship between fixation durations and disengagement latencies during the first year;

e) analyse the differences in fixation durations between infants and adults when viewing static and dynamic complex stimuli.

Results for the three studies will be discussed at the end of this chapter.

6.6 Longitudinal study 1: Three to 6 month-olds

6.6.1 Methods

6.6.1.1 Participants

In total, 12 typically developing infants (7 girls, 5 boys) participated in the longitudinal study. An additional 2 infants were excluded from the analysis due to their low quality data in at least two of the visits.

All infants scored within their age range on the Bayley Scales of Infant Development (Bayley, 1993). All infants completed three visits that were scheduled every 6 weeks at 3.5, 5 and 6.5 months calculated from their due date. Mean ages were 101.46 days (SD = 4.46), 144 days (SD = 4.39), and 186.31 (SD = 4.85). Visits were scheduled at times of the day when the caregiver

considered the infant was going to be most alert. In case the measurement session was not successful due to the infant being unhappy or other technical issues, another visit was scheduled as promptly as possible during the next 7 days.

Most infants (N = 10) were of white middle socioeconomic status. The infants were recruited via magazine advertisements, social networking media, and flyers. Families were given baby t-shirts or bags as gifts for their participation. The study protocol was approved by the Birkbeck, Psychological Sciences Ethics Committee. Table 6-1 shows the age of each participant at the times of visit.

Table 6-1 Information about the visits for the longitudinal study from 3 to 6 months

Subject	Gender	Age Visit 1 (Days)	Age Visit 2 (Days)	Age Visit 3 (Days)
1	Girl	109	153	192
2	Girl	98	140	179
3	Girl	105	147	188
4	Boy	104	146	188
5	Girl	101	143	190
6	Boy	107	146	190
7	Girl	94	140	182
8	Girl	97	139	184
9	Boy	98	141	181
10	Boy	98	140	183
11	Girl	101	143	184
12	Boy	101	143	185

6.6.1.2 Apparatus

The apparatus was identical to the one described in Chapter 4.

6.6.1.3 Stimuli

The spontaneous looking task included (1) 15 customized naturalistic videos in which three people performed baby-friendly actions, (2) 15 abstract videos created from the first set of naturalistic videos, and (3) 15 static complex images. The three types of stimuli were presented in colour and had the same size. For a precise description of each of the stimulus used for this study go to Appendix C.

6.6.1.3.1 Naturalistic videos

Fifteen naturalistic videos were created following the same procedure described for the cross-sectional study in Chapter 4. In fact, the present battery includes 9 of the videos described in Chapter 4 and 6 new ones. One of the previous naturalistic videos was excluded from this design because it was eliciting smooth pursuit eye-movements. Figure 6-1 (left side) shows a frame from one randomly selected naturalistic video.

Figure 6-1 Frames from a naturalistic video and the abstract video that was constructed from it.

6.6.1.3.1 Abstract videos

The abstract videos were derived from the naturalistic videos by applying a number of distortion filters from Final Cut Pro X Version 10.0.5. The resulting videos presented the same dynamics and maintained equal low-level visual features, such as colour or luminance, as the naturalistic videos. On the other hand, they lacked the semantic content and the familiarity that the naturalistic videos contain. Figure 6-1 shows a frame from one randomly selected abstract video paired with the naturalistic video from which it was constructed.

6.6.1.3.2 Static complex images

Fifteen static images were presented for 7 seconds each. The images not only presented high complexity in terms of their semantic and social content but also in terms of low-level visual features such as colour, contrast, and luminance. In fact, all the images contained bright and high contrast colours and displayed different scenarios including one or several characters. Figure 6-2 shows a randomly selected static image.

Figure 6-2 One of the static images used in the experiment.

6.6.1.4 Gaze-contingent paradigms

The current experiment includes the gap-overlap gaze-contingent task. The contingencies were created by using the T2T (Talk to Tobii) functions (Shukla et al., 2011). The design for this task was the same described in Chapter 4.

For the present study, a total number of 72 trials were presented. Forty per cent of these trials were overlap trials, 30% gap trials and 30% baseline trials. They were presented in groups of 12 trials (one iteration) and alternated with other tasks of the study. The peripheral stimulus was always the same (a cloud), and the central stimulus and the background colour changed after every 12 trials.

Disengagement latencies were calculated by subtracting mean baseline latencies to mean overlap latencies. For the analysis presented in this thesis the gap trials were excluded.

6.6.1.5 Design and Procedure

Infants and caregivers were welcomed in a lab waiting room where the infants acclimated to the experimenter and the lab. After this, they were tested individually in a small-darkened room while sitting on a baby-chair located 60cm away from the monitor or on their mother's lap. Prior

to starting the experiment the infants were calibrated with a 5 points calibration procedure (see Chapter 2). Once the eye-tracking study was completed, the experimenter evaluated the infant's gross-motor development using the Bayley Scales of Infant Development (Bayley, 1993).

The different tasks of the study were alternated: (1) thirteen gap overlap trials, (2) one abstract video presentation (20-25 seconds), (3) one static image presentation (7 seconds), (4) and one naturalistic video presentation (20-25 seconds). This process was repeated 6 times. Next, an additional 9 iterations of (1) abstract videos, (2) static images, and (3) naturalistic videos were presented. The order for the videos and the static images was always the same. In total, the infants viewed 15 naturalistic videos, 15 abstract videos, and 15 static images, and performed a maximum of 78 gap-overlap trials. In the case the infant became upset during the procedure the study was stopped. This procedure was the same for all infants and visits (within subjects design).
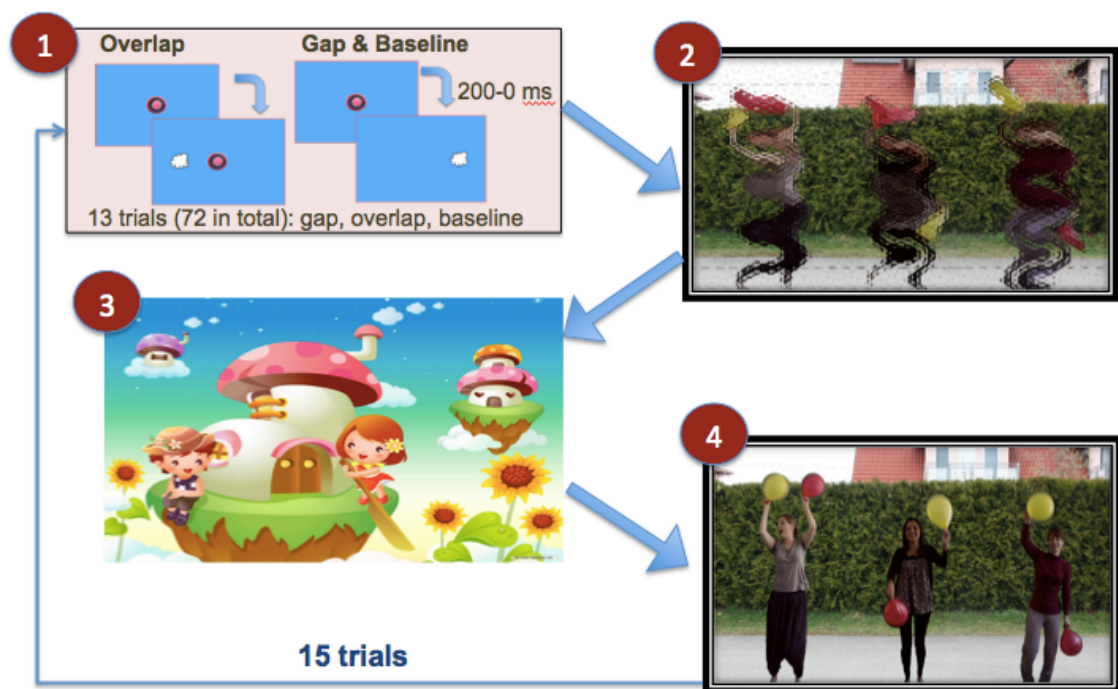
Figure 6-3 Experimental protocol.

6.6.2 Analysis

6.6.2.1 Fixation detection and coding

The fixation detection procedure was exactly the same as the one described in Chapter 4.

6.6.2.2 Data quality analysis

As in Chapter 4, spatial precision was measured with the root mean square (RMS) of inter-sample distances. The RMS was calculated per participant and visit. The mean RMS-s were 0.34° (SD = .16) for the first visit, 0.25° (SD = .09) for the second visit, and 0.17° (SD = .06) for the last visit.

A one-way ANOVA was used to test for the age effect in data quality. The degrees of freedom of the F test were corrected according to the Greenhouse-Geisser method. A significant effect was found showing that RMS decreased with age, $F (1.255, 13.809) = 12.454$, $p = .002$.

Additionally, correlational analysis between RMS-s and fixation durations were run for each age point in order to ensure data quality was not interfering with the fixation detection method. No correlations were found between data quality and spatial precision for any of the visits (see Table 6-2).

Table 6-2 Correlations between fixation durations and RMS in the three visits.

		Visit 1	Visit 2	Visit 3
Fixation durations vs. RMS	Pearson	-.230	-.388	-.468
	Sig. (2-tailed)	.472	.212	.125
	N	12	12	12

6.6.2.3 Cross-validation

6.6.2.3.1 Fixation durations

An external coder that was naive to expected outcomes was trained to code fixations from infants featuring low and high quality data. The coder had to (1) run GraFIX automatic detection algorithms using the parameters described in Chapter 3 (see Table 3-3), and then (2) manipulate the resulting outcome in order to remove artifactual fixations or add those undetected by following the predefined guidelines described in Chapter 4. In total, the external coder coded the 20 % of the data from the infants that participated in this study.

The inter-rater reliability between the means and the number of detected fixations was evaluated using the ICC (Hallgren, 2012). A strong agreement between the mean fixation

durations was found (with an ICC of .831, *p* = .007). Additionally, there was also an agreement in the number of fixations detected (with an ICC of .774, *p* = .004).

6.6.2.3.2 Gap-overlap

All the trials from the gap-overlap paradigm were reviewed in order to manually exclude those where (1) the accuracy was not good enough to trigger the gaze-contingencies, or (2) the infants looked away during the presentation of the peripheral stimulus. An external coder that was naïve to the expected outcomes evaluated the validity of the gap-overlap trials for 20% of the infants that participated in this study.

Interrater reliability was measured using the Kappa statistic with a view to determine consistency among raters. This analysis showed a strong agreement between the two coders, Kappa = .809 (*p* <.001), 95% CI (0.727, 0.891).

6.6.2.4 Statistical analysis

The mean fixation duration and the ex-Gaussian components for each participant and viewing condition were calculated (see Lacouture & Cousineau, 2008) after excluding all fixations with a duration that was two standard deviations above or below the initial mean. Hence, the fixations that were included in the analysis were those that accounted for about the 95 % of the set for each participant. This exclusion was performed in order to exclude very long or very short fixations on a participant basis. An alternative method is to establish maximum and minimum fixation thresholds for different age groups. In this case, due to the exploratory nature of this study in infants, these thresholds would have to be selected arbitrarily. For this reason I opted for the method that would preserve most of the fixations from each participant in the analysis.

In the gap-overlap paradigm only the first 40 trials were included in the analysis (10-15 trials per condition). This was to avoid the trials where infants could be predicting the location of the next target. Trials with latencies shorter than 200 ms were also excluded from the analysis, as in infant studies such fast eye-movements are considered to be eye-tracking errors or anticipatory saccades begun prior to stimulus onset (Canfield et al., 1995; Frick et al., 1999; S. A. Rose, Feldman, & Jankowski, 2002). Some studies also define a maximum latency limit (Elsabbagh et al., 2009; M. H. Johnson et al., 1991; Matsuzawa & Shimojo, 1997). However, since this study explores the variations in fixation durations in longitudinal data, one can expect to find visit and inter-individual differences in their disengagement abilities and reaction times. For this reason I

decided to establish a conservative criterion on a participant basis and exclude all the trials that were one standard deviation above the participant's mean latency (same as in Chapter 4). Disengagement latencies were calculated by subtracting the baseline mean latencies to the overlap mean latencies.

For the correlational analysis, a logarithmic transformation was applied to all the measures for fixation durations. The justification for this transformation has rested on the fact that fixation duration distributions are substantially skewed and the sample size is rather moderate. Thus, the logarithmic transformation was used to induce symmetry on the data.

6.6.2.4.1 The analysis of longitudinal variations in fixation durations

The longitudinal variations in fixation durations were analysed using repeated measures analyses of variances (ANOVAs). This analysis was possible considering that there were no missing points for any of the infants and visits.

This procedure was used in order to examine whether (1) fixation durations changed with age, (2) the viewing condition influenced fixation duration trajectories, and (3) the viewing conditions influenced fixation durations at each visit.

Three (visit) x 3 (stimulus) mixed ANOVAs were used to explore the overall age variations in fixation durations. Further, independent one-way ANOVAs were used in order to investigate the differences between different stimulus types and their independent longitudinal trajectories. In the cases where the correlation matrices between the different measurement points were heterogeneous the degrees of freedom of the F test were corrected according to the Greenhouse-Geisser method. Post-hoc pairwise comparisons were calculated based on Bonferroni correction.

6.6.3 Results

6.6.3.1 The development of fixation durations

6.6.3.1.1 Age effect: Fixation duration trajectories for different viewing conditions

A mixed 3 x 3 ANOVA revealed a significant main age effect in mean fixation durations ($F(2,22)= 14.911$, $p < .001$, $p\eta^2 = .575$) and a stimulus type effect ($F(1.169, 12.862)= 78.939$, $p < .001$, $p\eta^2 = .878$). Pairwise comparisons revealed significant differences in mean fixation durations between the first ($M = 592$ ms, $SD = .017$) and the second visits (M = 539 ms, SD =

.025, *p* = .012) but not between the second and the third visits (M = 523 ms, SD = .020, *p* = .710). Furthermore, there was an interaction effect between age and stimulus type (*F*(4,44)= 10.738, *p* < .001, *pη²*= .494). For a graphical visualization of the data see Figure 6-4.
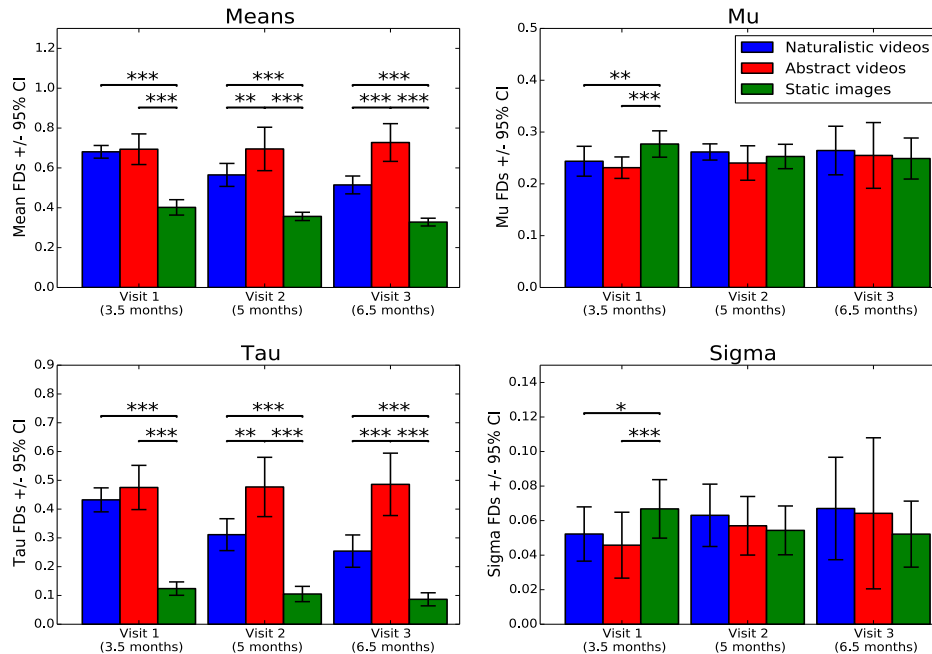
Figure 6-4 Longitudinal study 1: Fixation durations histograms for the means and the ex-Gaussian components.

The interaction between age and stimulus type was examined running three additional ANOVAs that examined the age effect on each stimulus type independently. Mean fixation durations decreased with age in naturalistic videos (*F*(2,22)=33.047, *p* < .001, *pη²*= .750) and static images (*F*(1.448,15.924)=21.468, *p* < .001, *pη²*= .661), but not in abstract videos. Post hoc t-tests revealed significant differences between the three visits for the naturalistic videos (visit one: *M* = 681 ms, *SD* = .050; visit two: *M* = 565 ms, *SD* = .091; visit three: *M* = 515 ms, *SD* = .070; visit one – visit two: *p* = .002; visit two – visit three: *p* = .068) and the static images (visit one: *M* = 402 ms, *SD* = .061; visit two: *M* = 357 ms, *SD* = .033; visit three: *M* = 328 ms, *SD* = .031; visit one – visit two: *p* = .022; visit two – visit three: *p* = .008).

In order to investigate and explain the origin of the effects described above further analyses were performed using the ex-Gaussian components (μ, σ, and τ) for the fixation duration distributions instead of the means. A mixed 3 x 3 ANOVA revealed a significant main stimulus

type effect in μ fixation durations ($F_{(2,22)}$=4.514, p = .023, $p\eta^2$= .291), but not significant age effect was found. Further, there was a significant interaction between age and stimulus type ($F_{(4,44)}$=4.465, p = .004, $p\eta^2$= .289). Three additional ANOVAs investigated this interaction independently revealing that μ fixation durations decreased with age in static images ($F_{(2,22)}$=4.671, p = .020, $p\eta^2$= .298), but not in naturalistic or abstract videos where no significant effects were found. Pairwise comparisons showed significant differences in μ fixation durations in static images between visit one (M = 278 ms, SD =.044) and visit three (M = 235 ms, SD =.042, p = .049).

A mixed 3 x 3 ANOVA revealed an interaction between age and stimulus type for σ fixation durations ($F_{(4,44)}$=3.595, p = .013, $p\eta^2$= .246), even though no main age or stimulus type effects were found. Three additional ANOVAs investigated this interaction independently revealing a marginally significant effect suggesting that σ fixation durations decreased with age in static images ($F_{(2,22)}$=3.173, p = .062, $p\eta^2$= .224), but not in naturalistic or abstract videos where no significant effects were found. However, post hoc pairwise comparisons showed no significant differences between σ fixation durations between the different visits for any of the viewing conditions.

The results for τ analysis were very similar to the ones obtained from the means. A mixed 3 x 3 ANOVA revealed a significant main age effect in τ fixation durations ($F_{(2,22)}$=7.166, p = .004, $p\eta^2$= .394) and a main effect of stimulus type ($F_{(1.175, 12.930)}$=75.377, p < .001, $p\eta^2$= .873). Pairwise comparisons revealed significant differences in τ fixation durations between the first (M = 343 ms, SD = .017) and the second (M = 289 ms, SD = .025, p = .012) visits but not between the second and the third (M = 287 ms, SD = .024, p = 1.000). Furthermore, there was an interaction effect between age and stimulus type ($F_{(4,44)}$= 10.945, p < .001, $p\eta^2$= .499).

As in previous analysis, three additional ANOVAs independently investigated the effect of age in τ fixation durations for each stimulus type. τ fixation durations decreased with age in naturalistic videos ($F_{(2,22)}$=32.872, p < .001, $p\eta^2$= .749) but not in abstract videos. A marginal effect was found for static images suggesting that τ also decreased with age for this viewing condition ($F_{(2.22)}$=2.935, p = .074, $p\eta^2$= .211). Post hoc t-test revealed significant differences between the first (M = 440 ms, SD = .019) and the second visits (M = 299 ms, SD = .024, p < .001) for

the naturalistic videos. However, there were no differences in τ between any of the visits in the static images condition.

6.6.3.1.2 Differences in fixation durations between stimulus types at different visits

A one-way ANOVA revealed a significant effect of stimulus type on mean fixation durations at visit one, $F(1.210, 13.315) = 65.310$, $p < .001$, $p\eta^2 = .856$. Post hoc paired comparisons revealed that static images elicited significantly shorter fixations (M = 402 ms, SD = .061) than naturalistic videos (M = 681 ms, SD = .050, $p < .001$), or abstract videos (M = 694 ms, SD = .121, $p < .001$). However, no significant differences were found between the mean fixation durations in naturalistic and abstract videos at this visit.

A significant effect of stimulus type on mean fixation durations was found at visits two ($F(1.315, 14.462) = 43.630$, $p < .001$, $p\eta^2 = .799$) and three ($F(1.200, 13.205) = 73.130$, $p < .001$, $p\eta^2 = .869$). Paired comparisons showed significant effects between all the three conditions at both visit two (Naturalistic videos: M = 565 ms, SD = .026; abstract videos: M = 695 ms, SD = .050; static images: M = 357 ms, SD = 009; all $p < .05$) and visit three (Naturalistic videos: M = 515 ms, SD = .020; abstract videos: M = 727 ms, SD = .043; static images: M = 328 ms, SD = 009; all $p < .001$).

In order to further understand the effects described above, the ex-Gaussian components (μ, σ, and τ) for the fixation duration distributions were also analysed. A significant effect of stimulus type on μ fixation durations was found at visit one ($F(2, 22) = 18.775$, $p < .001$, $p\eta^2 = .631$). Paired comparisons showed that μ fixation durations were significantly longer for static images (M = 278 ms, SD = .013) compared to naturalistic (M = 241 ms, SD = .014, $p = .010$) or abstract videos (M = 230 ms, SD = .010, $p < .001$). A marginally significant effect of stimulus type on μ fixation durations was found at visit two ($F(2, 22) = 2.766$, $p = .085$, $p\eta^2 = .201$), even though post hoc t-tests showed no differences between the three viewing conditions. No effects were found for visit three, suggesting a greater stabilization on eye-movement control at this age point.

A significant effect of stimulus type on σ fixation durations was found at visit one ($F(2, 22) = 5.675$, $p = .010$, $p\eta^2 = .340$). Pairwise comparisons showed a significant effect between static images (M = 68 ms, SD = .008) and naturalistic (M = 52 ms, SD = .008; $p = .50$) and abstract

stimuli (M = 50 ms, SD = .009; p = .004). Nevertheless, no differences between both dynamic conditions were found.

The results for τ were very similar to the ones obtained for mean fixation durations. Significant effects of stimulus type on τ fixation durations were found at visit one (F(1.256, 13.817) = 79.774, p < .001, $p\eta^2$= .879), visit two (F(1.349,14.839) = 45.755, p < .001, $p\eta^2$= .806), and visit three (F(1.211, 13.325) = 54.347, p < .001, $p\eta^2$= .832). Pairwise comparisons showed that τ in static images at visit one (M = 124 ms, SD = .012) was significantly shorter than in the two dynamic conditions (naturalistic: M = 440 ms, SD = .019, p < .001; abstract: M = 464 ms, SD = .036, p < .001). Nonetheless, no differences were found between naturalistic and abstract videos at visit one. At visits two (naturalistic: M = 299 ms, SD = .024; abstract: M = 465 ms, SD = .050; static: M = 104 ms, SD = .013; all p < .010) and three (naturalistic: M = 269 ms, SD = .023; abstract: M = 499 ms, SD = .052; static: M = 94 ms, SD = .008, all p < .001), however, there were significant differences between all three viewing conditions.

6.6.3.1.3 Distributions of fixation durations

Typically, distributions of fixation durations are positively skewed both in infants and adults (Harris et al., 1988). This effect is more evident in younger infants as a result of the higher variability in fixation durations that they present (see Chapters 4 and 5). Figure 6-5 shows the distributions of fixation durations and their ex-Gaussian fittings for all the visits and viewing conditions described in this study. These distributions include the data from all the participants in the study. More precisely, the first visit distribution includes 1905 fixations for the naturalistic condition, 1993 fixations for the abstract condition, and 1393 fixations for the static condition; the second visit distribution includes 3119 fixations for the naturalistic condition, 2722 fixations for the abstract condition, and 2037 fixations for the static condition; finally, the third visit distribution includes 3261 fixations for the naturalistic condition, 2071 fixations for the abstract condition, and 2011 fixations for the static condition.
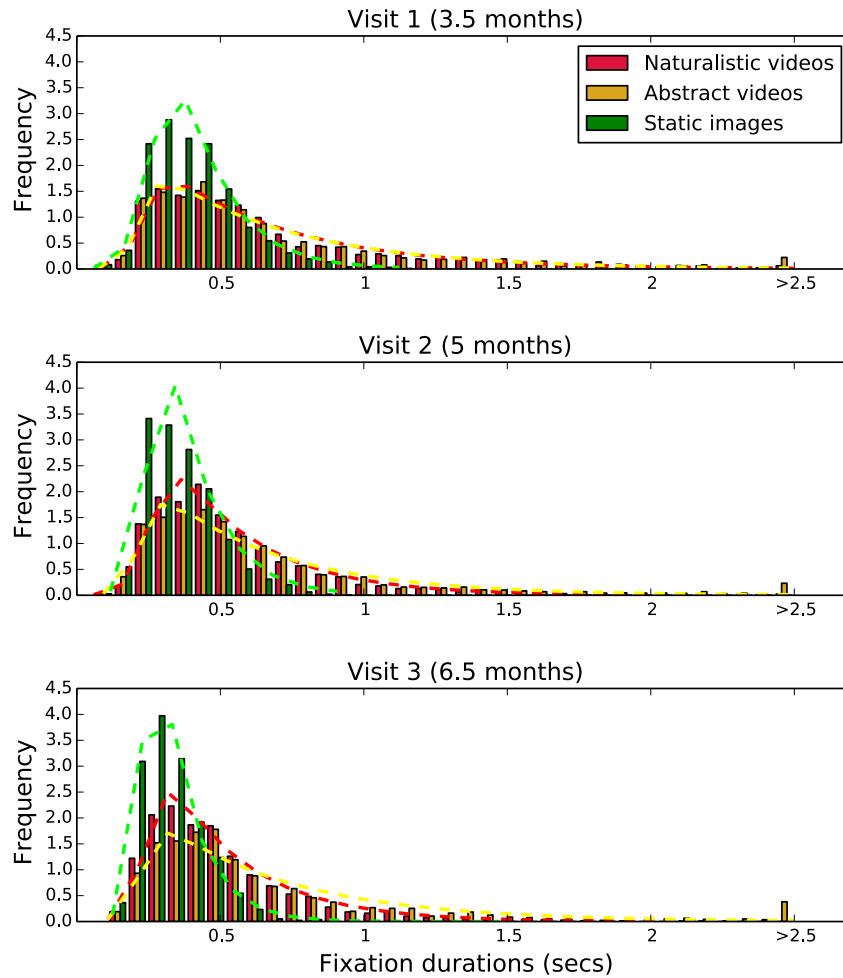
Figure 6-5 Longitudinal study 1: Distributions of fixation durations and the ex-Gaussian fitting for the three visits.

The visual inspection of these graphs reasserts the results from the previous section: While at visit one naturalistic and abstract distributions of fixation durations follow the same pattern, fixation durations for static images already show a much higher proportion of shorter fixations (<= 0.5 seconds, see Table 6-3) and less variability. From the second visit on, the differences between naturalistic and abstract distributions start being noticeable and become more evident at visit three. The percentage of short fixations (<= 0.5 seconds) increases over time for naturalistic and static conditions, while it remains the same for the abstract condition (see Table 6-3). This is evident in the increase in the peak (or the mode) of the naturalistic and static distributions over time and their decreasing slope. The percentage of very long fixations (>= 2.5 seconds) decreases for the naturalistic condition whereas for the abstract conditions it remains

stable across visits. No long fixations were detected for the static images condition for any of the visits.

Table 6-3 Long and short fixations. Descriptive statistics and percentages.

		Naturalistic		Abstract		Static	
		<= 0.5 secs	>= 2.5 secs	<= 0.5 secs	>= 2.5 secs	<= 0.5 secs	>= 2.5 secs
Visit 1	N	897	6	958	24	1112	0
	M	0.351	2.657	0.350	2.84	0.341	---
	SD	0.092	0.140	0.093	0.282	0.088	---
	%	47.09	0.32	48.07	1.21	79.83	0
Visit 2	N	1759	4	1356	39	1804	0
	M	0.356	2.775	0.346	3.139	0.325	---
	SD	0.088	0.278	0.094	0.63	0.085	---
	%	56.40	0.12	49.81	1.43	88.56	0
Visit 3	N	2017	0	1035	50	1861	0
	M	0.345	---	0.351	3.147	0.310	---
	SD	0.089	---	0.093	0.560	0.084	---
	%	61.85	0	49.98	2.41	92.84	0

6.6.3.2 Individual differences across viewing conditions and visits

6.6.3.2.1 Across viewing conditions

As explained in the introduction of this chapter, even though few studies have investigated individual differences in fixation durations in infancy (Papageorgiou et al., 2014; Wass & Smith, 2014), none of them analysed this stability through the first year of life. Table 6-4 shows the correlations and pairwise comparisons between mean fixation durations from the different viewing conditions for the three different visits. This table can be interpreted based on the following statements:

- Significant correlation + non-significant t-test: The fixation duration distributions for both viewing conditions are similar.

- Significant correlation + significant t-test: Even though mean fixation durations are significantly different they still correlate. These results reflect individual differences, i.e. one individual's mean fixation durations is consistently lower than another's.

- Non-significant correlation + non-significant t-test: Mean fixation durations for both conditions are not significantly different but do not correlate. The measures may not be stable enough to draw strong conclusions.

- Not significant correlation + significant t-test: Even though mean fixation durations are significantly different there is no relation between them.

Table 6-4 Correlations and pairwise comparisons between viewing conditions at different visits. All t-tests have been corrected with Bonferroni correction.

		Visit 1	Visit 2	Visit 3
Mean Naturalistic X Abstract	Pearson Corr.	$r(12) = -.001, p = .998$	$r(12) = .718, p = .009$	$r(12) = .631, p = .028$
	T-test	$t(11) = .342, p = 1.000$	$t(11) = 3.341, p = .020$	$t(11) = 6.015, p < .001$
Mean Naturalistic X Static	Pearson Corr.	$r(12) = .438, p = .154$	$r(12) = .744, p = .006$	$r(12) = .654, p = .021$
	T-test	$t(11) = 16.485, p < .001$	$t(11) = 10.110, p < .001$	$t(11) = 11.734, p < .001$
Mean Abstract X Static	Pearson Corr.	$r(12) = .621, p = .031$	$r(12) = .575, p = .050$	$r(12) = .250, p = .432$
	T-test	$t(11) = 10.543, p < .001$	$t(11) = 7.450, p < .001$	$t(11) = 9.489, p < .001$

At visit one it is already possible to find individual differences on mean fixation durations between abstract videos and static images (see Table 6-4). The relation between naturalistic videos and static images seems to follow the same tendency, even though the correlation is still not significant. The analysis between mean fixation durations on naturalistic and abstract videos, however, reveals no correlations or significant differences, which suggest that these two measures may still not be stabilized at this age point. These findings can be better explained by looking at the graphs (Figure 6-6, first column), where it is possible to see how some infants already presented longer mean fixation durations for the abstract videos compared to naturalistic videos (this trend will be more obvious at visit two) while others showed the opposite pattern, or no differences at all. This explains the lack of correlations and significant differences between these two measures and suggests that from 3 to 4 months infants may be going through a transitional age where the abstract stimuli starts to generate more interest.

At visit two individual differences were found between the means for all the viewing conditions (see Table 6-4). These results evidence a gain in top-down control from the previous visit.

At visit three, individual differences were found between naturalistic and abstract videos (*see* Table 6-4), as well as between naturalistic videos and static images. Surprisingly, no individual differences were found between abstract videos and static images (see Figure 6-6).
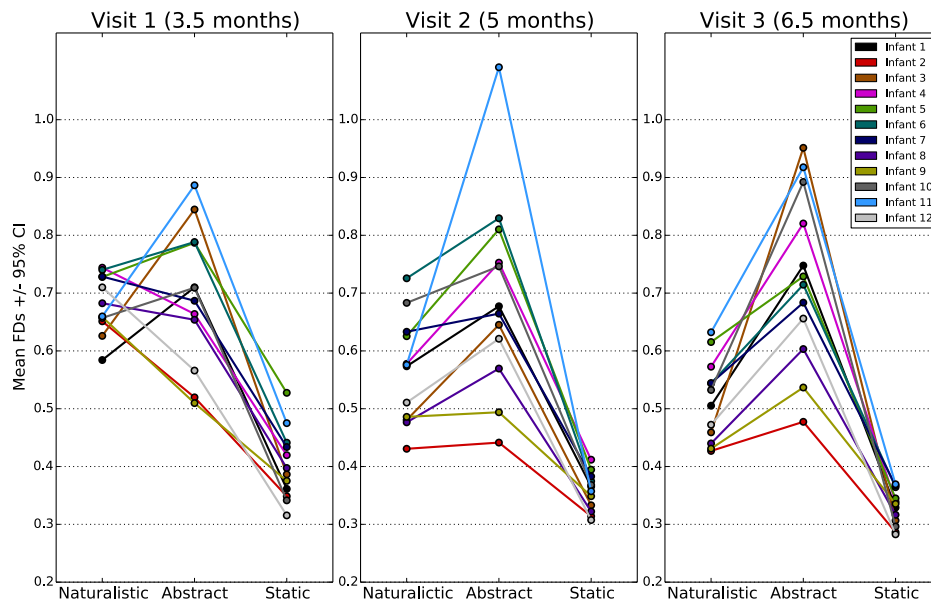
Figure 6-6 Longitudinal study 1: Individual differences across viewing conditions for the three visits.

6.6.3.2.2 Across age

Correlations were found between all the visits for almost all the viewing conditions (see Table 6-5). Nevertheless, the correlation between visits one and two for the naturalistic condition was still not significant.

Table 6-5 Correlations between different visits.

	Naturalistic	Abstract	Static
Means. Visits 1 X 2	$r(12) = .405, p = .191$	$r(12) = .841^{**}, p = .001$	$r(12) = .635^{*}, p = .026$
Means. Visits 2 X 3	$r(12) = .753^{**}, p = .005$	$r(12) = .787^{**}, p = .002$	$r(12) = .698^{*}, p = .012$

Abstract videos displayed the strongest correlations between the three different assessments, where mean fixation durations highly correlated across the three different visits (see Table 6-5). Figure 6-7 displays the graphs with all the individual means for each viewing condition across age.
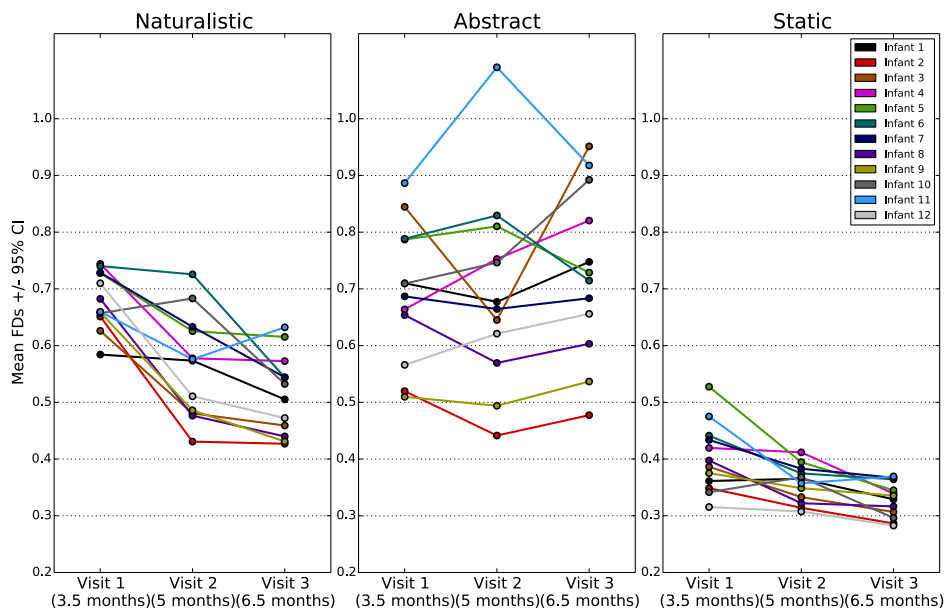
Figure 6-7 Longitudinal study 1: Individual differences across visits for the three viewing conditions.

The graphs show that the trajectories for fixation durations in abstract videos are very stable across visits. In other words, infants that presented fixations of a certain duration for abstract videos during the first assessment also tended to display fixations of a very similar duration on the forthcoming visits. Furthermore, infants displayed a high range of mean fixation durations for this viewing condition with some infants that presented relatively short fixation durations across visits (e.g., 509 ms) compared to others with longer durations (e.g., 887 ms). This stability is also consistent with previously reported results (section 6.6.3.1) that revealed no age effect in mean fixation durations for abstract stimuli.

Mean fixation durations in static images also present high stability across all the assessments. Further, previously reported results showed how fixation durations in static images decreased with age. That means that infants that fixated longer (or shorter) during the first assessment, also fixated proportionally longer (or shorter) in later assessments.

195

Naturalistic videos did not present high stability across visits. Even though there were positive correlations between the second and the third visits, no relationship was found between mean fixation durations on the first and the second visits.

6.6.3.3 The influence of disengagement in fixation durations

A gap-overlap effect was found in all the visits (visit one: $t(11) = 3.964$, $p = .002$; visit two: $t(11) = 4.968$, $p < .001$; visit three: $t(11) = 4.297$, $p = .001$).

A one way ANOVA showed that disengagement latencies diminished with age ($F(2, 22) = 5.683$, $p = .010$, $p\eta^2 = .341$). Post-hoc pairwise comparisons showed a decrease in disengagement latencies from visit two ($M = 264$ ms, $SD = .056$) to visit three ($M = 113$ ms, $SD = .034$; $p = .009$).

Gap ($F(1.236, 13.599) = 8.436$, $p = .009$, $p\eta^2 = .434$), and overlap ($F(1.356, 14.919) = 6.080$, $p = .019$, $p\eta^2 = .356$) latencies also diminished with age, while no effect was found for the baseline condition. Pairwise comparisons showed a significant decrease for gap latencies from visit 1 ($M = 581$ ms, $SD = .026$) to visit 2 ($M = 469$ ms, $SD = .022$; $p = .020$), and another significant decrease for the overlap latencies from visit 2 ($M = 771$ ms, $SD = .059$) to visit 3 ($M = 602$ ms, $SD = .033$; $p = .008$).

6.6.3.4 Correlations between disengagement latencies and fixation durations

There were no significant correlations between disengagement latencies and mean fixation durations at visits one and three (see Table 6-6). However, at visit two disengagement latencies correlated with mean fixation durations for naturalistic videos ($r(12) = .652$, $p = .021$) and static images ($r(12) = .649$, $p = .022$).

Table 6-6 Correlations between disengagement latencies and mean fixation durations.

	Visit 1	Visit 2	Visit 3
Naturalistic x disengagement	$r(12) = .133$, $p = .681$	$r(12) = .652$, $p = .021$	$r(12) = .280$, $p = .378$
Abstract x disengagement	$r(12) = -.024$, $p = .940$	$r(12) = .254$, $p = .426$	$r(12) = .335$, $p = .287$
Static x disengagement	$r(12) = .272$, $p = .392$	$r(12) = .649$, $p = .022$	$r(12) = .070$, $p = .829$

6.7 Longitudinal study 2: Six to 12 month-olds

6.7.1 Methods

6.7.1.1 Participants

In total, 19 typically developing infants (9 girls, 10 boys) participated in this longitudinal study. An additional infant was excluded from the analysis due to excessive interaction with the caregiver during the testing sessions and data quality issues.

All infants scored within their age range on the Bayley Scales of Infant Development (Bayley, 1993). Infants were tested every 8 weeks at approximately 6, 8, 10 and 12 months calculated from the due date. Fifteen of the infants carried out all four measurement sessions; three infants completed three sessions; one infant completed two sessions. Visits were excluded when data quality issues were found or when the infants did not provide enough fixations (50 fixations) for each task.

Mean ages were 192.53 days (SD = 11.673), 249.37 days (SD = 11.955), 304.47 days (SD = 12.365), and 359.84 (SD = 12.829). Visits were scheduled at times of the day when the caregiver considered the infant was going to be most alert. In case the measurement session was not successful due to the infant being unhappy or other technical issues, another visit was scheduled as promptly as possible during the next 7 days.

Most infants (N = 13) were of white middle socioeconomic status. The infants were recruited via magazine advertisements, social networking media, and flyers. Families were given baby t-shirts or bags as gifts for their participation. The study protocol was approved by the Birkbeck, Psychological Sciences Ethics Committee. Table 6-7 shows the age of each participant at the time of visit.

Table 6-7 Information about the visits for the longitudinal study from 3 to 6 months

Subject	Gender	Age Visit 1 (Days)	Age Visit 2 (Days)	Age Visit 3 (Days)	Age Visit 4 (Days)	Visits
1	Boy	184	245	298	354	1,2,3,4
2	Boy	195	249	305	361	1,2,4
3	Boy	197	254	310	367	1,2,3,4
4	Boy	203	259	315	371	1,2,3,4
5	Girl	200	254	309	364	1,2,3,4
6	Girl	173	229	285	341	1,2,3,4
7	Girl	207	266	322	375	1,2,3,4
8	Girl	199	259	319	378	1,3,4
9	Boy	183	238	294	348	1,2,3,4
10	Boy	181	237	293	348	1,3,4
11	Girl	210	266	324	374	1,2,3,4
12	Girl	207	263	315	375	3,4
13	Boy	252	310	368	252	1,2,3,4
14	Boy	245	300	357	245	1,2,3,4
15	Boy	255	304	360	255	1,2,3,4
16	Girl	255	304	360	255	1,2,3,4
17	Boy	251	307	358	251	1,2,3,4
18	Girl	237	294	348	237	1,2,3,4
19	Girl	224	277	330	224	1,2,3,4

6.7.1.2 Apparatus, stimuli, gaze-contingent paradigms and design and procedure

The apparatus was identical to the one described in Chapter 4. The stimuli, gaze contingent paradigms, and the design and procedure were identical to what was described in sections 6.6.1.3, 6.6.1.4, and 6.6.1.5 respectively.

6.7.2 Analysis

6.7.2.1 Fixation detection and coding

The fixation detection method and coding was identical to the one described in section 6.6.2.1.

6.7.2.2 Data- quality analysis

As in section 6.6.2.3, spatial precision was measured per participant and visit by calculating the RMS. The mean RMS-s were .26° (SD = .09) for the first visit, .26° (SD = .09) for the second visit, .24° (SD = .10) for the third visit, and .21° (SD = .08) for the last visit.

A one-way ANOVA showed no age effect in data quality (F (3,42) = 2.281, p = .093). Additionally, correlational analysis between RMS-s and fixation durations were run for each age point in order to ensure data quality was not interfering with the fixation detection method. No correlations were found between data quality and spatial precision for any of the visits (see Table 6-8).

Table 6-8 Correlations between fixation durations and RMS in the four visits.

		Visit 1	Visit 2	Visit 3	Visit 4
Fixation durations vs. RMS	Pearson	.293	-.042	.297	-.228
	Sig. (2-tailed)	.238	.870	.264	.348
	N	18	18	16	19

6.7.2.3 Cross-validation

6.7.2.3.1 Fixation durations

The process of cross-validating fixations was the same described in section 6.6.2.3.1. In total, the external coder coded 20 % of the data from the infants that participated in this study.

The inter-rater reliability between the means and the number of detected fixations was evaluated using the ICC (Hallgren, 2012). A strong agreement between the mean fixation

durations was found (with an ICC of .908, $p < .001$). Additionally, there was also an agreement in the number of fixations detected (with an ICC of .953, $p < .001$).

6.7.2.3.2 Gap-overlap

The process of cross-validating the gap-overlap trials was the same described in section 6.6.2.3.2. An external coder that was naïve to the expected outcomes evaluated the validity of the gap-overlap trials for the 20% of the infants that participated in this study.

Interrater reliability was measured using the Kappa statistic with a view to determine consistency among raters. This analysis showed a strong agreement between the two coders, Kappa = .891 ($p < .001$), 95% CI (0.869, 0.913).

6.7.2.4 Statistical analysis

The general statistical analysis was the same described in section 6.6.2.4.

6.7.2.4.1 The analysis of longitudinal variations in fixation durations

• *Multiple imputation and ANOVAs*

Unlike the previous cohort, the data for the present cohort is missing at random. The percentage of missing values was 5.36% for visit one (N=1), 15.79% for visit two (N=3), 5.36% for visit three (N=1), and 0% for visit four (N=0).

Analysis methods such as ANOVAs, require the use of additional techniques that deal with the missing data. The most common technique is *listwise deletion*, which involves the removal of all the data from those subjects with even a single missing data point. However, decreasing the effective sample size also decreases the power of the analysis and can potentially cause bias in the results. Other methods such as *imputation techniques* maintain the number of subjects by replacing missing data with a substituted value. For instance, when performing a *single imputation* the missing value can be replaced with the mean of the variable for all other cases. In these cases, even though the sample size and hence the power of the analysis is maintained, a correlational analysis may not be as reliable.

The ANOVA analysis was performed using a multiple imputation technique that aimed to maintain the statistical power. The technique used is called the Markov chain Monte Carlo (MCMC; Schafer, 1997) procedure and implements the Gibbs Sampler fully conditional specification (FCS) method (Van Buuren, Boshuizen, & Knook, 1999; Van Buuren, 2007).

Multiple imputation involves three main steps: (1) the selected method -MCMC- generates a number of plausible synthetic values or imputations (m) for each missing data point, which leads to m complete data sets; (2) each of the imputed data sets is analysed by the selected statistical method (e.g., t-test, ANOVA) generating a number (m) of slightly different and plausible results; and (3) the m estimates are pooled into a single estimate that combines the variations within and across the m imputed data sets. MCMC methods assume multivariate normality and are used to generate pseudorandom draws from multidimensional and otherwise intractable probability distributions via Markov chains. Markov chains are sequences of random variables in which the distribution of each element depends on the value of the previous one (Yuan, 2010). One of the benefits of this technique is that it is able to estimate missing values without making assumptions about expected distributions across the "population" of observations. Furthermore, it is especially suited for small datasets with a relatively small amount of missing data. The present MCMC method was programmed to compute 10 imputations.

The longitudinal variations in fixation durations for each of the imputed data sets were analysed using 3 (stimulus) x 4 (visit) mixed ANOVAs. Further, independent one-way ANOVAs were used in order to investigate the differences between different stimulus types and their independent longitudinal trajectories. Results were pooled together by calculating the means from the ANOVAs estimates from each imputation. In the cases where the correlation matrices between the different measurement points were heterogeneous the degrees of freedom of the *F* test were corrected according to the Greenhouse-Geisser method. Post-hoc pairwise comparisons were calculated based on Bonferroni correction.

This procedure was used to examine whether (1) fixation durations changed with age, (2) the viewing condition influenced fixation duration trajectories, and (3) the viewing conditions influenced fixation durations at each visit.

> • *Linear Mixed Models (LMM)*

The longitudinal variations of fixation durations for each viewing condition were also evaluated using linear mixed models (LMM; e.g., West, 2009). This inquiry aimed to reassure the results obtained with the ANOVAs and explain whether the data could also be described by a model that allows for inter-infant differences.

This type of multilevel model implements a regression procedure that allows for the analysis of individual changes across observations and the study of the relationships of variables in data sets with some type of dependency caused by a hierarchical study design. In a longitudinal design, the repeated measures are regarded as "nested" within individuals. These models present a number of advantages relative to other techniques used for the analysis of longitudinal data. For instance, they are able to accommodate unbalanced data sets with missing data points, which are very common in a longitudinal design.

Another advantage of LMMs over other techniques is that they permit the inclusion of fixed and random effects for the intercept and slope coefficients. Fixed effects are the constant regression coefficients associated either with the continuous covariates or the categorical factors where all the possible values are present in the data (e.g., gender). The levels of these factors will not vary for other theoretical replications of the study. On the other hand, random effects are those that will randomly vary for each study replication and allow different subjects to have unique trajectories.

The present LMMs evaluated the fixation duration trajectories for the different stimulus types including age as a fixed and a random effect. Furthermore, a random effect associated with the intercept for each subject was also included, in order to allow for different subjects to have different intercepts. The denominator degrees of freedom when fitting the LMM were computed using the Satterthwaite approximation method. For the longitudinal study described in section 6.6 a LMM analysis was not possible due to the low number of participants.

6.7.3 Results

6.7.3.1 The development of fixation durations

6.7.3.1.1 Multiple imputation

The ANOVA analyses described in this section were performed using the MCMC multiple imputation technique. First, ten plausible synthetic values or imputations were generated for each missing data point, leading to 10 different data sets. Secondly, each data set was analysed with the corresponding ANOVA, which generated 10 slightly different and plausible results. Finally, the 10 estimates were pooled into a single estimate by calculating their means, combining the variations within and across the 10 imputed data sets.

Appendix D shows the descriptive statistics for mean fixation durations and the ex-Gaussian components (μ, σ, and τ) for the original data set, the data sets generated with the MCMC multiple imputation technique, and the pooled data.

6.7.3.1.2 Age effect: Fixation duration trajectories for different viewing conditions

A mixed 3 x 4 ANOVA revealed a significant main age effect in mean fixation durations ($F_{(1.980, 35.647)}$= 5.817, p = .0147, $p\eta^2$= . 0.241) and a stimulus type effect ($F_{(1.09, 19.603)}$= 108.406, p < .001, $p\eta^2$= . 0.856). Pairwise comparisons revealed significant differences in mean fixation durations between the second (M = 526 ms, SD = .024) and the third (M = 466 ms, SD = .017; p = 0.052), the first (M = 543 ms, SD = .022) and the third (p = .008), and the first and the fourth (M = 493 ms, SD = .016; p = .019) visits. Furthermore, there was a marginally significant interaction effect between age and stimulus type ($F_{(2.381, 42.854)}$= 2.805, p = . 0.084, $p\eta^2$= . 0.134; See Figure 6-8).
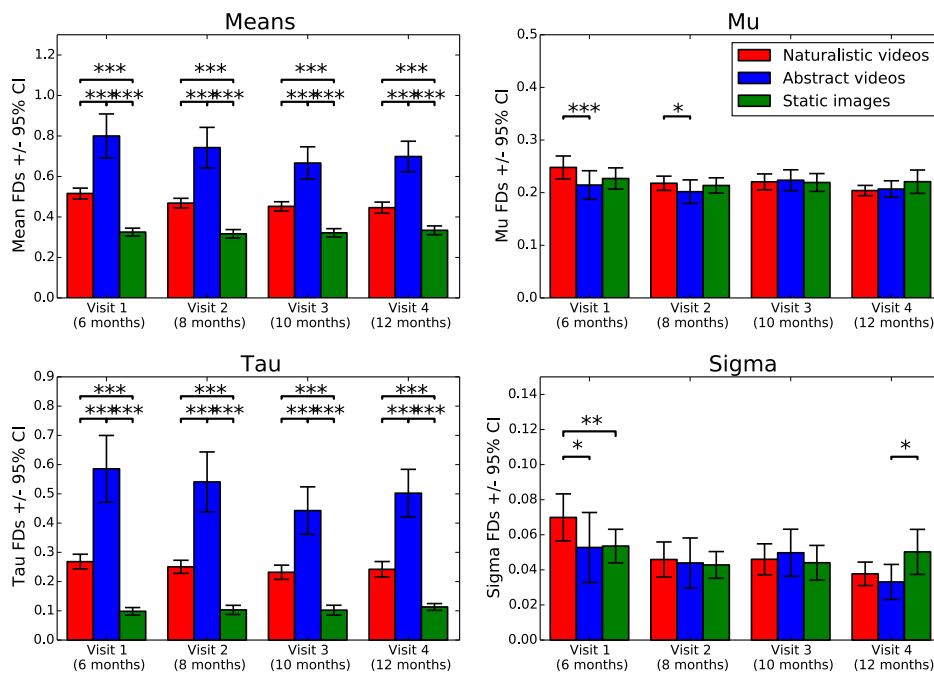
Figure 6-8 Longitudinal study 2: Fixation durations histograms for the means and the ex-Gaussian components.

The interaction between age and stimulus type was examined running three additional ANOVAs that examined the age effect on each stimulus type independently. Mean fixation durations

decreased with age in naturalistic videos ($F(3,54)$=12.976, $p < .001$, $p\eta^2$= .409) and abstract videos ($F(3,54)$=3.363, $p = .047$, $p\eta^2$= .156) but not in static images. Post hoc t-test revealed marginally significant differences for naturalistic videos between the first (M = 514 ms, SD = .013) and second visits (M = 477 ms, SD = .013; $p = .078$), and second and third (M = 448 ms, SD = .012; $p = .052$) visits. For abstract videos, there were marginally significant differences between the second (M = 780 ms, SD = .063) and the third (M = 661 ms, SD = .039; $p = .083$) visits, and significant differences between the first (M = 794 ms, SD = .053) and the third visits ($p = .029$).

In order to investigate and explain the origin of the effects described above further analyses were performed using the ex-Gaussian components (μ, σ, and τ) for the fixation duration distributions instead of the means. A mixed 3 x 4 ANOVA revealed a marginally significant main age effect in μ fixation durations ($F(1.867,33.607)$=3.017, $p = .089$, $p\eta^2$= .142), but not significant stimulus type effect was found. Moreover, there was a significant interaction between age and stimulus type ($F(6,108)$=3.369, $p = .022$, $p\eta^2$= .156). Three additional ANOVAs investigated this interaction independently revealing that μ fixation durations decreased with age for naturalistic videos ($F(1.844,33.192)$=8.200, $p = .007$, $p\eta^2$= .305), but not for abstract videos or static images where no significant effects were found. Pairwise comparisons showed significant differences in μ fixation durations in naturalistic videos between the first (M = 233 ms, SD = .011) and the second visits (M = 214 ms, SD =.009, $p = .034$).

A mixed 3 x 4 ANOVA revealed an age effect for σ fixation durations ($F(3,54)$=5.891, $p = .002$, $p\eta^2$= .247). Furthermore, an interaction between age and stimulus type was found ($F(3.581,64.464)$=3.049, $p = .020$, $p\eta^2$= .144). Post-hoc pairwise comparisons showed significant differences between the first (M = 60 ms, SD = .006) and the second visits (M = 45 ms, SD =.004, $p = .008$). Three additional ANOVAs investigated the age and stimulus type interaction in σ fixation durations independently revealing a significant age effect in naturalistic videos ($F(3,54) = 12.861$, $p < .001$, $p\eta^2$= .415) and a marginally significant effect in abstract videos ($F(3,54) = 2.347$, $p = .085$, $p\eta^2$= .115). On the other hand, no effects were found for static images. Post-hoc pairwise comparisons revealed significant differences in σ fixation durations for naturalistic videos between the first (M = 69 ms, SD =.006) and the second visits (M = 51 ms, SD =.006; $p = .013$), and a marginally significant effect for abstract videos between the third (M = 49 ms, SD =.006) and the fourth visits (M = 33 ms, SD =.005; $p = .055$).

A mixed 3 x 4 ANOVA showed significant age ($F(3,54) = 2.982$, $p = .047$, $p\eta^2 = .142$) and stimulus effects ($F(1.057,19.032) = 112.292$, $p < .001$, $p\eta^2 = .857$) in τ fixation durations. Nevertheless, no interaction effect was found. Post-hoc pairwise comparisons showed significant differences between τ fixation durations at visits one ($M = 322$ ms, $SD = .023$) and three ($M = 262$ ms, $SD = .017$; $p = .026$).

This same analysis was also performed with the LMM procedure. A LMM revealed an age effect on mean fixation durations for naturalistic videos, with the estimated fixed effect being -.026. This suggests that for each visit there was an estimated decrease on fixation durations of .026 ms, and that this fixed effect is significantly different from 0 based on the t-test ($t(17.568) = -5.412$, $p < .001$). A marginally significant age effect in mean fixation durations was also found for abstract videos ($t(17.565) = -1.809$, $p = .088$), with an estimated fixed effect of -.028. No age effect was found for the static images condition.

As before, further analyses were performed using the ex-Gaussian components (μ, σ, and τ) for the fixation duration distributions instead of the means. A LMM revealed a marginally significant age effect for μ fixation durations in naturalistic videos ($t(1.356) = -4.628$, $p = .086$) with an estimated fixed effect of -.013. Nevertheless, no age effects were found for the other viewing conditions.

LMMs showed a decrease in σ fixation durations for naturalistic videos ($t(15.009) = -5.025$, $p < .001$) with an estimated fixed effect of -.009. A marginal age effect was also found for the abstract condition ($t(17.096) = -1.853$, $p = .081$) with an estimated fixed effect of -.006. No age effect was found for the static images condition.

Finally, the age effect for τ fixation durations was analysed. Even though no age effects were found for the naturalistic and abstract videos, a marginally significant age effect was found for static images ($t(24.546) = 1.773$, $p = .089$) indicating that τ fixation durations increased with age with an estimated fixed effect of .004. The results obtained with the LMM procedure resembled those obtained in the previous ANOVAs analysis using multiple imputation.

6.7.3.1.3 Differences in fixation durations between stimulus types at different visits

Three one-way ANOVAs revealed significant effects of stimulus type on mean fixation durations at visits one ($F(1.098, 19.753) = 71.423$, $p < .001$, $p\eta^2 = .791$), two ($F(1.072, 19.303) = 45.853$, $p < .001$, $p\eta^2 = .711$), three ($F(1.152, 20.738) = 67.565$, $p < .001$, $p\eta^2 = .785$) and four ($F(1.117, 20.114) = 87.43$, $p < .001$, $p\eta^2 = .829$). Post-hoc pairwise comparisons showed significant differences in mean fixation durations between all viewing conditions in visits one (Naturalistic videos: $M = 514$ ms, $SD = .013$; abstract videos: $M = 794$ ms, $SD = .053$; static images: $M = 328$ ms, $SD = .009$; all $p < .001$), two (Naturalistic videos: $M = 477$ ms, $SD = .013$; abstract videos: $M = 780$ ms, $SD = .063$; static images: $M = 320$ ms, $SD = .009$; all $p < .001$), three (Naturalistic videos: $M = 448$ ms, $SD = .012$; abstract videos: $M = 661$ ms, $SD = .039$; static images: $M = 322$ ms, $SD = .009$; all $p < .001$), and four (Naturalistic videos: $M = 446$ ms, $SD = .013$; abstract videos: $M = 699$ ms, $SD = .036$; static images: $M = 334$ ms, $SD = .001$; all $p < .001$).

In order to further understand the effects described above, the ex-Gaussian components (μ, σ, and τ) for the fixation duration distributions were also analysed. A significant effect of stimulus type on μ fixation durations was found at visit one ($F(2,36) = 5.825$, $p = .016$, $p\eta^2 = .241$). This same effect was marginally significant at visit two ($F(2,36) = 4.321$, $p = .062$, $p\eta^2 = .791$), and disappeared at visits three and four. Post-hoc pairwise comparisons revealed significant differences between naturalistic and abstract videos at visits one (naturalistic videos: $M = 250$ ms, $SD = .012$; abstract videos: $M = 215$ ms, $SD = .014$; $p = .002$) and two (naturalistic videos: $M = 221$ ms, $SD = .009$; abstract videos: $M = 203$ ms, $SD = .012$; $p = .038$).

Significant effects of stimulus type on σ fixation durations were found at visits one ($F(2,36) = 3.900$, $p = .046$, $p\eta^2 = .177$) and four ($F(2,36) = 4.368$, $p = .020$, $p\eta^2 = .195$). Post-hoc pairwise comparisons at visit one revealed significant differences between naturalistic ($M = 72$ ms, $SD = .006$) and abstract videos ($M = 55$ ms, $SD = .009$; $p = .029$), and between naturalistic videos and static images ($M = 54$ ms, $SD = .004$; $p = .005$). On the other hand, at visit four significant differences between abstract videos ($M = 33$ ms, $SD = .005$) and static images were found ($M = 50$ ms, $SD = .006$; $p = .017$).

The results for τ analysis resembled the ones obtained from the means. Three one-way ANOVAs revealed significant effects of stimulus type on τ fixation durations at visits one

($F(1.054, 18.971) = 64.679$, $p < .001$, $p\eta^2 = .779$), two ($F(1.076, 19.369) = 39.209$, $p < .001$, $p\eta^2 = .630$), three ($F(1.121, 20.180) = 69.260$, $p < .001$, $p\eta^2 = .791$) and four ($F(1.116, 20.084) = 73.79$, $p < .001$, $p\eta^2 = .804$). Post-hoc pairwise comparisons showed significant differences in τ fixation durations between all viewing conditions in visits one (Naturalistic videos: $M = 269$ ms, $SD = .013$; abstract videos: $M = 597$ ms, $SD = .058$; static images: $M = 99$ ms, $SD = .007$; all $p < .001$), two (Naturalistic videos: $M = 259$ ms, $SD = .014$; abstract videos: $M = 561$ ms, $SD = 72$ ms; static images: $M = 110$ ms, $SD = .009$; all $p < .005$), three (Naturalistic videos: $M = 235$ ms, $SD = .012$; abstract videos: $M = 448$ ms, $SD = .039$; static images: $M = 104$ ms, $SD = .008$; all $p < .001$), and four (Naturalistic videos: $M = 242$ ms, $SD = .013$; abstract videos: $M = 502$ ms, $SD = .039$; static images: $M = 113$ ms, $SD = .006$; all $p < .001$).

6.7.3.1.4 Distributions of fixation durations

As described before, distributions of fixation durations are positively skewed both in infants and adults (Harris et al., 1988). Figure 6-9 shows the distributions of fixation durations and their ex-Gaussian fittings for all the visits and viewing conditions described in this study.
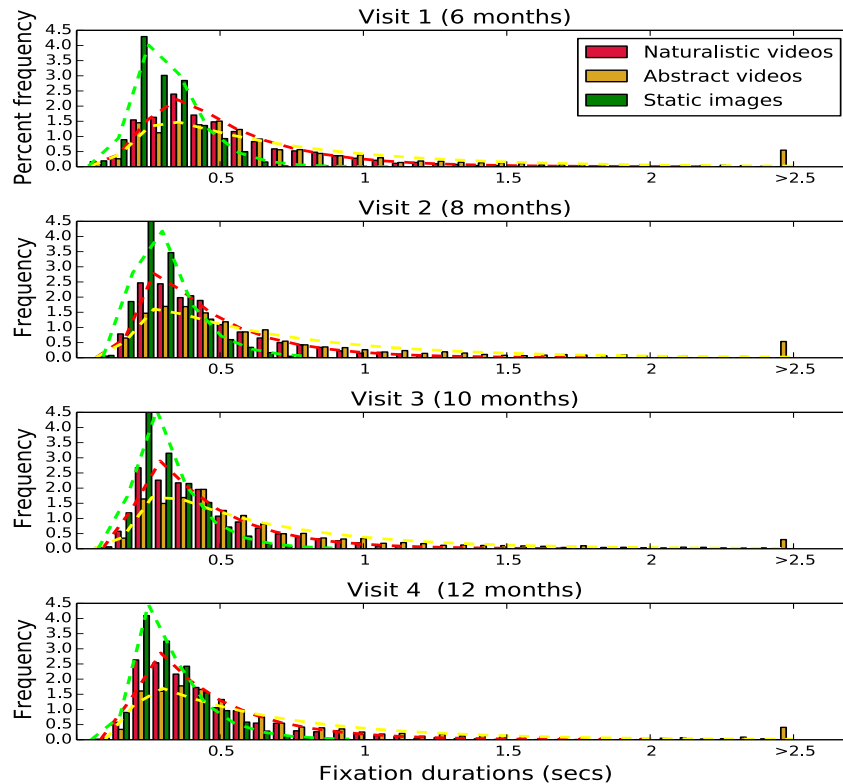
Figure 6-9 Longitudinal study 2: Distributions of fixation durations and their ex-Gaussian fittings for the four visits.

These distributions include the data from all the participants in the study. More precisely, the first visit distribution includes 4901 fixations for the naturalistic condition, 3148 fixations for the abstract condition, and 2712 fixations for the static condition; the second visit distribution includes 4635 fixations for the naturalistic condition, 2498 fixations for the abstract condition, and 2646 fixations for the static condition; the third visit distribution includes 5191 fixations for the naturalistic condition, 3031 fixations for the abstract condition, and 3158 fixations for the static condition; finally, the fourth visit distribution includes 5311 fixations for the naturalistic condition, 2797 fixations for the abstract condition, and 2899 fixations for the static condition (see Table 6-9).

Table 6-9 Long and short fixations. Descriptive statistics and percentages.

		Naturalistic		Abstract		Static	
		<=0.5secs	>= 2.5 secs	<=0.5secs	>= 2.5 secs	<=0.5secs	>= 2.5 secs
Visit 1	N	3044	0	1534	114	2539	0
	M	0.338	---	0.342	3.550	0.294	---
	SD	0.098	---	0.105	0.901	0.092	---
	%	62.11	0	48.73	3.62	93.62	0
Visit 2	N	3194	0	1283	87	2468	0
	M	0.329	---	0.339	3.707	0.294	---
	SD	0.092	---	0.096	1.041	0.085	---
	%	68.91	0	51.36	3.48	93.27	0
Visit 3	N	3699	0	1652	56	2932	0
	M	0.327	---	0.346	3.500	0.297	---
	SD	0.091	---	0.095	0.929	0.086	---
	%	71.26	0	54.50	1.848	92.84	0
Visit 4	N	3797	0	1521	76	2601	0
	M	0.318	---	0.339	3.283	0.300	---
	SD	0.090	---	0.094	0.574	0.088	---
	%	71.49	0	54.38	2.72	89.72	0

The visual inspection of these graphs reasserts the results from the previous section: The percentage of short fixations (<= 0.5) seems to increase from visit one to three for naturalistic and to a lesser extent for abstract videos. On the other hand, these values remained rather

stable for the static images condition. Very long fixations were only found in the abstract videos condition.

6.7.3.2 Individual differences across viewing conditions and age

6.7.3.2.1 Across viewing conditions

The results from the correlational analysis were interpreted following the same criteria described in section 6.6.3.2. Correlations and pairwise comparisons were calculated after excluding the missing points, thus the results from the t-tests described in this section may slightly differ from the ones reported earlier when missing values were imputed. Furthermore, all t-tests were corrected with Bonferroni adjusted alpha values.

Table 6-10 shows the correlations and pairwise comparisons between mean fixation durations from the different viewing conditions for all visits. Results revealed significant differences between mean fixation durations in all viewing conditions for all the visits, few significant correlations at visits one, three, and four, and marginally significant correlations at visits one, two, and three (see Table 6-10). Therefore, these results evidence individual differences in mean fixation durations for some of the viewing conditions and visits, but not for all.

With the view to further investigate these results the same analysis was performed using the ex-Gaussian components (μ, σ, and τ). Table 6-11, 6-12 and 6-13 display the correlations and pairwise comparisons for μ, σ, and τ components for all visits. Results from the τ component resembled those observed for mean fixation durations, thus this suggests that the lack of stability across all viewing conditions is due to the variability in fixation duration distributions and the differences in the slope of the distribution across viewing conditions.

Table 6-10 Individual differences for mean fixation durations. This table shows the correlations and pairwise comparisons between viewing conditions at different visits. All t-tests have been corrected with Bonferroni correction.

		Visit 1	Visit 2	Visit 3	Visit 4
Mean Nat. X Abs.	Pearson Corr.	$r(18) = .673, p = .002$	$r(16) = .430, p = .097$	$r(18) = .568, p = .014$	$r(19) = .355, p = .135$
	T-test	$t(17) = 6.504, p < .001$	$t(15) = 6.296, p < .001$	$t(17) = 6.584, p < .001$	$t(18) = 7.373, p < .001$
Mean Nat. X Static	Pearson Corr.	$r(18) = .418, p = .084$	$r(16) = .246, p = .358$	$r(18) = .424, p = .079$	$r(19) = .702, p = .001$
	T-test	$t(17) = 15.700, p < .001$	$t(15) = 11.473, p < .001$	$t(17) = 11.664, p < .001$	$t(18) = 11.840, p < .001$
Mean Abs. X Static	Pearson Corr.	$r(18) = .319, p = .197$	$r(16) = .042, p = .877$	$r(18) = .270, p = .279$	$r(19) = .375, p = .114$
	T-test	$t(17) = 9.625, p < .001$	$t(15) = 8.907, p < .001$	$t(17) = 9.444, p < .001$	$t(18) = 10.847, p < .001$

Table 6-11 Individual differences for μ fixation durations. This table shows the correlations and pairwise comparisons between viewing conditions at different visits. All t-tests have been corrected with Bonferroni correction.

		Visit 1	Visit 2	Visit 3	Visit 4
Mean Nat. X Abs.	Pearson Corr.	$r(18) = .735, p = .001$	$r(16) = .722, p = .002$	$r(18) = .439, p = .068$	$r(19) = .701, p = .001$
	T-test	$t(17) = 3.988, p = .012$	$t(15) = 2.198, p = .528$	$t(17) = -.338, p = 1.000$	$t(18) = .533, p = 1.000$
Mean Nat. X Static	Pearson Corr.	$r(18) = .607, p = .008$	$r(16) = .803, p < .001$	$r(18) = .582, p = .011$	$r(19) = .473, p = .041$
	T-test	$t(17) = 2.395, p = .336$	$t(15) = .956, p = 1.000$	$t(17) = .164, p = 1.000$	$t(18) = 1.750, p = 1.000$
Mean Abs. X Static	Pearson Corr.	$r(18) = .711, p = .001$	$r(16) = .697, p = .003$	$r(18) = .327, p = .185$	$r(19) = .631, p = .004$
	T-test	$t(17) = 1.303, p = 1.000$	$t(15) = 1.566, p = 1.000$	$t(17) = .417, p = 1.000$	$t(18) = 1.629, p = 1.000$

Table 6-12 Individual differences for σ fixation durations. This table shows the correlations and pairwise comparisons between viewing conditions at different visits. All t-tests have been corrected with Bonferroni correction.

		Visit 1	Visit 2	Visit 3	Visit 4
Mean Nat. X Abs.	Pearson Corr.	$r(15) = .648, p = .009$	$r(15) = .574, p = .025$	$r(17) = .405, p = .107$	$r(19) = .384, p = .142$
	T-test	$t(17) = 2.512, p = .264$	$t(15) = .378, p = 1.000$	$t(17) = .599, p = 1.000$	$t(18) = .951, p = 1.000$
Mean Nat. X Static	Pearson Corr.	$r(18) = .507, p = .032$	$r(16) = .448, p = .082$	$r(18) = .096, p = .704$	$r(19) = .239, p = .324$
	T-test	$t(17) = 3.012, p = .096$	$t(15) = .769, p = 1.000$	$t(17) = .343, p = 1.000$	$t(18) = 1.931, p = .828$
Mean Abs. X Static	Pearson Corr.	$r(15) = .885, p < .001$	$r(15) = .406, p = .133$	$r(17) = -.355 p = .162$	$r(19) = .382, p = .144$
	T-test	$t(17) = .091, p = 1.000$	$t(15) = .169, p = 1.000$	$t(17) = .665, p = 1.000$	$t(18) = 2.632, p = .204$

Table 6-13 Individual differences for τ fixation durations. This table shows the correlations and pairwise comparisons between viewing conditions at different visits. All t-tests have been corrected with Bonferroni correction.

		Visit 1	Visit 2	Visit 3	Visit 4
Mean Nat. X Abs.	Pearson Corr.	$r(18) = .717$, p = .001	$r(16) = .393$, p = .132	$r(18) = .532$, p = .023	$r(19) = .216$, p = .373
	T-test	$t(17) = 6.901$, p < .001	$t(15) = 6.375$, p < .001	$t(17) = 6.424$, p < .001	$t(18) = 6.717$, p < .001
Mean Nat. X Static	Pearson Corr.	$r(18) = .001$, p = .997	$r(16) = -.006$, p = .982	$r(18) = .358$, p = .145	$r(19) = .495$, p = .031
	T-test	$t(17) = 12.642$, p < .001	$t(15) = 11.068$, p < .001	$t(17) = 11.413$, p < .001	$t(18) = 11.856$, p < .001
Mean Abs. X Static	Pearson Corr.	$r(18) = -.220$, p = .380	$r(16) = -.159$, p = .555	$r(18) = .156$, p = .535	$r(19) = -.071$, p = .773
	T-test	$t(17) = 8.820$, p < .001	$t(15) = 8.772$, p < .001	$t(17) = 8.882$, p < .001	$t(18) = 9.807$, p < .001

6.7.3.2.2 Across age

Correlations were found between all the visits for all viewing conditions (see Table 6-14). Static images displayed the strongest longitudinal correlations between the four different assessments. Figure 6-10 displays the graphs with all the individual means for each viewing condition across age. The graphs show that the trajectories for fixation durations in static images are very stable across visits. In other words, infants that presented fixations of a certain duration for static images during the first assessment also tended to display fixations of a very similar duration on the forthcoming visits.

Table 6-14 Correlations between different visits.

	Naturalistic	Abstract	Static
Means. Visits 1 X 2	$r(16) = .714$, $p = .002$	$r(16) = .708$, $p = .002$	$r(16) = .915$, $p < .001$
Means. Visits 2 X 3	$r(15) = .554$, $p = .032$	$r(15) = .513$, $p = .050$	$r(15) = .924$, $p < .001$
Means. Visits 3 X 4	$r(18) = .800$, $p < .001$	$r(18) = .626$, $p = .005$	$r(18) = .803$, $p < .001$
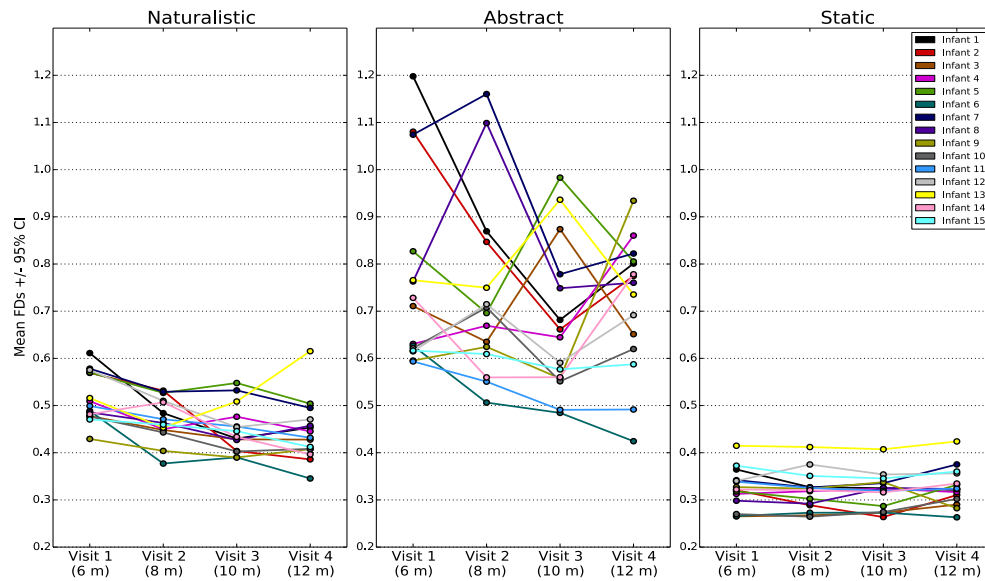
Figure 6-10 Longitudinal study 2: Individual differences across visits for the four viewing conditions.

6.7.3.3 The influence of disengagement in fixation durations

A gap-overlap effect was found for all visits (visit one: $t(17)$ = 5.231, p < .001; visit two: $t(15)$ = 5.006, p < .001; visit three: $t(17)$ = 4.710, p < .001; visit four: $t(18)$ = 6.832, p < .001).

A one-way ANOVA showed no age effect in disengagement latencies. Nevertheless, gap latencies decreased with age ($F(3,54)$ = 3.141, p = .033, $p\eta^2$= .149). Pairwise comparisons showed a significant decrease from visit 2 (M = 454 ms, SD = .014) to visit 4 (M = 419 ms, SD = .009; p = .030).

6.7.3.4 Correlations between disengagement latencies and fixation durations

Disengagement latencies significantly correlated with mean fixation durations in naturalistic videos at visit two ($r(16)$ = .583, p = .018), and with abstract videos at visit three ($r(18)$ = .488, p = .040). See all correlations in Table 6-15.

Table 6-15 Correlations between disengagement latencies and mean fixation durations.

	Visit 1	Visit 2	Visit 3	Visit 4
Naturalistic x disengagement	$r(18) = .198, p = .431$	$r(16) = .583, p = .018$	$r(18) = .320, p = .196$	$r(19) = .061, p = .805$
Abstract x disengagement	$r(18) = .379, p = .121$	$r(16) = .398, p = .127$	$r(18) = .488, p = .040$	$r(19) = .314, p = .190$
Static x disengagement	$r(18) = .422, p = .081$	$r(16) = .005, p = .986$	$r(18) = .310, p = .210$	$r(19) = .089, p = .716$

6.8 Study 3: Adults

6.8.1 Methods

6.8.1.1 Participants

Twenty adults participated for payment (ten male; mean age 26.7 years, range 18 - 38). All participants had normal or corrected to normal vision and were naive with respect to the purposes of the study. The study protocol was approved by the Birkbeck, Psychological Sciences Ethics Committee.

6.8.1.1 Apparatus, stimuli, and gaze-contingent paradigms

The apparatus was identical to the one described in Chapter 4. The stimuli, and the gaze contingent paradigms, were identical to what was described in sections 6.6.1.3, and 6.6.1.4 respectively.

6.8.1.2 Design and procedure

Participants were told to look at the screen freely without any instructions. They were tested individually in a small-darkened room while sitting on a chair located 60 cm away from the monitor. Participants were calibrated with the same 5 point infant calibration described in Chapter 2 and section 6.6.1.5. The calibration was repeated until the tracking was satisfactory. The study design was the same described in section 6.6.1.5.

6.8.2 Analysis

6.8.2.1 Fixation detection and coding

The fixation detection method and coding was identical to the one described in section 6.6.2.1

6.8.2.2 Data- quality

As in sections 6.6.2.2 and 6.7.2.2, spatial precision was measured for each participant by calculating the RMS. The mean RMS was .14° (SD = .03).

A correlational analysis between RMS-s and fixation durations was run in order to ensure data quality was not interfering with the fixation detection method. No significant correlations were found between data quality and spatial precision.

6.8.2.3 Cross-validation

6.8.2.3.1 Fixation durations

The process of cross-validating fixations was the same described in section 6.6.2.3.1. In total, the external coder coded 20 % of the data from this study.

The inter-rater reliability between the means and the number of detected fixations was evaluated using the ICC. A strong agreement between the mean fixation durations was found (with an ICC of .871, p = .04). Additionally, there was also an agreement in the number of fixations detected (with an ICC of .761, p = .005).

6.8.2.3.2 Gap-overlap

Adult participants were instructed to look at the screen during the experiment, thus all the trials stated when the participant looked at the central stimulus and ended when they looked at the peripheral one. For this reason trials were not manually reviewed as in the previous studies with infants.

6.8.2.4 Statistical analysis

As in the studies described earlier, the mean fixation duration and the ex-Gaussian components for each participant and viewing condition were calculated after excluding all fixations with a duration that was two standard deviations above or below the initial mean. Hence, the fixations that were included in the analysis were those that accounted for about the 95 % of the set for each participant. This exclusion was performed in order to exclude very long or very short fixations on a participant basis rather than establishing a minimum and maximum fixation duration threshold for all the participants together.

The gap-overlap trials with latencies shorter than 80 ms were excluded from the analysis since latencies below this value are considered to be anticipatory or express saccades (Fischer & Weber, 1993). As in the studies described earlier, all the trials that were one standard deviation above the mean latency were excluded.

The differences in fixation durations across viewing conditions were analysed using one-way ANOVAs. In the cases were the correlation matrices between the different measurement points were heterogeneous the degrees of freedom of the *F* test were corrected according to the Greenhouse-Geisser method. Post-hoc pairwise comparisons were calculated based on Bonferroni correction.

For the correlational analysis, a logarithmic transformation was applied to all the values in the data set. The justification for this transformation has rested on the fact that all these distributions are substantially skewed and the sample size is rather moderate. Thus, the logarithmic transformation is used to induce symmetry on the data.

6.8.3 Results

6.8.3.1 Differences in fixation durations across viewing conditions

A one-way ANOVA revealed a significant effect of stimulus type on mean fixation durations ($F(1.245, 26.646) = 94.771$, $p < .001$, $p\eta^2 = .833$). Post-hoc pairwise comparisons showed significant differences between all viewing conditions (Naturalistic: $M = 318$ ms, $SD = .009$; abstract: $M = 441$ ms, $SD = .020$; static: $M = 252$ ms, $SD = .007$; all $p < .001$; See Figure 6-11).
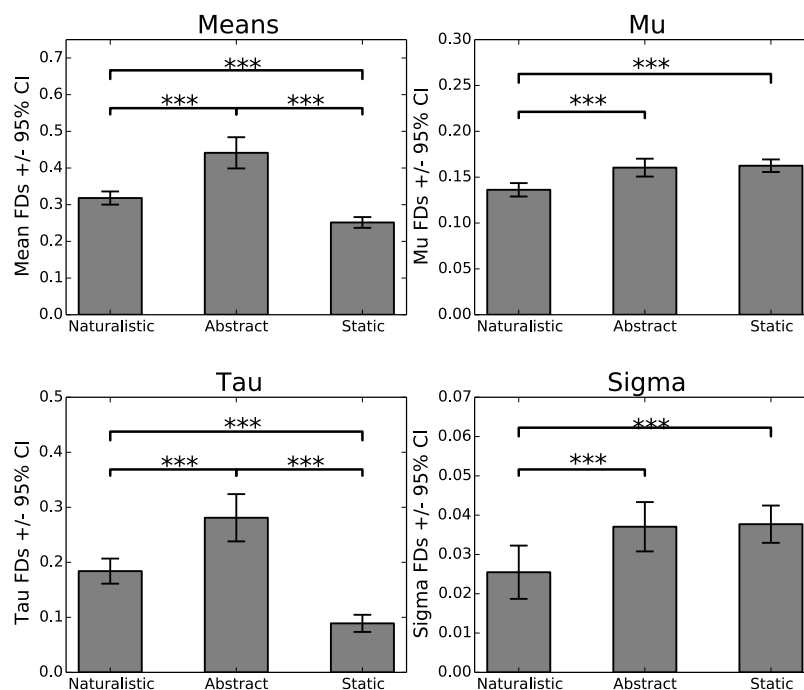
Figure 6-11 Adults study: Mean fixation durations and the ex-Gaussian components for all viewing conditions in adults.

In order to investigate and explain the origin of the effects described above further analysis were performed using the ex-Gaussian components (μ, σ, and τ) for the fixation duration distributions instead of the means. A one-way ANOVA revealed a significant effect of stimulus type on μ fixation durations ($F_{(2, 38)}$= 21.727, p < .001, $p\eta^2$= .533). Pairwise comparisons showed significant differences between naturalistic (M = 136 ms, SD = .003) and abstract videos (M = 160 ms, SD = .005; p < .001), naturalistic videos and static images (M = 162 ms, SD = .003; p < .001), but not between abstract videos and static images.

A one-way ANOVA showed a significant effect on σ fixation durations ($F_{(2, 38)}$= 7.616, p = .002, $p\eta^2$= .286). As for μ fixation durations analysis, pairwise comparisons showed significant differences between naturalistic (M = 25 ms, SD = .003) and abstract videos (M = 37 ms, SD = .003; p = .014), naturalistic videos and static images (M = 38 ms, SD = .002; p < .007), but not between abstract videos and static images.

Results analysing the differences in τ fixation durations resembled those observed for the means. A one-way ANOVA showed a significant effect on τ fixation durations ($F_{(1.454, 27.624)}$= 74.909, p < .001, $p\eta^2$= .798). Post-hoc pairwise comparisons showed significant differences between all viewing conditions (Naturalistic: M = 184 ms, SD = .011; abstract: M = 281 ms, SD = .021; static: M = 89 ms, SD = .007; all p < .001).

6.8.3.2 Individual differences across viewing conditions

Mean fixation durations for all viewing conditions correlated with each other (see Table 6-16). Furthermore, pairwise comparisons showed significant differences between all viewing conditions. These results evidence individual differences in mean fixation durations between short and long lookers.

Table 6-16 This table shows the correlations and pairwise comparisons between viewing conditions for adults. All t-tests have been corrected with Bonferroni correction.

	Mean naturalistic X Abstract	Mean naturalistic X Static	Mean abstract X Static
Pearson Corr.	$r_{(20)}$ = .823, p < .001	$r_{(20)}$ = .583, p = .007	$r_{(20)}$ = .553, p = .012
T-test	$t_{(19)}$ = -8.614, p < .001	$t_{(19)}$ = 8.638, p < .001	$t_{(19)}$ = 10.560, p < .001

6.8.3.3 The influence of disengagement in fixation durations

A t-test showed a gap-overlap effect on the adults group ($t(19) = 3.978$, $p = .001$). No correlations were found between disengagement latencies and mean fixation durations for any of the viewing conditions (see Table 6-17).

Table 6-17 Correlations between disengagement latencies and mean fixation durations for adults.

	Adults
Naturalistic x disengagement	$r(20) = .222$, $p = .347$
Abstract x disengagement	$r(20) = .334$, $p = .150$
Static x disengagement	$r(20) = -.117$, $p = .622$

6.8.3.4 Comparing adult and infant fixation durations

As the results from the three studies described in this chapter have shown, whilst generally fixation durations tended to decrease during the first year of life, this decrease did not happen at the same rate for all viewing conditions. Furthermore, the ex-Gaussian component that seemed to resemble mean fixation duration results best was the exponential τ component, suggesting that it is the variation in the tail of the distribution that is mostly responsible for the mean fixation durations change across viewing conditions. In this section I analyse whether infant fixation durations had reached adult levels at 12 months.

In this section, I compared mean fixation durations in 12-month-olds and adults. Results showed age differences in mean fixation durations ($F(1.163, 43.026) = 166.300$, $p < .001$, $p\eta^2 = .818$) as well as a stimulus and age interaction ($F(1.163, 43.026) = 17.060$, $p < .001$, $p\eta^2 = .316$). Pairwise comparisons revealed significant differences between both age groups for all viewing conditions (all $p < .001$). Similarly, the ex-Gaussian components µ ($F(1.682, 62.223) = 11.512$, $p < .001$, $p\eta^2 = .237$), σ ($F(2,74) = 6.896$, $p = .002$, $p\eta^2 = .157$), and τ ($F(1.229,45.482) = 135.313$, $p < .001$, $p\eta^2 = .785$) were shorter for adults as compared to 12-month-olds. There was a marginally significant interaction between µ fixation durations and stimulus type ($F(1.682, 62.223) = 2.759$, $p = .080$, $p\eta^2 = .069$). Significant interaction effects were found for σ fixation durations ($F(2,74) = 3.768$, $p = .028$, $p\eta^2 = .092$) and τ fixation durations ($F(1.229,45.482) = 17.512$, $p < .001$, $p\eta^2 = .321$). Pairwise comparisons showed a significant decrease for all ex-

Gaussian components in all viewing conditions (all p < .050), with the exception of the σ

component for shapes and static images (see Figure 6-12).
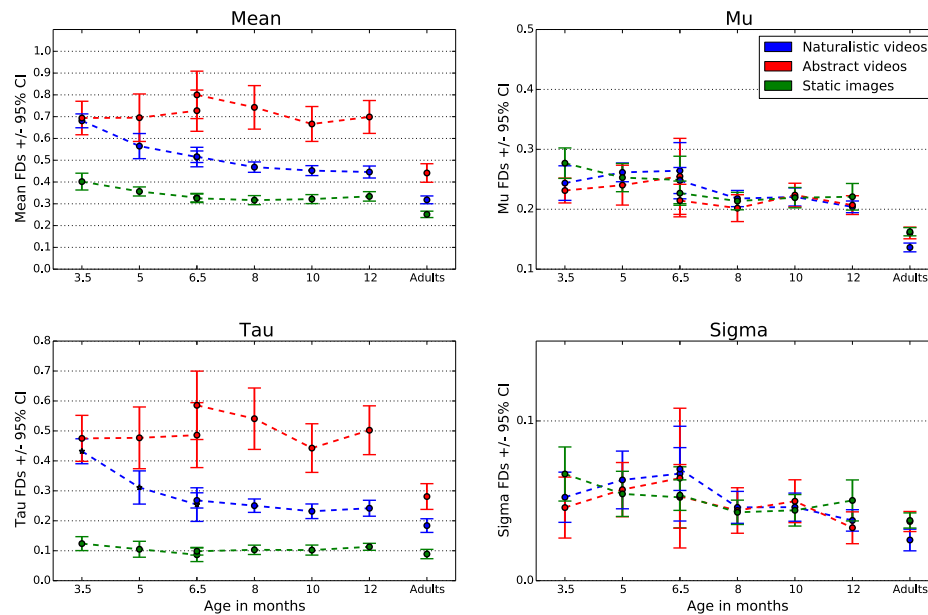
Figure 6-12 Longitudinal trajectories of fixation durations (means and ex-Gaussian components) for the two longitudinal studies and adults.

6.9 Discussion

The development of look duration across infancy and toddlerhood has been widely investigated

in a number of studies. While many of these studies have reported a decrease in look duration

during the first year of life (Colombo & Mitchell, 1990; Kagan et al., 1971; M. Lewis et al., 1969;

Mayes & Kessen, 1989; Shaddy & Colombo, 2004), others found different trends in look

duration trajectories for different stimulus types (Courage et al., 2006; Frank et al., 2014, 2009;

Shaddy & Colombo, 2004). Results from the second group of studies seemed to indicate that

more complex forms of visual stimuli elicited an increase in looking times during the second half

of the first year. Furthermore, looking times have also been the focus of many studies

investigating individual differences in infancy and childhood (Colombo et al., 1995, 1991;

Colombo & Mitchell, 1990; Colombo, 1995; Cuevas & Bell, 2013; Jankowski & Rose, 1997).

In contrast, little research has been concerned with fixation duration trajectories in infancy or on

individual differences in such measures (but see Bronson, 1994; Hunnius & Geuze, 2004a;

Papageorgiou et al., 2014; Wass & Smith, 2014). Such investigations are not only important to understand the development of saccadic control in infancy, but also to study the mechanisms underlying fixation durations and visual processing during the first years of life. Furthermore, fixation durations in infancy have been proven to be a useful tool to investigate continuity in infancy and childhood (Papageorgiou et al., 2014).

In this chapter I have presented two longitudinal studies and a study with adult participants that aimed to investigate the development of fixation durations and saccadic control over the first year of life when viewing a battery of dynamic and static stimuli and performing the gap-overlap paradigm. Furthermore, the data from these studies revealed individual differences across viewing conditions and visits. In this section I will review and discuss the results from these studies.

6.9.1 Fixation duration trajectories for dynamic and static complex scenes

With the studies presented in this chapter I aimed to describe fixation duration trajectories for different types of stimuli during the first and the second half of the first year, and analyse the differences in the processing of dynamic and static stimuli during this period.

Results showed an overall decrease in mean fixation durations over the first and the second half of the first year of life. Nevertheless, different viewing conditions showed different trajectories in mean fixation durations, evidencing the influence of cognitive and visual processing in saccadic control even in the group with the youngest infants (3.5 months). During the first half of the first year, results showed a significant decrease in mean fixation durations for naturalistic videos and static images, but not for the abstract videos that remained fairly stable. From 6 months on, mean fixation durations continued to decrease for naturalistic videos (especially from 6 to 10 months), as well as for abstract videos (from 8 to 10 months). In contrast, during the same period mean fixation durations for static stimuli seemed to have reached certain stability and did not decrease. Interestingly, from 10 to 12 months mean fixation durations did not decrease nor increase for any of the viewing conditions. The comparison between mean fixation durations for the 12-month-olds and adults revealed significant differences for all the viewing conditions, confirming that saccadic control and fixation durations still did not reach adult levels at the end of the first year.

The decrease in mean fixation durations during the first half of the first year resemble the findings from studies investigating looking times during the same period (Colombo & Mitchell, 1990; Kagan et al., 1971; M. Lewis et al., 1969; Mayes & Kessen, 1989; Shaddy & Colombo, 2004). The significant decrease in disengagement latencies and reaction times that I found during this period (these results are discussed later in this chapter) suggests the decrease in mean fixation durations can -to an extent- be explained by the improvement in the disengagement of attention that happens as a result of the development of the frontal cortical structures concerned with saccade planning and execution (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011).

Some studies investigating looking times during the second half of the first year of life found that for more complex forms of visual intake (e.g., complex dynamic stimuli, looking at faces, playing with multiple toys) looking times tend to increase (Courage et al., 2006; Frank et al., 2014, 2009; Shaddy & Colombo, 2004). The results presented in this chapter did not show any increase for any of the viewing conditions, even though the stimuli presented was fairly complex, both in terms of the semantic and social content and the low level visual features of the scene such as colour or luminance. This supports the claim that studies using looking times may not be measuring the same processes as those using fixation durations. While looking times can be indicative of selective attention, measuring fixation durations may add a level of precision in the analysis of the processes underlying eye-movements. Nevertheless the present study was not designed to specifically look at the differences between these measures thus these conclusions are still speculative.

The decrease in fixation durations during the second half of the first year could be due to the advances in memory functions –in either encoding, storage, retrieval, or some combination of these processes- that occur during this period (Baillargeon & Graber, 1988; Cohen & Gelber, 1975; Colombo, 1993; Richmond & Nelson, 2009; Ross-sheehy, Oakes, & Luck, 2003; Ruff & Rothbart, 1996). For instance, short-term memory, which appears to be strongly dependent on the functioning of regions on the prefrontal cortex (Miller, Erickson, & Desimone, 1996; Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000; E. E. Smith & Jonides, 1997), suffers a major development during the second half of the first year (Ross-sheehy et al., 2003) presumably as a consequence of the dramatic development of prefrontal functioning that happens during this period (Bell, 1998; M. H. Johnson, 1995a; Nelson, 1995).

Another factor that may explain the decrease in fixation durations during the first and the second half of the first year is the improvement in processing efficiency (Colombo et al., 1991). Infants get more and more experience processing visual information and this may undeniably improve their capacity for visual and cognitive encoding.

Interestingly, fixation durations in the abstract videos did not suffer any substantial change during the first half of the year, which, as suggested in previous studies that also used abstract videos of faces and found similar results (Hunnius & Geuze, 2004a), may be due to the novelty of the stimuli. During the first half of the first year infants gain increased experience at looking at faces, and as a consequence they may need less time to process social content as they get older. This can also explain the decrease in fixation durations for the naturalistic and the static conditions during this period, since both conditions display social scenes. On the other hand, the abstract condition remains novel across visits. During the second half of the first year there was a decrease in mean fixation durations for the abstract condition, more particularly from 8 to 10 months. This coincides with the advances in memory capacity that occur during this period (Baillargeon & Graber, 1988; Cohen & Gelber, 1975; Colombo, 1993; Richmond & Nelson, 2009; Ross-sheehy et al., 2003; Ruff & Rothbart, 1996).

Fixation duration distributions have been reported to be highly skewed in adults and more particularly in infants (Harris et al., 1988). Because of this reason, reporting other measures beyond the means –such as the ex-Gaussian components- can add an extra level of precision and help disentangling some of the effects that the experimental manipulations have on the data. For the studies presented in this chapter I calculated the ex-Gaussian components (μ, σ, and τ) for each fixation duration distribution (Lacouture & Cousineau, 2008). The most obvious finding is that the main component that modulates mean fixation durations is the exponential τ component. Higher τ values represent more positively skewed distributions and hence indicate more variability in the data. This suggests that the differences in mean fixation durations are mainly a consequence of the variability in the distributions, and hence infants flagged as short or long lookers are not necessarily the ones with shorter or longer fixations, but those with more or less variability in their fixation duration distributions. For instance, a distribution that has a low μ and high τ, indicate that while the infant was able to reliably move her/his eyes relatively fast, he/she still presented high variability in fixation durations, maybe due to the stimulus demands or other factors.

During the first half of the first year μ fixation durations –the mean from the Gaussian fitting, representing the peak of the distribution- only decreased for the static condition, while no age effect was found for any of the dynamic conditions. The σ component –the standard deviation of the Gaussian fitting, representing the variation or dispersion from the average- showed similar – but not statistically significant- trends. These results suggest that fixation durations are stabilizing for the static condition during this period. From 6 months onwards, although there was an overall decrease in μ and σ fixation durations, these components only decreased for the naturalistic condition from 6 to 8 months and remained fairly stable after this.

Interestingly, the LMM showed an almost significant increase in τ fixation durations for static images during the second half of the first year. An explanation for this effect could be the emergence of very elementary executive functions (Cuevas & Bell, 2013). Some of the studies that found an increase in looking times at about 13 months have speculated that this increase could be a consequence of the emergence of executive functions at the end of the first year (Kagan et al., 1971; M. Lewis et al., 1969; Ruff & Rothbart, 1996). Nevertheless, some recent studies have also analysed the emergence of elementary executive functions during the second half of the first year (Cuevas & Bell, 2013), which could also explain the slight increase in mean fixation durations in static images and not dynamic scenes, where information processing is more complex and eye-movements take longer to stabilize (Hunnius & Geuze, 2004a).

The differences in fixation durations between the three viewing conditions were also analysed at each visit. At 3.5 months, infants showed significantly shorter mean fixation durations for the static condition as compared to the two dynamic conditions (naturalistic and abstract videos). In turn, at this age no differences were found between the means or ex-Gaussian components from these two dynamic viewing conditions, which displayed the same low-level visual features but different semantic and social content. Infants from 5 months on and adults presented differences in mean fixation durations between all viewing conditions. The fact that at 3.5 months infants could only differentiate between viewing conditions presenting different low-level visual features evidences how at this age infant saccadic control is mainly driven by bottom-up factors such as colour or luminance. As they gain more top-down control they start being able to differentiate between the two dynamic conditions containing the same bottom-up information (from the 5 months visit).

Furthermore, mean fixation durations for static images were significantly shorter than for the other dynamic conditions, both for infants –in all the visits- and adults. Visual information is processed through several neural pathways that become functional at different developmental stages (see Chapter 1; Atkinson, 2000; M. H. Johnson, 2011). The cortical pathway that goes to the superior colliculus from the primary visual cortex (V1) through the middle temporal area (MT), is thought to be involved in motion detection and the smooth tracking of moving stimuli (M. H. Johnson, 1990, 2011; Schiller, 1985, 1998) and is already functioning at about 2 months of age (M. H. Johnson, 1990, 2011). Dynamic visual information is thus also processed through the MT pathway –which presumably is not involved in the processing of static visual information-, leading to prolonged fixation durations. Additionally, a dynamic scene may contain more semantic content that needs to be processed -such as the understanding of the actions that the different characters are performing- that can prolong fixation durations.

From 5 months on mean fixation durations for the abstract videos were significantly longer than for the remaining viewing conditions. Besides the novelty effect that I have discussed earlier, prolonged fixation durations in the abstract condition can also be explained as a result of the higher demands in cognitive and perceptual processing that a novel stimuli may elicit (see discussion from Chapter 4).

6.9.2 Individual differences in fixation durations over the first year of life

A number of studies have investigated individual differences in looking times (Colombo et al., 1995, 1991; Colombo & Mitchell, 1990; Colombo, 1995; Cuevas & Bell, 2013; Jankowski & Rose, 1997) and their use as a predictor for later cognitive and intellectual function in childhood and adolescence (see Chapters 1 and 4; Choi & Vaswani, 2014; Colombo et al., 1995; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997; Kavšek, 2004; Lawson & Ruff, 2004; Sigman et al., 1986; A Slater, 1995). Nevertheless, few studies have analysed individual differences in infant fixation durations (but see Papageorgiou et al., 2014; Wass & Smith, 2014).

The studies presented in this chapter also served to explore individual differences in fixation durations across visits and stimulus types. In the first longitudinal study individual differences were found across all the visits (3.5, 5 and 6 months) for almost all the viewing conditions. Correlations were not found, however, for mean fixation durations in naturalistic videos between the 3.5 and 5 months visits. Interestingly, abstract videos displayed the stronger correlations

across the three different assessments, followed by static images. In the second longitudinal study individual differences were found across all the visits (6, 8, 10 and 12 months) for all viewing conditions, showing a great stability in fixation durations during the second half of the first year. During this period, static images displayed the strongest correlations between the four different assessments. These results evidence the increase in top-down control from 3.5 to 5 months and demonstrate certain stability in fixation durations from a very early age.

Individual differences were also analysed for each visit by looking at the stability across viewing conditions. Even though correlations were found in all visits, not all viewing conditions correlated with each other in every visit. At 3.5 months fixation durations in naturalistic videos did not correlate with the other two viewing conditions, whereas there was a correlation between fixation durations in abstract videos and static images. As described earlier, the correlations for the naturalistic condition between 3.5 and 5 months were not significant. Looking at the graphs from Figure 6-7 it is possible to identify two different trends in the naturalistic condition at this age: while mean fixation durations decreased from 3.5 to 5 months for some of the infants, they remained stable for others. As I have explained earlier, during this period the frontal neural structures concerned with saccade planning and execution are developing rapidly (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), thus it is not surprising to find more variability in fixation durations at this transitional age. On the other hand, results suggest a period of stability at 5 and 6 months, where high correlations were found between all viewing conditions, with the exception of abstract videos and static images at 6 months.

Results from the second longitudinal study revealed few significant correlations at 6 months (naturalistic x abstract), 10 months, (naturalistic x abstract) and 12 months (naturalistic x static), and marginally significant correlations at 6 months (naturalistic x static), 8 months (naturalistic x abstract), and 10 months (naturalistic x static). Therefore, these results indicate individual differences in mean fixation durations for some of the viewing conditions and visits, but not for all, suggesting that saccadic control and fixation durations are still undergoing important changes during the first year of life and still have not reached stability. In the third study with adult participants strong correlations were found between fixation durations for all viewing conditions.

6.9.3 Reaction times and disengagement latencies over the first year of life

In this chapter I also explored the development of reaction times and disengagement latencies during the first year of life and analysed their relationship with fixation durations during this period.

Not surprisingly, as in previous studies using the gap-overlap paradigm (Atkinson et al., 1992; Blaga & Colombo, 2006; Butcher et al., 2000; Domsch et al., 2010; Elsabbagh et al., 2009, 2013; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991; Kano et al., 2011; Kano & Tomonaga, 2011b; Kikuchi et al., 2011; Kopecz, 1995; Reuter-Lorenz et al., 1991) the gap effect was found for all visits in infants and adults.

During the first half of the first year disengagement latencies diminished with age. Furthermore, reaction times for the gap and overlap conditions also decreased during this period. These findings go in line with previous research reporting improvements in the ability to disengage (e.g., Butcher et al., 2000; Matsuzawa & Shimojo, 1997) during the first half of the first year as a result of the development of the frontal structures concerned with saccade planning and execution (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011). On the other hand, during the second half of the first year results did not show any age effect in disengagement latencies, even though there was a decrease in reaction times for the gap condition. This suggests that while their ability to disengage already reached stability during the first half of the first year saccadic programming and execution still continued to increase in efficiency from 6 months on.

The correlations between mean fixation durations and disengagement latencies were not very stable for any of the visits –including adults-, even though there were a few positive correlations for some of the viewing conditions and visits. More precisely, there were no significant correlations between disengagement latencies and mean fixation durations at 3.5 and 6 months (first longitudinal study), even though some of the correlations were close to significant. However, at 5 months disengagement latencies correlated with mean fixation durations for naturalistic videos and static images. In the second longitudinal study disengagement latencies significantly correlated with mean fixation durations in naturalistic videos at 8 months and with abstract videos at 10 months. In adults there were no correlations between disengagement latencies and fixation durations for any of the conditions. These results are rather surprising and

contradict the findings reported in the cross-sectional study from Chapter 4, where correlations were found for 6-month-olds and adults. The main difference between the present studies and the cross-sectional study from Chapter 4 is the double-step paradigm, which was removed from the design of the studies described in the present chapter. Having the gap-overlap and the double-step in the same design may have made the targets less predictable than they were for the present studies, where infants and adults were most likely predicting the appearance of the next target.

6.10 Conclusions

In sum, in the present chapter I presented two longitudinal studies and a study with adult participants with the view to investigate the development of fixation durations and saccadic control over the first year of life when viewing a battery of dynamic and static stimuli and performing the gap-overlap paradigm. With these studies I aimed to describe fixation duration trajectories for different types of stimuli (dynamic and static) during the first year of life as well as to explore individual differences in fixation durations across visits and stimulus types. Additionally, I investigated the relationship between fixation durations and disengagement latencies during the first year.

Results showed an overall decrease in mean fixation durations even though different viewing conditions decreased at different rates and times. This decrease could be explained by an improvement in processing efficiency (Colombo et al., 1991), the improvements in disengagement abilities during the first 6 months (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), and/or the advances in memory capacity that occur during the second half of the first year (Colombo, 1993). Unlike some of the studies investigating the development of looking times when viewing more complex forms of visual stimuli (Courage et al., 2006; Frank et al., 2014, 2009; Shaddy & Colombo, 2004), mean fixation durations did not increase for any of the viewing conditions. Results also showed individual differences across visits and viewing conditions, even though they did not appear to be as stable as the ones found in adults.

In the next chapter I will use the CRISP computational model (Nuthmann et al., 2010) to model the present empirical data in order to investigate the development of the mechanisms underlying fixation durations and saccadic control during the first year of life. More specifically,

the simulations from the next chapter are directed to investigate the development of saccade timing and saccade programming in young infants. Additionally, I include a simulation study that investigates different developmental trajectories in saccade timing and saccade programming for infants presenting long and short mean fixation durations.

Chapter 7: Modelling longitudinal fixation duration trajectories

7.1 Introduction

In Chapter 5 I investigated the unexplored mechanisms underlying saccadic control in infancy by determining the generalizability and potential limitations of extending the CRISP theoretical framework and computational model of fixation durations in adult scene-viewing (Nuthmann et al., 2010) to infants. The CRISP model was used to investigate the underlying mechanisms modulating fixation durations in 6-month-old infants and adults (see also Chapter 4). As this generalizability was confirmed, it was implied that at 6 months saccades are also programmed in two stages, an initial labile phase during which saccade programs can be altered or cancelled and a subsequent non-labile phase in which programs cannot be cancelled. In addition, the CRISP model was used to examine the extent to which fixation durations at 6 months are affected by developmental aspects of the visuo-motor system and/or by visual and cognitive processing, with results suggesting that the main factor affecting infant saccadic control at this age may already be visual and cognitive processing, even though the developmental state of the visuo-motor system still appears to be affecting infants' saccadic control. Nevertheless, due to the cross-sectional nature of the design these results only provide insight into a particular point in time and hence the developmental changes that both saccade timing and saccade programming undergo over the first year of life had to be hypothesized. With the view to investigate the development of fixation durations and saccadic control over the first year of life, in Chapter 6 I presented two longitudinal studies (the first group of infants was tested at approximately 3.5, 5 and 6.5 months, and the second at 6, 8, 10 and 12 months) and a study with adult participants. In these studies eye-movements were monitored while participants looked at dynamic and static complex stimuli, and performed the gap-overlap paradigm. In each assessment infants and adults were presented with: (1) a set of naturalistic videos in which three people performed baby-friendly actions, (2) a set of abstract videos created from the first set, and (3) static complex images.

The goal of the present chapter is to investigate the development of saccade timing and saccade programming parameters during the first year of life by modelling fixation data from the longitudinal studies presented in Chapter 6. Additionally, this chapter includes a simulation study that investigates different developmental trajectories for saccade timing and saccade programming parameters in two infants presenting long and short mean fixation durations.

For a literature review on the development of saccadic control and fixation durations in infancy go to Chapters 1, 4, 5 and 6.

7.2 Modelling longitudinal data with the CRISP model

The CRISP model architecture and theoretical assumptions were described in Chapter 5. In order to revisit the CRISP model architecture, in the following lines I include the model's three main assumptions on saccade timing and programming:

(1) The accumulation of activity to a saccade threshold is implemented via a random walk process and is responsible for generating inter-saccadic intervals and hence variations in fixation durations.

(2) Saccade programming occurs in two stages: an initial, labile stage that is subject to cancellation, and a later, non-labile stage.

(3) Processing difficulty can inhibit saccade timing and programming in a moment-to-moment fashion. The latter can happen in two ways. First, in case of increased processing demands the random walk saccade timer slows down, which delays the initiation of a new saccade program and eventually leads to longer fixation durations. Secondly, processing difficulties can even cancel an ongoing labile saccade program, which extends the duration of the current fixation.

The following CRISP parameters contribute toward systematic differences in fixation durations: (1) the mean (t_{sac}) and variance (i.e., number of states N) of the random walk timing signal, (2) the mean duration of the labile saccade program (τ_{lab}), and (3) the mean duration of the non-labile saccade program (τ_{nlab}) (see Nuthmann et al., 2010).

In Chapter 5, some of the aspects related to the development of saccade timing and saccade programming parameters had to be hypothesized due to the cross-sectional nature of the study. In the present chapter I aim to explore and confirm these hypothesis:

    a)   *Random walk timing and speed of processing.*

In Chapter 6 results showed an overall decrease in mean fixation durations over the first year of life, especially during the first half of the year.

In CRISP, decreases in processing speed and increases in processing difficulty will slow down the random walk saccade timer. This delays the initiation of a new saccade program, which in

turn leads to longer fixation durations. As I argued in Chapter 5, saccade timing parameters can be indicative of information processing and thus I assumed that as infants get older they would show a decrease in the mean values for the random walk timer. Whilst the differences in saccade timing parameters between 6-month-olds and adults were evident, this assumption could not be tested with a cross-sectional design. In this chapter I will explore whether as expected, the mean values for the random walk timer decrease –and hence processing efficiency increase- across the first year of life, and more particularly during the second half of the first year as infants' memory capacity develops (Baillargeon & Graber, 1988; Cohen & Gelber, 1975; Colombo, 1993; Richmond & Nelson, 2009; Ross-sheehy et al., 2003; Ruff & Rothbart, 1996).

b) Two-stage saccade programming.

In CRISP, saccade programming is completed in two stages: an initial, labile stage that is subject to cancellation and a subsequent non-labile stage in which the program can no longer be cancelled (see Chapters 1,4 and 5). The notion of the two-stage saccade programming is motivated by findings from double-step experiments (e.g., Becker & Jurgens, 1979).

While the two-step notion of saccade programming has never been specifically investigated in behavioural studies with infants, results from the simulation studies from Chapter 5 suggested that as for adults, 6-month-old infants may also program saccades in two consecutive stages. Nevertheless the development of the saccade programming parameters (the labile and the non-labile stages) had to be hypothesized based on previous findings from the anti-saccade paradigm (e.g., M. H. Johnson, 1995; see Chapter 1), which suggested that by 4 months infants can reliably inhibit the saccade to a first target in order to move their eyes to a second more attractive target. This ability to inhibit a saccade program coincides with the development of the premotor areas of the frontal lobes that occurs from 3 to 4 months (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), and suggests that by this age infants should be able to cancel a saccade program, and hence program saccades in two stages (see also findings from Chapter 4 on saccade cancellation at 6 months).

Since infants' ability to cancel a saccade may not be efficient enough before 4 months, it could be the case that the non-labile program is longer than the labile program before this age. In Chapter 5 I predicted that from 4 months onwards there would be (1) a relative increase of the

labile stage with respect to the non-labile stage program, and (2) a gradual decrease in the absolute times for both labile and non-labile programs as the infant's saccadic control increases in efficiency. Additionally, research has shown that by 6 months infant saccadic control has not reached adult levels (Butcher, Kalverboer, & Geuze, 2000; Csibra, Tucker, & Johnson, 2001; see also Chapter 4), thus both labile and non-labile stages are likely to be longer at this age compared to adults.

The simulation studies presented in this chapter will serve to test these assumptions.

7.3 The current studies

In the present chapter I describe three simulation studies with the CRISP model (Nuthmann et al., 2010) in which fixation data from Chapter 6 were modelled. The overarching goal was to explore the trajectories for saccade timing and saccade programming parameters over the first year of life and to investigate how these trajectories would differ for short and long lookers.

In simulation study 1, fixations from the naturalistic condition from the first longitudinal study described in Chapter 6 were modelled. In this study, infants from 3 to 6 months were eye-tracked in three consecutive visits (3.5, 5 and 6.5 months) while looking at dynamic and static complex scenes and performing the gap-overlap paradigm. This simulation study aimed to (a) explore the development of saccade timing and saccade programming parameters during the first half of the first year, a period during which the frontal structures concerned with saccade programming and execution are developing rapidly (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011). In simulation study 2 fixations from the naturalistic condition from the second longitudinal study described in Chapter 6 (with visits at 6, 8, 10 and 12 months) were modelled. The goal for this simulation study was to (b) explore the development of saccade timing and saccade programming during the second half of the first year. Finally, simulation study 3 aimed to investigate the trajectories for saccade timing and saccade programming parameters in infants classified as long and short lookers. As it was described earlier in this thesis (see Chapter 4), during the first year of life long lookers have been associated with greater disengagement difficulties (Frick et al., 1999; see also Blaga & Colombo, 2006) or poorer processing speed (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997), thus I expected to find a delayed trajectory in

saccade timing and/or saccade programming parameters for long lookers compared to short lookers.

7.4 Simulation study 1: Three to 6 months

The goal of the present simulation study was to investigate the developmental changes on the mechanisms underlying saccadic control in a group of infants that was tested at 3.5, 5 and 6 months (see longitudinal study 1 from Chapter 6), a period during which infants go through major changes in areas of the brain involved in eye-movement generation (e.g., the frontal eye fields, FEF), consequently improving their ability to disengage from a focal point (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011). More specifically, these simulations served to investigate the development of saccade timing and saccade programming parameters during this developmental stage, and to test some of the hypothesis made in Chapter 5 where fixation durations from 6-month-olds and adults were modelled (see Introduction from this chapter).

7.4.1 Behavioural data

For the current simulation study, fixation data from the first longitudinal study described in Chapter 6 were modelled. In this study, a group of infants was eye-tracked at 3.5, 5 and 6.5 months (every 6 weeks) while looking at a battery of dynamic and static complex scenes and performed the gap-overlap paradigm. With the view to reduce the number of free parameters in the model simulations only the naturalistic viewing condition was included in the present analysis.

The statistical analysis showed an overall decrease in mean fixation durations over the first half of the first year, with different viewing conditions displaying different fixation duration trajectories (see section 6.6.3.1). In particular, mean fixation durations in the naturalistic condition significantly decreased over this period. Furthermore, at the same time disengagement latencies as well as gap reaction times diminished with age, evidencing the maturational changes that the frontal structures responsible for saccade programming and execution undergo during this period (see section 6.6.3.3).

The empirical data set was comprised of 1920 fixations at 3.5 months, 3133 fixations at 5 months and 3210 fixations at 6.5 months. The differences in the number of fixations across

ages were mainly a consequence of the amount of time infants at different ages looked at the stimuli.

7.4.2 Model adjustments

The model adjustments and assumptions were exactly the same as those described in the simulation study 1 from Chapter 5.

Each model simulation replicates the results from 10 fictive participants thus it is possible to perform a statistical analysis on the simulated data. Nevertheless, since the CRISP model does not generate individual differences the variations between fictive participants are minimal and the significance of the statistical analysis does not necessarily explain the results. Still, the statistical analysis is used as a way of explaining the data in a more rigorous way than just commenting on the model parameters. The longitudinal variations in mean fixation durations and saccade programming parameters were analysed using repeated measures ANOVAs. In the cases where the correlation matrices between the different measurement points were heterogeneous the degrees of freedom of the F test were corrected according to the Greenhouse-Geisser method. Post-hoc pairwise comparisons were calculated based on Bonferroni correction.

7.4.3 Modelling results

The genetic algorithm optimization technique that was used in previous simulation studies with the CRISP model (simulation study 2 in Nuthmann & Henderson, 2012; Nuthmann et al., 2010; see also Chapter 5) was the procedure used to estimate free parameters. This technique estimates the input parameters minimizing a goodness-of-fit measure and evaluating how much the simulated data differs from the empirical data. For each visit, a total of four free parameters were fit: the mean durations for the labile and non-labile stages of saccade programming; and the mean and variance of the random walk timing signal for the naturalistic condition. The ranges for estimating these values were based on the same principles described in Chapter 5. Each visit was modelled independently. For more information about the particularities of the genetic algorithm optimization technique and the goodness-of-fit function used for these simulations see simulation study 1 from Chapter 5.

Figure 7-1 displays fixation duration distributions for the empirical data paired with the simulated data. Despite the very skewed fixation duration distributions –particularly at 3.5 months- the model seems to fit the empirical data well (see Figure 7-1 and errors from Table 7-1).
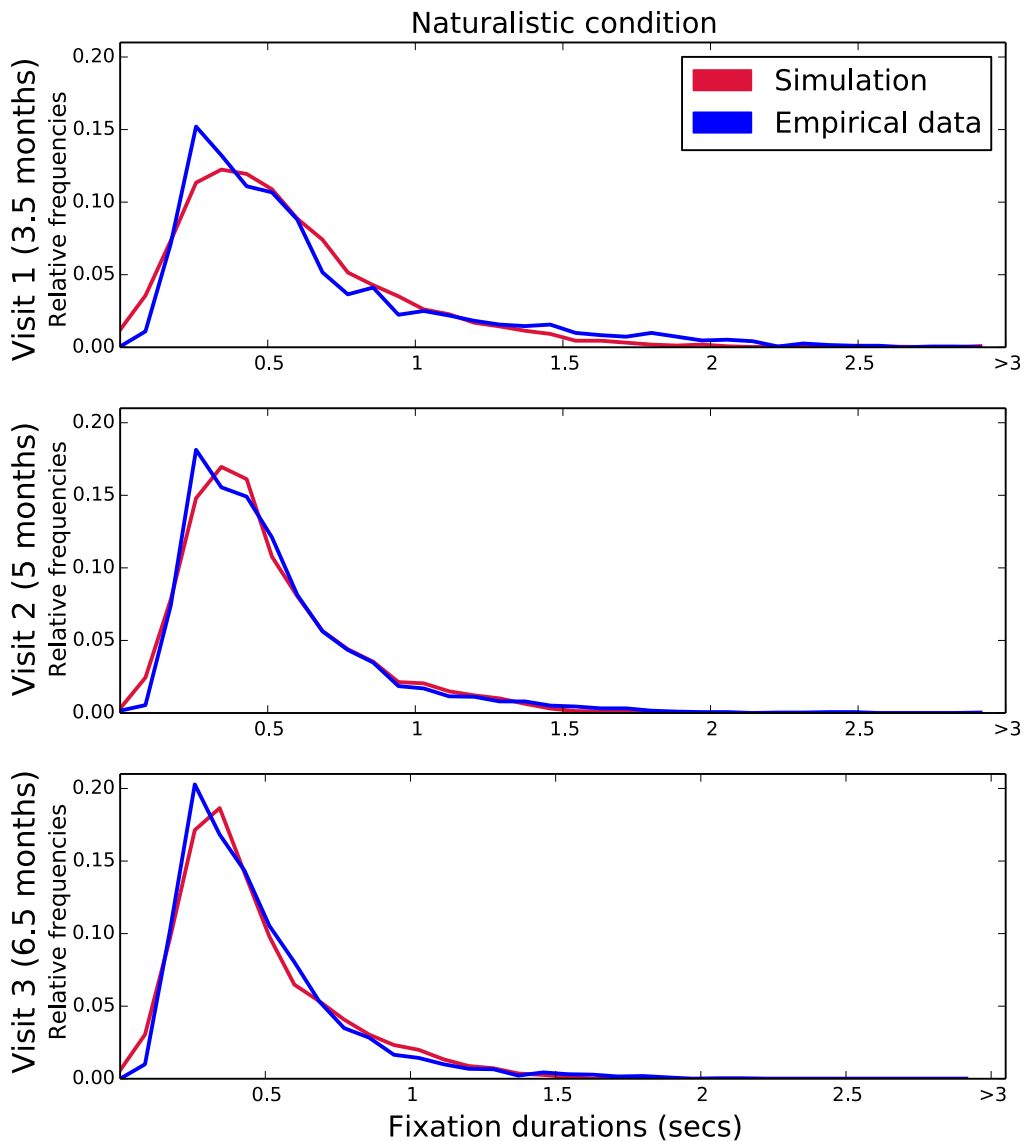
Figure 7-1 Simulation study 1. Simulations paired with empirical data for the naturalistic condition on the three visits.

The parameter values that best resembled the empirical distributions for the three visits are described in Table 7-1.

Table 7-1 Simulation study 1. Best fitting parameters.

Parameter	Symbol	Range	3.5 months	5 months	6.5 months
Saccade timing					
Mean (ms)	t_{sac}	200-650	443	440	506
Variance	N	5-20	20	20	18
Saccade execution	t_{ex}	---	50	50	50
Saccade programming					
Labile stage (ms)	t_{lab}	30-400	363	353	329
Non-labile stage (ms)	t_{nlab}	30-350	342	83	62
Standard deviation (ms)	σ	---	0.33	0.33	0.33
Error	E	---	9.620	2.644	2.054
Error cumulative distribution	E_c	---	10.284	2.809	3.195

The statistical analysis performed on the simulated data resembled the empirical findings. Mean fixation durations for the naturalistic condition decreased with age ($F(2,18)= 294.941$, $p < .001$, $p\eta^2= .970$) both from 3.5 ($M = 563$ ms, $SD = .004$) to 5 months ($M = 490$ ms, $SD = .002$; $p < .001$) and from 5 to 6.5 months ($M = 461$ ms, $SD = .003$; $p < .001$).

Additionally, age effects were statistically analysed for saccade programming parameters. This analysis was possible since for each realization of the model simulation –and hence, for each fixation-, parameter values for the different saccade programming phases (τ_{lab}, τ_{nlab}, and τ_{ex}) were drawn from gamma distributions, introducing unsystematic variability in the duration of saccade programming stages. As predicted, there was a gradual decrease in the absolute times for both labile and non-labile programs during this period. In particular, the labile stage decreased ($F(2,18)= 216.175$, $p < .001$, $p\eta^2= .960$) both from 3.5 ($M = 311$ ms, $SD = .590$) to 5 months ($M = 308$ ms, $SD = 1.101$; $p = .025$) and from 5 to 6.5 months ($M = 292$ ms, $SD = .670$; $p < .001$). Similarly, the non-labile stage also decreased ($F(2,18)= 27574.214$, $p < .001$, $p\eta^2= 1.000$) from 3.5 ($M = 335$ ms, $SD = 1.518$) to 5 months ($M = 83$ ms, $SD = .249$; $p < .001$) and from 5 to 6.5 months ($M = 62$ ms, $SD = .234$; $p < .001$). Moreover, further statistical analyses also showed a relative increase of the labile stage with respect to the non-labile stage program ($F(2,18)= 35103.135$, $p < .001$, $p\eta^2= 1.000$). Specifically, at 3.5 months the non-labile stage constituted 51.829 % ($SD = .125$) of the saccade program and it decreased to the 21.244 % ($SD = .077$) at 5 months ($p < .001$). Whilst not at the same rate, the proportion of the non-labile

stage with respect to the labile stage continued to decrease and at 6.5 months the non-labile stage constituted the 17.511 % (*SD* = .074) of the saccade program (*p* < .001).

Saccade timing parameters could not be statistically analysed the same way as saccade programming parameters. As explained earlier, in CRISP, decreases in processing speed and increases in processing difficulty will slow down the random walk saccade timer delaying the initiation of a new saccade program and hence leading to longer fixation durations. Saccade timing parameters are used to calculate the rate of the diffusion process for the random walk (see Chapter 5) and the number of steps needed for the random walk timing signal to progress and reach a threshold, where a new saccade program is initiated. This means that the model does not provide a saccade timing "value" per fixation, and thus these parameters need to be analysed by examining the best-fitting parameters from Table 7-1. Results showed how the mean values for the random walk timer did not change from 3.5 (443 ms) to 5 months (440 ms), but increased considerably from 5 to 6.5 months (506 ms).

7.5 Simulation study 2: Six to 12 months

The goal of the present simulation study was to investigate the developmental changes on the mechanisms underlying saccadic control in a group of infants that was tested at 6, 8, 10 and 12 months (see longitudinal study 2 from Chapter 6). More specifically, the present simulations investigated the development of saccade timing and saccade programming parameters during the second half of the first year by testing some of the hypothesis made in Chapter 5 where fixation durations from 6-month-olds and adults were modelled (see Introduction from this chapter).

7.5.1 Behavioural data

For the current simulation study, fixation data from the second longitudinal study described in Chapter 6 were modelled. As for the previous simulation study, only the naturalistic viewing condition was included in the analysis. The results from the behavioural studies showed an overall decrease in mean fixation durations over the second half of the first year of life. During this period mean fixation durations continued to decrease for naturalistic videos, especially from 6 to 10 months. From 10 to 12 months mean fixation durations did not change. The comparison between mean fixation durations for the 12-month-olds and adults revealed significant

differences for all the viewing conditions, suggesting that saccadic control and fixation durations still did not reach adult levels at the end of the first year.

The empirical data set was comprised of 5037 fixations at 6 months, 4738 fixations at 8 months, 5256 fixations at 10 months and 5447 fixations at 12 months. The differences in the number of fixations across ages were mainly a consequence of the amount of time infants at different ages looked at the stimuli.

7.5.2 Model adjustments

The model adjustments and assumptions were exactly the same as those described previously in this chapter.

7.5.3 Modelling results

The model's free parameters were estimated using the genetic algorithm optimization technique described earlier in this chapter (see simulation study 1, or the simulation studies from Chapter 5). As for the simulation study 1, for each visit a total of four free parameters were fit: the mean durations for the labile and non-labile stages of saccade programming; and the mean and variance of the random walk timing signal for the naturalistic condition. The ranges for estimating these values were based on the same principles described in Chapter 5. The parameter values that best resembled the empirical distributions for the four visits are described in Table 7-2. Additionally, Figure 7-2 displays fixation duration distributions for the empirical data paired with the simulated data.

Table 7-2 Simulation study 2. Best fitting parameters for the visit at 6 months.

| Parameter	Symbol	Range	6 months	8 months	10 months	12 months
Saccade timing						
Mean (ms)	t_{sac}	200-650	397	369	340	350
Variance	N	5-20	20	20	20	20
Saccade execution	t_{ex}	---	50	50	50	50
Saccade programming						
Labile stage (ms)	t_{lab}	200-400	333	310	290	290
Non-labile stage (ms)	t_{nlab}	30-120	58	41	40	40
Standard deviation (ms)	σ	---	0.33	0.33	0.33	0.33
Error	E	---	3.047	2.729	3.600	1.834
Error cumulative distrib.	E_c	---	1.500	4.881	14.580	4.641
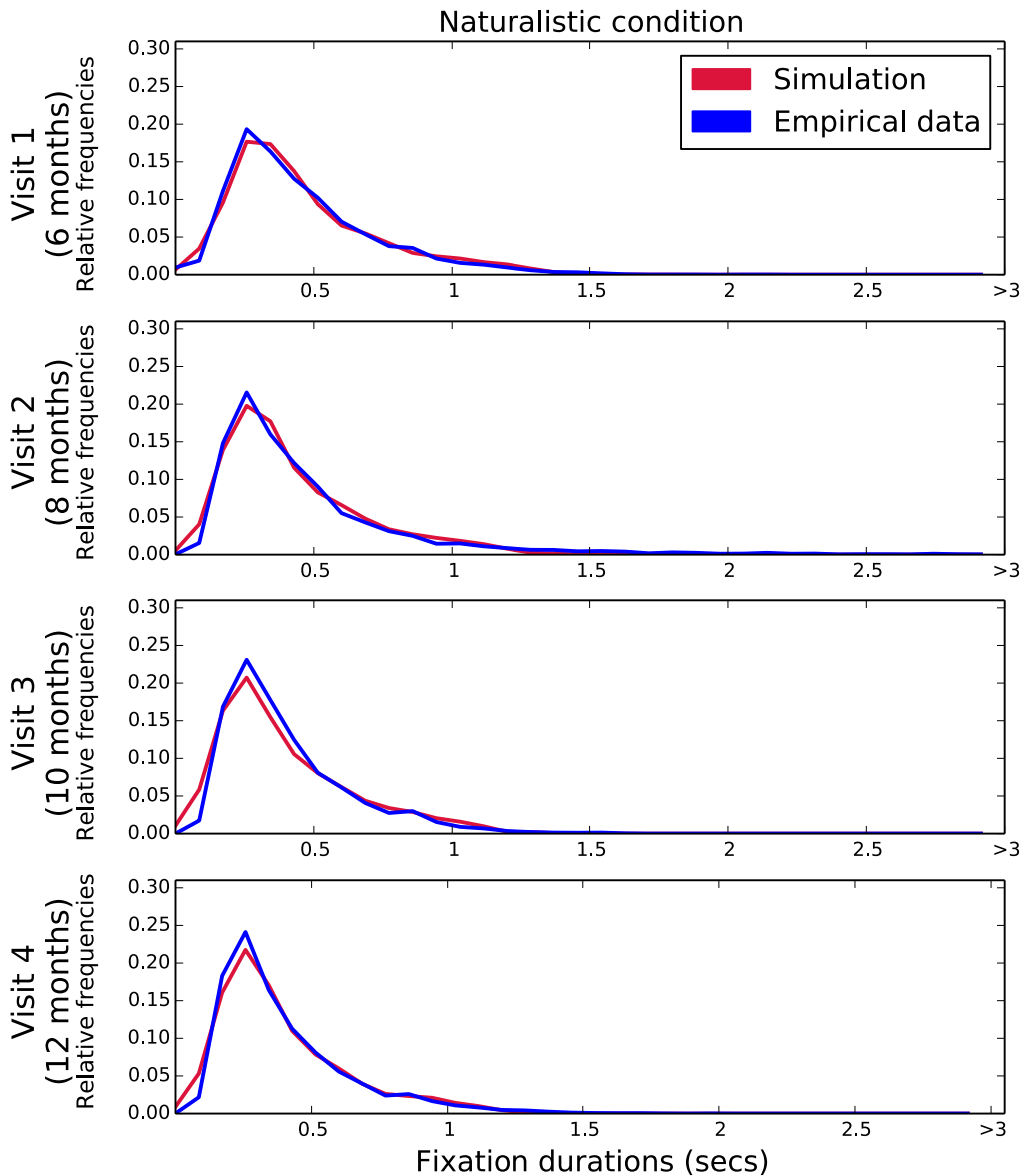
Figure 7-2 Simulation study 2. Simulations paired with empirical data for the naturalistic condition on the four visits.

The results from the statistical analysis performed on the simulated data resembled those found on the empirical data. Mean fixation durations decreased with age ($F(3,27)= 124.979$, $p < .001$, $p\eta^2= .933$), particularly from 6 ($M = 462$ ms, $SD = .003$) to 8 months ($M = 426$ ms, $SD = .003$; $p < .001$) and from 8 to 10 months ($M = 403$ ms, $SD = .002$; $p < .001$).

There was also a significant decrease in the labile stage of saccade programming ($F(3,27)= 360.664$, $p < .001$, $p\eta^2= .973$). As for mean fixation durations, pairwise comparisons showed a

significant decrease from 6 (M = 290 ms, SD = .807) to 8 months (M = 267 ms, SD = .780; p < .001) and from 8 to 10 months (M = 260 ms, SD = .736; p < .001). While an overall significant decrease was also found for the non-labile stage of saccade programming ($F_{(3,27)}$= 7447.349, p < .001, $p\eta^2$= .999), pairwise comparisons revealed that it was only significant from 6 (M = 58 ms, SD = .164) to 8 months (M = 40 ms, SD = .059; p < .001). Additionally, as in simulation study 1 results revealed a relative increase of the labile stage with respect to the non-labile stage program ($F_{(3,27)}$= 1939.947, p < .001, $p\eta^2$= .995). This change was significant from 6 to 8 months where the non-labile stage went from constituting the 16.659 % (SD = .043) of the saccade program to being the 13.011 % (SD = .033; p < .001). Noticeably, from 8 to 10 months this percentage suffered a slight increase (from 13.011 %, SD = .033, to 13.312 %, SD = .045; p = .013), mainly due to the sudden decrease of the labile stage program during this period.

With respect to saccade timing parameters, the mean values for the random walk timer decreased considerably from 6 to 10 months (6 months = 397 ms; 8 months = 369 ms; 10 months = 340 ms), but remained fairly stable from 10 to 12 months (350 ms). The saccade timing variance (number of steps for the random walk, N), on the other hand, did not vary across visits.

Figure 7-3 summarizes the results from simulation studies 1 and 2 and shows an overview off saccade timing and saccade programming parameter trajectories for the first year of life. Interestingly, for the overlapping visits at 6 months saccade programming parameters were very similar. On the other hand, saccade timing parameters were considerably larger for the results in simulation study 1. Since this parameter is assumed to be indicative of information processing, it might be the case that different groups of infants processed the visual information differently. These individual differences could be less obvious by adding more participants in each group (see Chapter 8).
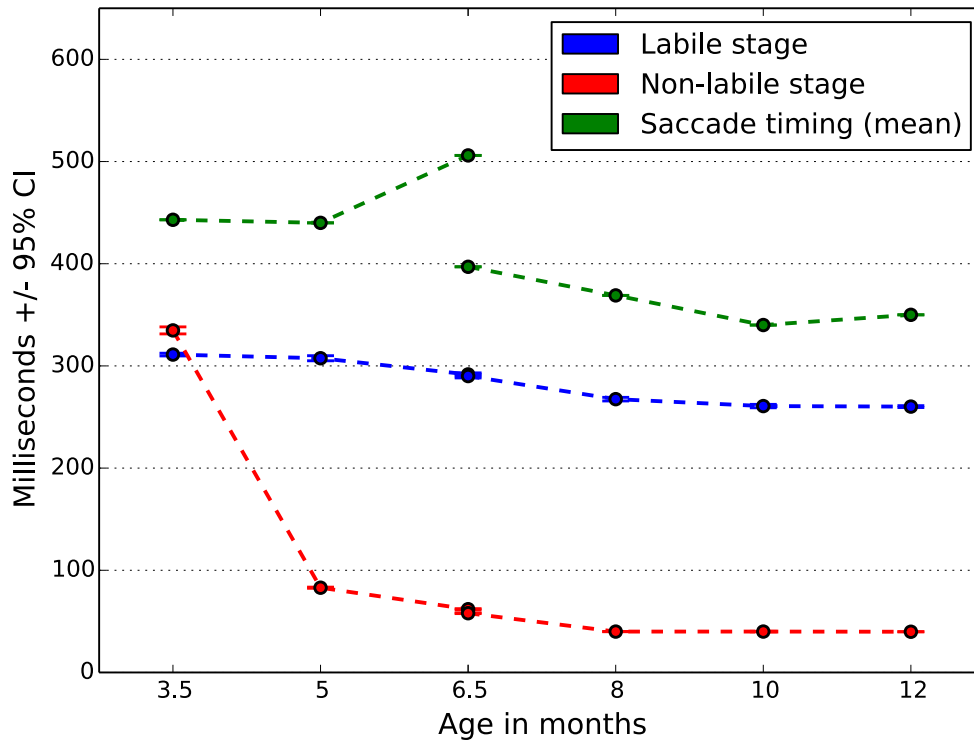
Figure 7-3 Simulation studies 1 and 2. Saccade timing and saccade programming parameter trajectories over the first year of life.

7.6 Simulation study 3: Individual differences between short lookers and long lookers

In the present simulation study I modelled the fixation data from two infants (one long and one short looker) from the group of infants that was tested at 3.5, 5 and 6 months (see longitudinal study 1 from Chapter 6). During this period, the frontal areas of the brain involved in saccade planning and execution are developing rapidly (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), leading to greater disengagement abilities and shorter fixation durations.

During the first year of life long lookers have been associated with larger disengagement difficulties (Frick et al., 1999; see also Blaga & Colombo, 2006) or poorer processing speed (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997). The present simulation study will analyse whether long lookers present a delayed trajectory in saccade timing and/or saccade programming parameters with respect to short

lookers. Furthermore, this section also represents the first demonstration of how the CRISP model can be applied to a single participant.

7.6.1 Behavioural data

Two infants from the first longitudinal study from Chapter 6 were selected and flagged as short and long lookers. The criteria to select these infants were (1) there were at least 50 fixations per visit; (2) for the long looker, mean fixation durations for each visit were above the mean; (3) for the short looker, mean fixation durations for each visit were bellow the mean. The long looker was infant number 6 from the first longitudinal study described in Chapter 6, while the short looker was infant number 2 (see Figure 6-6 and Figure 6-7). As for the previous simulation studies described in this chapter, with the view to reduce the number of free parameters only the naturalistic condition was modelled.

The empirical data set for the infant flagged as long looker was comprised of 166 fixations at 3.5 months, 175 fixations at 5 months and 242 fixations at 6.5 months. The empirical data set for the infant flagged as short looker was comprised of 66 fixations at 3.5 months, 226 fixations at 5 months and 180 fixations at 6.5 months.

7.6.2 Model adjustments

The model adjustments and assumptions were exactly the same as those described previously in this chapter.

7.6.3 Modelling results

As for the previous simulations, the model's free parameters were estimated using the genetic algorithm optimization technique described earlier in this chapter (see simulation study 1, or the simulation studies from Chapter 5). For each visit a total of four free parameters were fit: the mean durations for the labile and non-labile stages of saccade programming; and the mean and variance of the random walk timing signal for the naturalistic condition. The ranges for estimating these values were based on the same principles described in Chapter 5.

The parameter values that best resembled the empirical distributions for the long and the short lookers are described in Table 7-3 and Table 7-4 respectively.

Table 7-3 Simulation study 3. Best fitting parameters for the long looker at 3.5 months.

Parameter	Symbol	Range	3.5 months	5 months	6.5 months
Saccade timing					
Mean (ms)	t_{sac}	200-650	420	470	490
Variance	N	5-20	20	18	20
Saccade execution	t_{ex}	---	50	50	50
Saccade programming					
Labile stage (ms)	t_{lab}	30-400	355	370	330
Non-labile stage (ms)	t_{nlab}	30-350	307	87	57
Standard deviation (ms)	σ	---	0.33	0.33	0.33
Error	E	---	46.665	60.40	7.225
Error cumulative distribution	E_c	---	88.145	86.230	24.363

Table 7-4 Simulation study 3. Best fitting parameters for the short looker at 3.5 months.

Parameter	Symbol	Range	3.5 months	5 months	6.5 months
Saccade timing					
Mean (ms)	t_{sac}	200-650	530	421	414
Variance	N	5-20	20	20	20
Saccade execution	t_{ex}	---	50	50	50
Saccade programming					
Labile stage (ms)	t_{lab}	30-400	300	274	280
Non-labile stage (ms)	t_{nlab}	30-350	88	49	41
Standard deviation (ms)	σ	---	0.33	0.33	0.33
Error	E	---	10.552	6.082	5.999
Error cumulative distribution	E_c	---	163.157	6.012	6.136

Figure 7-4 displays the empirical distributions of fixation durations paired with the simulated data for long and short lookers respectively. These graphs exhibit less smooth empirical distributions as a consequence of the reduced number of fixations. Nevertheless the simulated distributions seem to capture the main trends of the empirical distributions well.
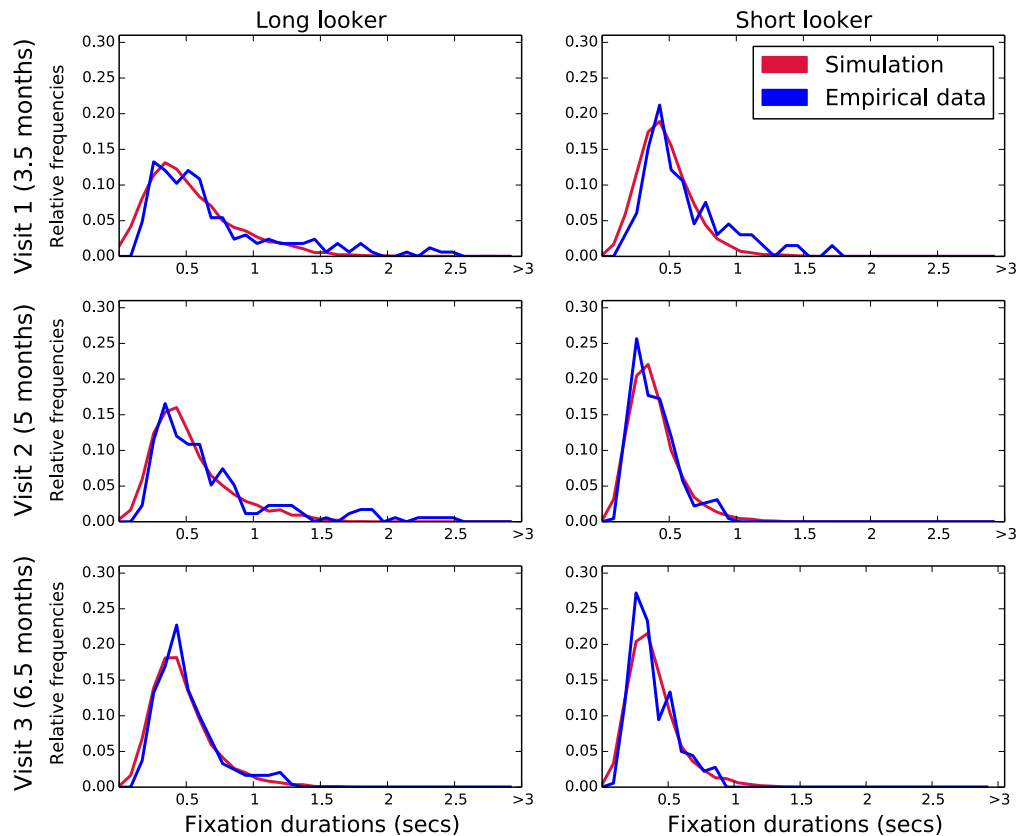
Figure 7-4 Simulation study 3. Simulations for the long and short looker participants paired with the empirical data for the naturalistic and abstract conditions on the three visits.

Whilst a statistical analysis was not suitable for exploring the differences between the long and short lookers in the empirical data (there is only one participant per group), it was possible to perform this analysis on the simulated data. As it was explained in the introductory chapter, computational modelling allows us to describe, predict and explain data that is itself unobservable (Lewandowsky & Farrell, 2011). For the present simulation, the model was able to simulate the data from 10 participants per condition, permitting a statistical analysis that compares the trajectories for the different variables in both groups (see simulation study 1).

A repeated measures ANOVA with a within-subjects factor was performed on the simulated data revealing an overall decrease in mean fixation durations ($F(1.466,26.399)= 315.323$, $p <$ .001, $p\eta^2=$ .946) and an interaction effect between age and group (long vs. short lookers) ($F(1.466,26.399)= 86.415$, $p < .001$, $p\eta^2=$ .828). In order to analyse the independent trajectories of mean fixation durations for short and long lookers two independent ANOVAs were performed.

Mean fixation durations decreased both for the long (F(1.211, 10.894)= 59.427, *p* < .001, *pη²*= .868) and the short lookers (F(2,18)= 1031.044, *p* < .001, *pη²*= .991) groups. Nevertheless, while for long lookers the significant decrease happened from 5 (*M* = 527 ms, *SD* = .003) to 6.5 months (*M* = 483 ms, *SD* = .002; *p* < .001), for the short lookers mean fixation durations only decreased from 3.5 (*M* = 492 ms, *SD* = .001) to 5 months (*M* = 403 ms, *SD* = .002; *p* < .001). Additionally three independent t-tests confirmed how mean fixation durations in each visit were shorter for the short lookers group (all p < .001). These results already suggest a delayed trajectory in mean fixation durations –and potentially a delayed development of the visuo-motor system- in the long lookers group as compared with the short lookers group.

Saccade programming trajectories were also statistically analysed. A repeated measures ANOVA with a within-subjects factor revealed an age effect in the labile stage of saccade programming (F(2,36)= 175.914, *p* < .001, *pη²*= .907) and an interaction effect between age and group (long vs. short lookers) (F(3,36)= 521.438, *p* < .001, *pη²*= .967). Two independent ANOVAs were used to analyse the labile stage independent trajectories for long and short lookers and results revealed age effects both in long (F(2,18)= 229.388, *p* < .001, *pη²*= .962) and short lookers groups (F(2,18)= 539.594, *p* < .001, *pη²*= .984). The labile stage significantly increased for the long looking group from 3.5 (*M* = 312 ms, *SD* = .922) to 5 months (*M* = 332 ms, *SD* = .918; *p* < .001), and significantly decreased from 5 to 6.5 months (*M* = 307 ms, *SD* = .850; *p* < .001). On the other hand, for the short looking group the labile stage significantly decreased from 3.5 (*M* = 284 ms, *SD* = .874) to 5 months (*M* = 254 ms, *SD* = .471; *p* < .001), and significantly increased from 5 to 6.5 months (*M* = 259 ms, *SD* = .784; *p* < .001). Furthermore, three independent t-tests confirmed how the labile stages in each visit were shorter for the short lookers group (all p < .001) as compared to the long lookers group. The slight differences in the trajectories evidence the effect of individual differences in the development of saccadic control in young infants. Unsurprisingly, as with mean fixation durations, labile stages were shorter for the short lookers group reinforcing the view that – according to the assumption stating that longer labile stages during the first year of life indicate a less developed visuo-motor system (see Chapter 5)- the short lookers had a more developed visuo-motor system at each visit.

Trajectories for the non-labile stage of saccade programming were investigated with a repeated measures ANOVA with a within-subjects factor that revealed an age effect (F(1.119,20.140)=

64893.735, $p < .001$, $p\eta^2 = 1.000$) and an interaction effect between age and group (long vs. short lookers) ($F_{(1.119, 20.140)} = 30643.241$, $p < .001$, $p\eta^2 = .999$). Two independent ANOVAs analysed the non-labile stage independent trajectories for short and long lookers. Results revealed an overall decrease in the non-labile stages of saccade programming for both groups (long lookers: $F_{(1.092, 9.828)} = 48830.630$, $p < .001$, $p\eta^2 = 1.000$; short lookers: $F_{(2,18)} = 29289.011$, $p < .001$, $p\eta^2 = 1.000$). The non-labile stages decreased from 3.5 (long: $M = 303$ ms, $SD = 1.001$; short: $M = 88$ ms, $SD = .200$) to 5 (long: $M = 87$ ms, $SD = .247$; short: $M = 49$ ms, $SD = .095$; all $p < .001$) and from 5 to 6.5 months (long: $M = 56$ ms, $SD = 210$; short: $M = 41$ ms, $SD = .083$; all $p < .001$) in both groups. Furthermore, the non-labile stages for the short lookers were significantly shorter than for long lookers in each visit (all $p < .001$). Interestingly, at 3.5 months the non-labile stage for long lookers constituted the 49.27 % ($SD = .413$) of the saccade programming, while for short lookers it was just the 23.52 % ($SD = .169$). Long lookers' non-labile stage experienced an important decrease from 3.5 to 5 months thus by the second visit it constituted just 20.79 % ($SD = .203$) of the saccade programming. According to the assumptions made at the introduction for this chapter, a shorter non-labile stage is associated with a greater ability to cancel a saccade and hence with a more developed visuo-motor system. Thus according to these results from 3.5 months short lookers already showed a more developed visuo-motor system as compared to long lookers.

Pertaining to saccade timing parameters, while the mean values for the random walk timer increased for long lookers (3.5 month: 420 ms; 5 months: 470 ms; 6.5 months: 490 ms), they decreased for short lookers (3.5 month: 530 ms; 5 months: 421 ms; 6.5 months: 414 ms). These results suggest that the processing speed between short lookers and long lookers also differs. More specifically, At 5 and 6.5 months long lookers presented longer means supporting past research where long lookers have been associated with poorer processing speed (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997). Furthermore, the conflicting trajectories for the mean values for the random walk timer suggest that as proposed by a few studies (Colombo et al., 1995, 1991) long and short lookers present different strategies of visual intake. The saccade timing variance (number of steps for the random walk, N) almost did not vary across visits.

For an overview off the saccade timing and saccade programming trajectories for long and short lookers see Figure 7-5.
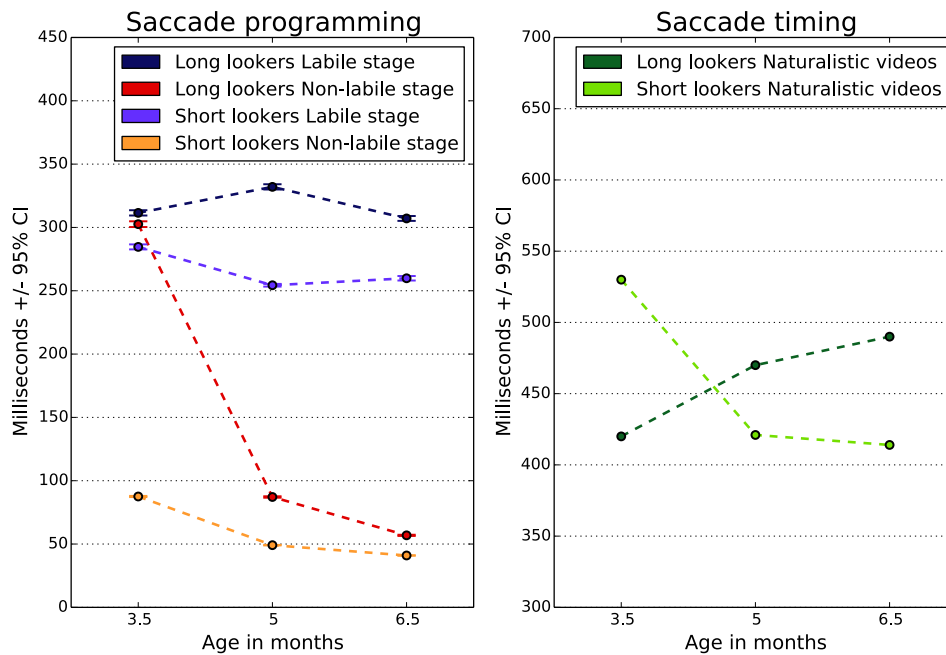
Figure 7-5 Simulation study 3. Saccade timing and saccade programming parameter trajectories for short and long lookers.

7.7 Discussion

The present chapter aimed to investigate the development of the mechanisms underlying saccadic control and fixation durations in early infancy. More specifically, the studies described in this chapter served to investigate saccade timing and saccade programming parameter trajectories during the first year of life and to examine how these trajectories differed for short and long lookers.

For this purpose I modelled fixation durations data from the longitudinal studies described in Chapter 6. With the view to reduce the number of free parameters in the model simulations only fixations from the naturalistic viewing condition were modelled.

Simulation studies 1 and 2 aimed to explore the development of saccade timing and saccade programming during the first year of life. These studies modelled the data from the groups of infants tested at 3.5, 5 and 6.5 months, and at 6, 8, 10 and 12 months respectively. More specifically, these simulation studies intended to test the assumptions about the development of saccade timing and saccade programming parameters made in Chapter 5. In this chapter I hypothesized that (a) saccade timing parameters -the mean values for the random walk timer-

would decrease over the first year of life as a result of improvements in processing efficiency. Furthermore, I predicted that from 4 months onwards there would be (b) a relative decrease of the non-labile stage with respect to the labile stage program, and (c) a gradual decrease in the absolute times for both labile and non-labile programs as the infant's saccadic control increases in efficiency.

Results from the statistical analysis performed on the simulated distributions of fixation durations and investigating mean fixation duration trajectories resembled those from the empirical data, verifying the validity of the model fitting.

Corroborating one of the assumptions from Chapter 5 (c), results showed how the labile stage of saccade programming kept decreasing until the 10 months visit. Similar results were found for the non-labile stage, even though it stabilized slightly earlier, at 8 months. The decrease of the labile and the non-labile stages of saccade programming suggest that the infant visuo-motor system kept developing and increasing in efficiency until infants were 8-10 months. These findings go in line with past evidence from various neurophysiological and behavioural studies suggesting that at 6 months postnatal age the neural mechanisms underlying saccade programming and saccade execution (such as certain eye-movement related areas in the superior parietal lobe) are still undergoing development or need to increase in efficiency, despite hundreds of thousands of practice saccades (Butcher et al., 2000; Csibra et al., 1998; Hood & Atkinson, 1993; Matsuzawa & Shimojo, 1997).

Furthermore, results also revealed a relative increase of the labile stage with respect to the non-labile stage program until the 8 months visit (b). This increase –and the concurrent decrease of the non-labile stage with respect to the labile stage- is particularly noticeable from 3.5 to 5 months, where the non-labile stage went from constituting the 51.829 % (*SD* = .125) of the saccade program to constituting the 21.244 % (*SD* = .077) at 5 months. These results support the assumption made in Chapter 5 maintaining that when infants' saccadic control is mainly driven by subcortical structures the non-labile stage of saccade programming will constitute most of the saccade program. As infants gain more volitional control over their eye-movements at 3-4 months due to the development of the eye-movement related areas of the prefrontal lobe (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), they will improve their ability to cancel a saccade program, and hence as the present results suggest

one could expect a relative increase of the labile stage with respect to the non-labile stage program.

In terms of the saccade timing parameters, the mean for the random walk timer did not necessarily decrease for all visits. From 3.5 to 5 months it remained stable while it slightly increased by 6 months, indicating an increase in information processing that could be due to the infants showing more interest in the stimuli at this age (see Frank et al., 2011). On the other hand, the mean values for the random walk timer decreased considerably from 6 to 10 months. These advances in processing speed could be the result of the developments in memory capacity that occur during the second half of the first year (Baillargeon & Graber, 1988; Cohen & Gelber, 1975; Colombo, 1993; Richmond & Nelson, 2009; Ross-sheehy et al., 2003; Ruff & Rothbart, 1996). The variance (number of steps for the random walk, N) for the random walk timer remained fairly stable across visits.

Finally, the simulation study 3 intended to analyse the trajectories for saccade timing and saccade programming parameters in infants classified as long and short lookers. Certainly, there is a large body of research investigating individual differences between short and long lookers, even though the majority of past studies have used looking times to classify long and short lookers rather than fixation durations (but see Papageorgiou et al., 2014; Wass & Smith, 2014). While some of this research has found a relationship between look duration and disengagement abilities –with long lookers presenting greater difficulties to disengage during the first year of life- (Frick et al., 1999; see also Blaga & Colombo, 2006), others found an association between long lookers and poorer processing speed (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997). Most likely, as discussed throughout this thesis long fixation durations during the first year of life are the result of both poorer disengagement abilities –a less developed visuo-motor system- and processing speed. For this reason I expected to find a delayed developmental change in saccade timing and/or saccade programming parameters for long lookers as compared to short lookers. The analysis performed on the simulated mean fixation duration trajectories already suggested a delayed trajectory in mean fixation durations in the long lookers group as compared to the short lookers group, that could potentially indicate a delay in the development of the visuo-motor system for long lookers. Similar results and conclusions were extracted from the analysis performed on saccade programming parameters. Labile as well as non-labile stages were shorter for the short

lookers in each visit. Interestingly, the labile and non-labile stages for long lookers at 3.5 months resembled the results obtained in simulation study 1 at this same age. On the other hand, 3.5-month-old short lookers already displayed labile and non-labile stages that were closer to 5-month-olds from simulation study 1. These results indicate a delayed trajectory of saccade programming parameters in long lookers, suggesting a slower development of the visuo-motor system in this group compared to short lookers.

Saccade timing parameters in long and short lookers displayed opposed results. While the mean values for the random walk timer increased for long lookers, they decreased for short lookers, indicating differences in processing speed between short and long lookers. From 5 months on long lookers showed longer mean values than short lookers, supporting past research where long lookers have been associated with poorer processing speed (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997). Moreover, the conflicting trajectories for the mean values for the random walk timer suggest that as proposed by few studies (Colombo et al., 1995, 1991) long and short lookers present different strategies of visual intake. The saccade timing variance (number of steps for the random walk, N) almost did not vary across visits.

In Chapter 5 I discussed whether disengagement latencies should belong to the labile stage of saccade programming or to the non-labile stage. While the present results do not reject the idea that disengagement abilities could affect the labile stage, they reinforce the idea that the primary disengagement problems lie in the non-labile stage. The non-labile stage at 3.5 months (simulation study 1), a time when the frontal structures associated with saccade programming and execution are still developing, constitutes around the half of the saccade programming. Interestingly, similar results were found for the long lookers in simulation study 3. Thus this suggests that at least to an extent the prolonged non-labile stages are can also be due to problems disengaging.

7.8 Conclusion

The present chapter presented three simulation studies that aimed to investigate the development of saccade timing and saccade programming parameters through the first year of life. For this purpose I modelled fixations data from the two longitudinal studies presented in

Chapter 6. Additionally, simulation study 3 investigated the different developmental trajectories in saccade timing and saccade programming in infants classified as long and short lookers.

Results showed a relative increase of the labile stage with respect to the non-labile stage program, more evident at 3.5-5 months and significant until 8 months. There was also a gradual decrease in the absolute times for both labile (until 10 months) and non-labile programs (until 8 months) as the infants' saccadic control increased in efficiency. These results indicate the influence in fixation durations of the developmental state of the visuo-motor system during the first and also the second half of the first year, even though both stages of saccade programming seemed to stabilize before 10 months. With regard to saccade timing parameters, the mean for the random walk timer did not necessarily decrease for all visits. Nevertheless, it decreased from 6 to 10 months, maybe due to the developments on memory capacity that occur at this age (Colombo, 1993). These results suggest that both the developmental state of the visuo-motor system and processing speed affect saccadic control until at least 10 months postnatal age.

Simulation study 3 investigated the trajectories for saccade timing and saccade programming in infants classified as long and short lookers. Results displayed delayed trajectories in saccade programming parameters for long lookers as compared to short lookers, indicating a faster development of the visuo-motor system in the later group. Furthermore, the analysis of saccade timing parameters suggested that long lookers have slowed processing (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997) and that they present different strategies of visual intake studies as compared to short lookers (Colombo et al., 1995, 1991).

Chapter 8: General discussion, limitations and conclusion

8.1 Revisiting initial goals and introduction to the final chapter

The overarching goal of this work was to investigate the unexplored mechanisms underlying saccadic control and fixation durations during the first year of life using both behavioural and computational methods. Additionally, a secondary goal of the present work was to develop new analysis tools able to extract fixations reliably from low and high quality eye-tracking data.

Chapter 1 provided the general literature background for this thesis, discussing past research on vision, visual orienting and visual attention. It particularly focused on reviewing previous empirical work and theoretical and computational models investigating visual attention and the mechanisms underlying oculomotor control and fixation durations in infants and adults. I reviewed adult studies investigating the locations where fixations land and their durations. Following this review I moved onto describing past literature on the development of looking behaviour and the few studies investigating fixation durations and saccadic control in infants. In infants, due to the numerous technical and practical constraints, very few studies have attempted to investigate the development of fixation durations and saccadic control during the first year of life and thus the understanding of the development of infant oculomotor control is restricted. This realization lead to the formulation of the motivation for the empirical studies carried out in the present thesis.

In Chapters 2 and 3 I introduced the methods, experimental tools, and analytical techniques. More specifically, in Chapter 2 I described the particularities and technical difficulties of recording infants' eye-movements and I outlined the techniques used within this thesis in order to minimize these problems (e.g., calibration techniques, equipment, room set-up). In Chapter 3 I introduced GraFIX, a new method and software that I developed to detect fixations reliably in low and high quality eye-tracking data.

Chapters 4-7 described the empirical work carried out within this thesis. In the present chapter, I discuss critically the findings of my experiments from a broader perspective. I examine the limitations of this work and how these might be addressed and overcome in the future. Moreover, I provide a discussion of issues to be considered when designing studies investigating infant visual attention in general and saccadic control in particular, and I propose future directions to follow the work described in this thesis.

8.2 Critical discussion

In this section I will describe the achievements of this thesis and the ways in which it might contribute to the wider field. Since throughout the present thesis I have provided detailed discussions of the results of my work (see Chapters 3, 4, 5, 6, and 7), I will focus here on some individual findings in an attempt to discuss them from a broader perspective. I will start discussing how the idiosyncrasies and technical difficulties that testing infants entail can affect experimental results. The attempts to overcome some of these issues are also discussed, including GraFIX, the new method and software that I developed in order to facilitate and improve the analysis of low and high quality eye-tracking data. I will move on discussing the main findings from the cross-sectional and longitudinal behavioural studies described in this thesis (Chapters 4 and 6), as well as the simulation studies that were performed in these data (Chapters 5 and 7). Finally, I will review how the use of computational models helped answering questions about infant saccadic control that could not be explained with more traditional methods.

8.2.1 Methods: Gathering data from infant participants

As obvious as it may seem, testing infants is not an easy task. To begin with, they do not attend to experimental instructions and will not necessarily pay attention to what is presented to them, especially for an extended period of time (e.g., > 10 minutes). For these reasons most adult studies are not suitable for infants and have to be adapted, which is not always feasible.

The first challenge that the researcher needs to face when planning an infant experiment is to design a study that infants at the targeted age will enjoy. Designing a fun experiment will not only considerably increase the chances of obtaining a decent amount of data, but also might improve the quality of it by encouraging infants to stay still in front of the experimental stimuli. This does not only apply to eye-tracking experiments but also to studies using any other techniques directed to test infant participants. In order to be able to model the fixation data described in this thesis, it was essential to gather as much data as possible (e.g., in Nuthmann et al., 2010, each adult participant was tested for about 40 minutes). Infant studies usually last 5 to 10 minutes, which is not enough time to collect an acceptable number of fixations per viewing condition. To overcome this issue the experiments described in this thesis went through a period of piloting during which the stimuli was repeatedly changed and adapted until infants enjoyed it enough to attend to it for around 20 minutes. For instance, the initial battery of

naturalistic stimuli that was used for the studies described in Chapters 4 and 6 was composed of videos previously used to test adult participants. Since infants did not find these stimuli particularly appealing, I decided to design my own infant-friendly naturalistic videos.

Secondly, infants are "noisy". It does not matter how much we try, infants may always have something else on their minds unrelated to the experimental stimuli they are being presented with. This is an issue to keep in mind when interpreting our experimental results. For instance, it is still not clear whether the very long fixations found in the semi-naturalistic (Chapter 4) and abstract conditions (Chapter 6) are a consequence of information processing or any other inner process unrelated to the visual information being encoded. These processes are indeed difficult to assess.

Thirdly, infant data is particularly noisy. As it was discussed in Chapter 3, the data quality from an eye-tracker can -and will- vary depending on the eye-tracker model and manufacturer, the eye physiology, the calibration procedure, the position of the participant relative to the eye-tracker, the degree of head motion (Holmqvist et al., 2011, 2012), or even ethnicity (Blignaut & Wium, 2014). Data quality is particularly problematic in infant populations for a number of reasons, such as the difficulties of using the corneal reflection technique with their very watery eyelids (Gredebäck et al., 2009), or the unexpected movements the infant makes in front of the eye-tracker. Data quality is in fact a very important problem that has still not received the attention it deserves, especially in fields like developmental psychology. As recent articles have pointed out (Holmqvist et al., 2012; Saez de Urabain et al., 2014; Wass et al., 2013), low data quality can substantially alter the experimental outcomes. While noise in the data cannot be avoided when using the current remote eye-trackers (also in adults), in Chapters 2 and 3 I described some guidelines (e.g., infant calibration, lighting conditions, the participant's distance and angle with respect to the eye-tracker) that could improve infant data quality.

Finally, eye-tracking results are not always adequately reported in scientific journals. While there are specific guidelines for publishing eye-tracking studies (McConkie, 1981; Oakes, 2010), not many articles include them, hindering the reproducibility of the results. In this thesis I included all the details to explain not only how the data was collected and treated, but also what the quality of the data was and its potential impact on the experimental results.

8.2.2 Methods: Detecting fixations in low and high quality eye-tracking data

8.2.2.1 Summary of main findings

In Chapter 3 I presented GraFIX, a new method and software to reliably detect fixations in low and high quality data. As it was described earlier, the quality of infant data tends to be particularly low both in terms of spatial and temporal precision. Previous fixation detection approaches were not ready to handle this kind of data. Automatic detection approaches can be remarkably imprecise, while the alternative hand-coding methods are highly time consuming.

Developing GraFIX I aimed to solve these problems by presenting a new method that combines both automatic and hand-coding approaches in a user-friendly desktop application. More precisely, in GraFIX data is initially parsed by using velocity-based algorithms whose input parameters are adapted by the user, and then manipulated using the graphical interface, allowing accurate and rapid adjustments of the algorithms' outcome. The automatic algorithms (1) smooth the raw data, (2) interpolate missing data points, and (3) apply a number of post-hoc criteria to automatically evaluate and remove artifactual fixations. These post-hoc evaluations include (3.1) merging adjacent fixations with a similar location, (3.2) removing all fixations with a root mean square (RMS) of inter-sample distances higher than a threshold, and (3.3) removing all fixations shorter than a given threshold. The user interface allows for rapid adjustments of the algorithm's input parameters (e.g., velocity threshold, interpolation latency), thus the user is able to immediately visualize the effect of different filters and parameter values on the data.

This method was successfully assessed from four different perspectives. First, the agreement between two different raters was evaluated using the intra-class correlation coefficient (ICC) in two groups of infants presenting low and high quality data. Secondly, I compared hand-coding results with the two-step approach (a combination of automatic detection and hand-coding), demonstrating that both techniques should generate exactly the same results. Thirdly, the outcome from GraFIX automatic algorithms was compared to the two-step approach. This evidenced how even though in high quality data the results from the algorithms alone were close enough to hand-coding results, it was not the case for low quality data. Finally, I compared hand-coding results with GraFIX automatic algorithms and previous automatic detection algorithms (the velocity-based algorithms from Wass, et al., 2013; the adaptive velocity-based algorithms from Nyström and Holmqvist, 2010; and the I-VT filter velocity based

algorithm as implemented in Tobii-studio 3.0.0). This illustrated the problems of using automatic detection approaches to handle variable eye-tracking data quality.

8.2.2.2 Contributions and general considerations

The empirical work described in this thesis could not have been accomplished without developing GraFIX. When I started investigating the techniques available to extract fixations from infant eye-tracking data, it became obvious the need for new analysis tools able to handle the particularities of low quality data. Certainly, as outlined earlier, previous approaches were highly imprecise and/or time consuming.

The main contribution of this new method is the use of a combination of automatic and hand-coding approaches. The outcome from this procedure is very similar to what we could obtain from a purely hand-coding approach (see validation 2 from Chapter 3), but it is much faster and precise. For instance, participants featuring high data quality will need only few manual adjustments after executing the automatic algorithms and thus the whole detection process may just take few minutes. The higher the data quality is, the less time the researcher will have to invest in hand-coding. Additionally, the use of the automatic approach adds precision in detecting the onsets and offsets of fixations reliably.

While the combination of both methods has been shown to be the most reliable technique for detecting fixations in variable quality eye-tracking data, both procedures alone already represent an improvement with respect to previous detection techniques. The automatic algorithms included post-hoc validation filters specifically implemented to handle low quality data, such as the filter to delete fixations with a RMS higher than a threshold. The idea of including post-hoc validation criteria is not new (Wass et al., 2013), but this particular filter is. Furthermore, the user interface allows the user to play with the input parameter values and see the effects that different values have in the data within seconds. This feature is particularly useful for users that are not familiar with fixation detection techniques. Many times these users tend to use the values that have been previously reported in scientific papers or those recommended by manufacturers, without considering how these values will affect their experimental outcomes. This is very concerning, since this practice can seriously alter the experimental outcomes, especially when dealing with variable quality data (Holmqvist et al., 2011).

Moreover, GraFIX hand-coding tool also constitutes an improvement compared to previous hand-coding approaches. First of all, fixations can be created, merged or deleted by simply clicking on them with the mouse. This method is considerably quicker than previous hand-coding procedures where fixations had to be coded by analysing videos frame by frame. Additionally, I implemented different visualization options where raw and processed eye-tracking data can be replayed in real time. This feature permits the user to look at the detection outcomes from different perspectives and make more constructive decisions of whether a particular event in the data is a fixation or not.

A controversial contribution of this method is the idea of using different sets of input parameters for different participants featuring various levels of data quality. As I discussed in Chapter 3, a particular set of parameters will affect low and high data quality in different ways. For instance, a low velocity threshold will detect fixations reliably in high precision data, whilst it will detect multiple very short artifactual fixations in low precision data. On the other hand, a high velocity threshold may work better in low precision data, but will merge fixations together in high quality data. For this reason I believe that the use of different parameters for different participants is justified. Nevertheless, it remains an open question whether it should be acceptable to include participants featuring very different data quality in the same analysis.

The main limitation of this approach is that it can still be time consuming compared to automatic algorithms. This might not be a problem if we have a relatively small number of participants (e.g., < 200), but when processing a large amount of data it can be tempting to choose a less rigorous method.

8.2.3 Factors affecting saccadic control and fixation durations at 6 months

8.2.3.1 Summary of main findings

In Chapter 4 I investigated different aspects related to infant saccadic control and fixation durations. For this purpose I gathered fixation data from 6-month-olds and adults that where presented with two different complex dynamic viewing conditions (naturalistic and semi-naturalistic videos) and performed the gap-overlap (Atkinson et al., 1992; Butcher et al., 2000; Elsabbagh et al., 2009; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; M. H. Johnson et al., 1991) and double-step (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954) paradigms.

As expected, results showed longer mean fixation durations for infants as compared to adults for both naturalistic and semi-naturalistic viewing conditions. Furthermore, mean fixation durations where also longer for the semi-naturalistic condition compared to the naturalistic one, both for infants and adults, evidencing how at 6 months infants also adapt their eye-movements to different viewing conditions.

The gap-overlap paradigm served to analyse participants' disengagement abilities. In the case of infants, longer disengagement latencies have been typically associated with a less developed visuo-motor system (Butcher et al., 2000; Farroni et al., 1999; Hood & Atkinson, 1993; M. H. Johnson et al., 1991). Results showed correlations between disengagement latencies and fixation durations not only in 6-month-olds, but also in adults for the naturalistic condition. This indicates that the ability to disengage does not only depend on the developmental state of the visuo-motor system but also in the visual stimulus being processed (see discussion from Chapter 5).

Individual differences between long and short lookers were found both in infants and adults, demonstrating a great control in gaze allocation at 6 months.

The double-step paradigm was used to investigate saccade cancellation in infants. While the limited amount of data did not allow for the same analysis performed in previous studies with adults (Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954), reaction times when infants made a single saccade to the second target were longer than when they made two consecutive saccades to the first and the second target, suggesting that in the former case they were cancelling the saccade program to the first target. The data and findings from this study constituted the basis for modelling the data with the CRISP model in Chapter 5.

8.2.3.2 Contributions and general considerations

The findings reported in Chapter 4 indicate that most 6-month-olds already have a well developed top-down control that modulates gaze allocation (as suggested by the differences in mean fixation durations found between different viewing conditions, and the individual differences between short and long lookers), even though the mechanisms underlying saccadic control may still need to further develop or increase in efficiency (as suggested by the differences in mean fixation durations between 6-month-olds and adults). These findings go in

line with past behavioural and neurological evidence suggesting that even though at 6 months the neural structures involved in saccade planning and execution are in place, they may still be going through a "calibration phase" to increase in efficiency (Butcher et al., 2000; Csibra et al., 1998; Hood & Atkinson, 1993; Matsuzawa & Shimojo, 1997). Saccade planning and execution involves subcortical areas such as the superior colliculus or the substantia nigra, and cortical areas such as the frontal (e.g., the frontal eye fields, FEF; the supplementary eye fields, SEF) and superior parietal lobes (Kurtzberg & Vaughan, 1982; Posner, 1988; Schiller, 1985, 1998). Csibra, Tucker, and Johnson (1998) reported that at 6 months the saccade-related potentials (SRPs), associated with the eye-movement related areas of the parietal cortex that are found to precede the onset of a saccade (Kurtzberg & Vaughan, 1982), are weak, less synchronized to the saccade execution, or completely absent. As the findings from this thesis, these findings suggest that at 6 months the infant visuo-motor system is still undergoing development.

Six-month-olds as well as adults showed significant differences in mean fixation durations between viewing conditions, indicating the influence of visual and cognitive processing in saccadic control. Even though some previous studies have investigated this issue, they did not use particularly complex visual stimuli (but see Hunnius & Geuze, 2004). In this study I aimed to move forward and investigate infants' eye-movements under complex dynamic conditions that are more ecologically valid. Regardless of the change in the stimuli, results were similar to previous studies where differences between viewing conditions were also found at the same age (Hunnius & Geuze, 2004a). Interestingly, mean fixation durations remained significantly longer for the semi-naturalistic condition as compared to the naturalistic one, which was visually and semantically more complex. This raises the question of whether longer fixation durations could be to an extent a consequence of higher requirements in processing more complex visual stimuli, or if on the other hand they are not reflecting any cognitive activity related to the visual stimulus. In Chapter 4 I argued how both alternatives are feasible and not mutually exclusive. This question could be further explored in future experiments by including heart-rate measures (Courage et al., 2006; Reynolds & Richards, 2007; see section 8.3.2).

Another relevant finding from this study is the positive relationship between mean fixation durations and disengagement latencies from the gap-overlap paradigm in infants and to a lesser extent also in adults, with longer fixation durations being associated to longer disengagement latencies. As explained earlier in this thesis (see Chapters 1 and 4) disengagement latencies in

infancy have been typically associated with the developmental state of the visuo-motor system (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011). This could explain the stronger correlations in 6-month-olds, even though at this age disengagement difficulties are thought to have somewhat diminished. Nevertheless, as suggested by the adult correlations in the naturalistic condition, disengagement can also be affected by the visual stimuli being processed. This conclusion is backed up by previous studies from Kikuchi and colleagues (2011), where they reported larger disengagement latencies and saccade-related event-related potentials (ERPs) when children disengaged from faces.

The double-step paradigm was used to investigate saccade cancelation in infants. Whilst this paradigm has been widely used in adult research to investigate saccade programming (see Chapter 1; Becker & Jurgens, 1979; Findlay & Harris, 1984; Walker & McSorley, 2006; Westheimer, 1954), in infancy research it has only been applied to identify the shift from retinotopic to spatiotopic coordinate systems (e.g. Gilmore & Johnson, 1997; Kaufman, Gilmore, & Johnson, 2006; Senju et al., 2011). Technical and practical constraints in infant testing are the main reason why equivalent analyses were not feasible in infants and adults (see discussion from Chapter 4). Unlike infants, adults were expected to follow the experimental instructions and to perform a high number of trials. The version of the double-step paradigm described in Chapter 4 was designed to encourage infants to look at the second target by making it visually more attractive, as it was done in previous infant studies utilizing the anti-saccade task (M. H. Johnson, 1995b). The visual targets were colourful and engaging in order to keep the infant interested for longer periods, and the time lapse between the first and the second target's appearances was adapted to each participant's responses using the modified binary search (MOBS) thresholding algorithm. Still, it was not possible to gather enough valid trials in order to determine the threshold between the labile and non-labile stages of saccade programming, as estimated in adult studies (Becker & Jurgens, 1979). Results from this study suggested, however, that 6-month-olds were able to cancel a saccade. In any case these results were not satisfactory enough for the purposes of this thesis and hence I decided to remove this paradigm from future designs and focus on gathering more fixations from more viewing conditions (see Chapters 6 and 7). Future studies aiming to identify the labile and non-labile stages of saccade programming in infants should consider having multiple testing sessions close in time in order to gather enough data per participant.

The results and the data from this study were not only essential to model infant fixation durations in Chapter 5, but also served to build a strong experimental design that was used in the longitudinal studies described in Chapter 6.

8.2.4 The mechanisms underlying saccadic control and fixation durations at 6 months

8.2.4.1 Summary of main findings

The simulation studies described in Chapter 5 intended to explore the mechanisms underlying saccadic control and fixation durations in 6-month-olds by determining the generalizability of extending the CRISP theoretical framework and computational model of fixation durations in adult scene-viewing (Nuthmann et al., 2010) to infants. Furthermore, the model was used to test specific developmental questions on saccadic control and explore the limitations of the CRISP adult theoretical framework when applied to infant data.

Results from simulation study 1 demonstrated the CRISP model and its theoretical framework's generalizability –such as the two-step notion of saccade programming– to data from 6-month-old infants. Furthermore, results showed great differences in saccade programming and saccade timing parameters between infants and adults, suggesting that -unsurprisingly- 6-month-olds are slower processors than adults and that their visuo-motor system is still undergoing development. Simulation studies 2 and 3 tested whether fixation durations in infancy are influenced by the developmental state of the visuo-motor system and/or by visual and cognitive processing speed, with results suggesting that while both factors influence oculomotor control at this age it is visual and cognitive processing what contributes most to developmental change.

8.2.4.2 Contributions and general considerations

Computational models are able to describe, predict and explain data that is itself unobservable (Lewandowsky & Farrell, 2011) allowing us to investigate the mechanisms that are not directly accessible through experimentation (Braitenberg, 1984; Schlesinger & McMurray, 2012). This technique is particularly useful for investigating infant development, where the limitations testing infants pose important constraints (see Chapters 2 and 3). Current behavioural techniques have been unable to investigate certain aspects of the development of saccadic control, such as saccade cancellation or the two-step notion of saccade programming. Probably one of the most

important contributions of the simulation studies presented in Chapter 5 is the use of a computational approach to investigate the unanswered questions related to the development of saccadic control and fixation durations in 6-month-olds. While the simulation studies were grounded in assumptions and the conclusions that were taken from them were not based on any "real" data (the simulated data), these results provide a feasible interpretation of the data that could be further investigated with other future studies.

But, why use the CRISP model? Might other models of oculomotor control not explain infant data just as well? I chose this model over others for a number of reasons. First of all, the CRISP model investigates fixation durations in scene viewing, which is also the focus of the studies described in this thesis. Additionally, the model's architecture is adequate to answer specific questions about the development of the mechanisms underlying fixation durations and saccadic control. In particular, I was able to use the free parameters from the CRISP model as an index for the developmental state of the visual system (saccade programming parameters) and visual-cognitive processing (saccade timing parameters). Finally, the use of this model allowed investigating saccade cancellation and the two-step notion of saccade programming, which due to practical and technical difficulties has not been specifically analysed before. Nevertheless, this does not mean that the CRISP model is the only option able to explain infant oculomotor control. Other models could also provide new valid perspectives and potentially extend the work presented in this thesis.

Free model parameters were fitted using a genetic algorithm optimization technique. Nevertheless, due to the very skewed infant fixation duration distributions, sometimes the results had to be slightly adapted by hand (this was also the case for the baseline simulation in Nuthmann et al., 2010). Even though adapting free parameters by hand is relatively common in computational modelling, it can be a matter of concern. How do we know that the parameters we are reporting are the best fitting? Certainly, other parameter combinations could also provide a good fit. Some parameter combinations can be discarded based on the plausibility of the results. For instance, very short labile stages in 3-month-old infants can be discarded since these results are not physiologically plausible. In this thesis, the genetic algorithm was used as an approximation technique whose results were later optimized manually. The final parameter values were never massively changed from the results obtained with the genetic algorithm, with the view to be as rigorous as possible.

Results showed that the CRISP computational model was applicable to 6-month-olds, implying that its theoretical framework is also feasible for infants. One of the most relevant assumptions from the CRISP theoretical framework states that saccades are programmed in two stages, an initial labile stage during which the saccade is subject to cancellation, and a subsequent non-labile stage where the saccade can no longer be cancelled. This notion has not previously been investigated in infants, nor has saccade cancellation. Results from Chapters 4 and 5 suggested that saccades at 6 months were programmed in these two stages. Nevertheless, due to the cross-sectional nature of the study the development of saccade programming had to be hypothesized and tested later in subsequent studies (see Chapters 6 and 7).

Results from these simulation studies also suggested that at 6 months fixation durations are mainly affected by information processing speed, even though the developmental state of the visuo-motor system still plays an important role. I proposed that most likely these two factors are linked together, as suggested by previous studies that concluded that a less developed brain requires more time for information processing (Rose, Feldman, & Jankowski, 2002).

These simulations also indicated some limitations when applying CRISP to infant data described in detail in Chapter 5 and reviewed later in this chapter.

8.2.5 The development of saccadic control and fixation durations over the first year of life

8.2.5.1 Summary of main findings

The studies described in Chapter 6 intended to investigate the development of fixation durations and saccadic control during the first year of life. For this purpose I presented two longitudinal studies and a study with adult participants that were eye-tracked while viewing a battery of dynamic (naturalistic and abstract) and static stimuli and performed the gap-overlap paradigm. Specifically, these studies intended to analyse fixation duration trajectories for different types of stimuli (dynamic and static) during the first year of life as well as to explore individual differences in fixation durations across visits and stimulus types. Moreover, the gap-overlap paradigm was used to explore the relationship between fixation durations and disengagement latencies during the first year.

Overall, mean fixation durations decreased during the first year of life, with different viewing conditions decreasing at different rates and times. While some previous studies reported an

increase in looking times when viewing more complex forms of visual stimuli, (Courage et al., 2006; Frank et al., 2014, 2009; Shaddy & Colombo, 2004), in the studies from Chapter 6 mean fixation durations did not increase for any of the viewing conditions. The decrease in fixation durations could be explained by an improvement in processing efficiency (Colombo et al., 1991), the improvements in disengagement abilities during the first 6 months (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011), and/or the advances in memory capacity that occur during the second half of the first year (Colombo, 1993).

Individual differences across visits and viewing conditions were also found, even though they did not appear to be as stable as the ones found in adults.

With regard to disengagement latencies, they diminished during the first half of the first year, probably as a consequence of the development of the areas in the frontal cortex concerned with saccade programming and execution (Atkinson, 2000; Bronson, 1974; Frick et al., 1999; M. H. Johnson, 1990, 2011). While there were few positive correlations between mean fixation durations and disengagement latencies for some of the viewing conditions and visits, they were not very stable.

8.2.5.2 Contributions and general considerations

Whilst the studies described in Chapters 4 and 5 provided a detailed account about fixation durations and saccadic control at 6 months, they left some open questions to be investigated in subsequent chapters. The results from the cross-sectional study described in these chapters served to design the longitudinal studies from Chapters 6 and 7. For instance, in the simulation studies from Chapter 5 it became evident that more fixations per participant were needed in order to obtain more reliable results from the model, and to potentially be able to model participants individually. To solve this issue I decided to include more free-viewing scenes and to exclude the double-step paradigm from the protocol, since results from Chapter 4 were not satisfactory and the task seemed to make infants tire easily (see section 8.2.3.2). Additionally I included a third viewing condition: static images.

Few studies have investigated the differences in looking at dynamic and static stimuli in infants and children (Bronson, 1990; Hunnius & Geuze, 2004a; Shaddy & Colombo, 2004; Stoesz & Jakobson, 2014). For instance, Bronson (1990) found that 2-14-week-old infants had more difficulties disengaging when the stimuli was flickering. Similarly, Hunnius and Geuze (2004)

reported that scanning patterns for dynamic videos only stabilized at around 18 weeks, which is slightly later than what is usually reported for the scanning of static images. Results from the longitudinal studies in Chapter 6 backed these findings, with trajectories of mean fixation durations in static images stabilizing already during the first half of the first year. From 3.5 months, infants showed significantly shorter fixation durations in the static images condition as compared to the other two dynamic conditions. As discussed in Chapter 6, the big changes in fixation durations between static and dynamic stimuli could be the result of two different processes. First, the processing of motion may be adding an extra step when encoding visual information from a dynamic scene. The cortical pathway that goes to the superior colliculus from the primary visual cortex (V1) through the middle temporal area (MT), is thought to be involved in motion detection and the smooth tracking of moving stimuli (M. H. Johnson, 1990, 2011; Schiller, 1985, 1998) and is already functioning at about 2 months of age (M. H. Johnson, 1990, 2011). Dynamic visual information is thus also processed through the MT pathway –which presumably is not involved in the processing of static visual information-, leading to prolonged fixation durations. Secondly, a dynamic scene may contain more semantic content that needs to be processed -such as the understanding of the actions that the different characters are performing- that can prolong fixation durations.

Overall static images were the viewing condition that could most predict performance in later testing sessions, as evidenced by the strong correlations found across visits. This might be because this condition is less complex than the dynamic scenes (no motion), and thus stabilizes earlier.

The poor correlations between disengagement latencies and mean fixation durations were unexpected. Findings from the cross-sectional study in Chapter 4 showed strong correlations between these two measures in 6-month-olds, thus I expected to find similar results particularly in infants younger than 6 months, since during this period their disengagement abilities are though to be undergoing development. One possible explanation for the poor correlations reported in Chapter 6 is that infants could be predicting the target's next location (left or right) and thus the paradigm might as well be recording anticipatory eye-movements. In Chapter 4, trials from the gap-overlap paradigm were alternated with double-step trials, making it difficult to predict where the target was going to appear next. In future studies with the gap-overlap task I

would recommend to also include random targets at different locations in order to avoid infants predicting the next target location.

In Chapter 6 I used the Markov chain Monte Carlo (MCMC; Schafer, 1997) multiple imputation technique, a new procedure to deal with missing data. While this technique has rarely been used in similar psychological studies –and never for longitudinal studies-, it proved to be a reliable way to interpolate missing values. Nevertheless, it is still not clear whether it should be used when the number of missing points is high relative to the number of participants gathered. These aspects should be further investigated in future studies.

8.2.6 The development of the mechanisms underlying saccadic control and fixation durations over the first year of life.

8.2.6.1 Summary of main findings

The simulation studies described in Chapter 7 aimed to investigate the development of saccade timing and saccade programming parameters over the first year of life and to analyse how parameter trajectories differ for short and long lookers. As in Chapter 5, data was modelled with the CRISP model (Nuthmann et al., 2010).

In simulation studies 1 and 2, fixations from the naturalistic condition from the first and the second longitudinal studies described in Chapter 6 were modelled. As it was hypothesized in Chapter 5, results showed a relative increase of the labile stage with respect to the non-labile stage program, more evident at 3.5 - 5 months and significant until 8 months. There was also a gradual decrease in the absolute times for both labile (until 10 months) and non-labile programs (until 8 months) as the infants' saccadic control increased in efficiency. Regarding saccade timing parameters, the mean for the random walk timer –indicative of visual and cognitive processing speed- did not necessarily decrease for all visits. However, it decreased from 6 to 10 months, maybe due to the developments on memory capacity that occur at this age. These results suggest that both the developmental state of the visuo-motor system and processing speed affect saccadic control until at least 10 months postnatal age.

Simulation study 3 investigated the trajectories for saccade timing and saccade programming parameters in infants classified as long and short lookers. I found delayed trajectories in saccade programming parameters for long lookers as compared to short lookers, suggesting a faster development of the visuo-motor system in the later group. Additionally, the analysis of

saccade timing parameters suggested that long lookers are slower processors (Colombo et al., 1991; Colombo & Mitchell, 1990; Cuevas & Bell, 2013; Jankowski & Rose, 1997) and that they present different strategies of visual intake compared to short lookers (Colombo et al., 1995, 1991).

8.2.6.2 Contributions and general considerations

As in Chapter 5, in Chapter 7 I used computational modelling to investigate mechanisms that are not directly accessible through experimentation. Findings from the simulation studies confirmed the assumptions made in Chapter 5 and provided a feasible interpretation to explain the development of fixation durations and saccadic control in young infants. As for the findings in Chapter 5, this interpretation needs to be further investigated in future empirical studies.

According to the modelling results, both the developmental state of the visuo-motor system and processing speed affect saccadic control until at least 10 months postnatal age. These findings suggest that much of the variance in fixation durations that is observed during the first year (and particularly during the first 6 months) come from the development of the mechanisms underlying saccade programming and saccade execution. This finding is very important in order to interpret the results from studies using various looking measures.

Moreover, the differences between short and long lookers in saccade programming and saccade timing parameters were striking, with short lookers resembling results from older infants in each visit. This also evidences the substantial variability in fixation durations that can be found in young infants at the same age.

8.3 Limitations

Having discussed some contributions of my findings, I will now focus on some more general limitations of my approach. While some of the limitations were inevitable due to choices that had to be made prior to my studies and other restrictions (e.g., the number of participants), others emerged only during the testing or the analysis phase (e.g., data quality). Disentangling the weaknesses of the studies presented in this thesis will help other researchers to consider their experimental designs and methods for future studies more carefully. As such, reporting the limitations is an essential process in ensuring the highest research standards are maintained.

8.3.1 Experimental stimuli

Most research investigating eye-movements in infants is based on results from basic oculomotor paradigms such as the gap-overlap, or the presentation of relatively simple static and dynamic stimuli. This approach has been often criticized due to the limited ecological validity of the stimuli (Hunnius & Geuze, 2004a; Neisser, 1976; Schmuckler, 2001). Some studies have attempted to design more naturalistic stimuli. For instance, Hunnius and Geuze (2004) investigated the development of scanning in dynamic scenes, and for this they recorded the infants' mothers' face, which was more ecologically valid than a random face.

In the present thesis I have attempted to design stimuli that were as naturalistic as possible (see Chapters 4 and 6). Nevertheless, as in most current eye-tracking infant studies the videos were still displayed on a computer screen, which reduces the ecological validity of the experiment. Recently few head-mounted eye-tracker models have become available for use with infants (Aslin, 2012; Corbetta et al., 2011; Franchak & Adolph, 2010; Franchak et al., 2010). These new advances in eye-tracking equipment will allow researchers to investigate cognitive development and eye-movements in the real world. Even though using these devices implies new difficulties in testing and analysis that still need to be addressed, head-mounted eye-trackers allow researchers to gather a baste amount of data of more natural eye-movement behaviour.

8.3.2 Experimental design

In Chapters 4 and 6 significantly longer mean fixation durations were reported for certain viewing conditions. The interpretation of these results, however, was not straightforward: were the long fixation durations in certain conditions consequence of the higher cognitive demands of the stimuli? Or rather the result of other inner processes unrelated to the stimuli being viewed? The design of the studies presented in this thesis does not allow giving a definite answer to these questions.

Some infant researchers have argued that measures of look duration do not reflect information processing directly, and have attempted to use heart-rate measures to determine which portion of a look is relevant in attentional processes (Courage et al., 2006; Reynolds & Richards, 2007). For instance, Reynolds and Richards (2007) measured look duration together with the participant's heart-rate and concluded that it is only in the portion of looking that happens during the heart-rate defined phase of sustained attention when the infants engage in actively

processing visual information. It is important to note that these studies measured look duration, and not fixation durations. While it would be worth trying similar research on fixation durations, these may still be too short to be able to relate them to the heart-rate phases of attention.

8.3.3 Participants and statistical analysis

The number of participants per group was not so high, particularly in the first longitudinal study where only 12 infants could be tested. This was a problem for performing the statistical analysis on the data using the Linear Mixed Model (LMM) procedure. An LMM allows for the analysis of trajectories from two different cohorts as long as they share a point. The two longitudinal studies share the 6 months visit (which is the last visit for the first longitudinal study, and the first for the second longitudinal study), thus technically it should be possible to analyse both cohorts together in a single analysis. This analysis was not possible, however, due to the low number of participants in the first cohort. Most likely this also affected the correlational analyses, for which usually many more participants are required. Future studies should increase the number of participants participating in longitudinal studies and also factor in a larger proportion of drop-outs than would be expected in studies involving older children and adults, especially if the later time points are after most parents have returned to work and have less free-time to devote to participating in such studies.

8.3.4 Data quality

Low data quality is a major issue when testing infants (see Chapters 2 and 3). Regardless of my attempts to minimize the problems associated with infant testing and data analysis, there were still differences in data quality across groups. These problems will remain present until more effective high-speed eye-trackers become available for infants.

As discussed in Chapter 4, it is still debated whether participants featuring very different levels of data quality should be included in the same analysis. In fact, when these differences occur across groups (e.g., 3-month-old, vs. adults) it could potentially alter the experimental results. In general, to date not many studies have paid attention to these issues (particularly in developmental psychology). Hopefully the present work will help raising awareness on the importance of data quality in research.

8.3.5 Modelling infant fixation durations with CRISP

The particularities of infant fixation data add an extra level of complexity for modelling fixation durations with the CRISP model. First of all, it is important to gather a large number of fixations per participant, which is not always feasible when testing infants. For instance, the cross-sectional study described in Chapter 4 did not include enough free-viewing scenes in order to gather a large number of fixation data per participant. Consequently, modelling participants individually was not possible. The design for the studies described in Chapter 5 included more free-viewing scenes and this problem was solved.

In Chapter 5 I also discussed the difficulties of modelling very skewed fixation duration distributions. As a result of the high variability in fixation durations in infants, their fixation duration distributions are particularly skewed and might include very long fixations (> 2-3 secs). This can be due to problems disengaging and/or less efficient information processing. Adult fixation duration distributions are also to a lesser extent positively skewed, even though they tend to present much shorter tails. Moreover in adults long fixations are fairly uncommon and tend to be discarded from the analysis, the rationale being that such long fixations may not be indicative of any perceptual activity or information processing (Inhoff & Radach, 1998). For the present studies I decided not to include an upper thershold that discards these long fixations.

The CRISP model captured the tails of the distributions to a certain degree, but was not able to simulate the very long fixations from infant fixation duration distributions. The genetic algorithm optimization technique did not handle the very skewed distributions well, and frequently results had to be manually fitted. Surely, more work needs to be done to improve the model's optimization technique. Furthermore, the model could be extended to include a *disengagement component* able to further explain the very skewed distributions in very young infants (see *Future considerations*).

8.3.6 Modelling individual differences

Given a set of parameters, the CRISP model is able to generate a number of fictitious participants, each with slightly different fixation duration distributions. While these distributions are different, all derive from the same set of parameters and thus they are very similar. In Chapter 7 I used a statistical analysis to explain the modelling results. Even though this can help explaining the data, it is important to take into account that since the model is not able to

generate individual differences, the significance of the analysis is not valid. In the future, using the data and results presented in this thesis the CRISP model could be extended to account for individual differences.

8.4 Future directions

While this thesis already provided valuable insights into the development of fixation durations and saccadic control, it also raised new questions and opened up new strands of research. In this section I overview the future directions of this thesis work.

8.4.1 Improving fixation detection techniques

In Chapter 3 I presented GraFIX, a new method and software to detect fixations in low and high quality eye-tracking data. As demonstrated with the validation analyses, GraFIX provides a better alternative for detecting fixations compared to previous automatic and manual approaches. Nevertheless, it is still not perfect.

As any hand-coding method, the researcher needs to establish some guidelines that determine what will be coded as a fixation. While a reliability analysis can quantify the consistency of coding fixations within the same study, different researchers may have different ways of coding. For instance, a researcher may not pay attention to small microsaccades and will merge them together in a long fixation, while another researcher will consider fixational eye-movements when coding eye-tracking data. This way, without looking at the actual data it is not possible to determine how fixations were coded.

A way to solve this problem is by making eye-tracking data publicly available in a single database, where other researchers can download it. Similarly, publishing raw data together with scientific articles can allow replicated analysis as well as meta-analyses. This could help standardizing the coding guidelines as well as would make research more transparent. COGAIN[6] has lead the main initiative to standardize eye-tracking data-quality by investigating how eye-trackers compare in their performance, and designing methods and measures for valid comparison. More specifically, they have focused on eye data quality standardization, standardization of eye-data terminology, designing a standardized set of artificial eyes, designing a standard experimental protocol for the collection of data for data quality measurement, investigating physiological measures of eye data quality, implementing

---

[6] http://www.cogain.org/info/eye-data-quality

commercial eye trackers to the same experimental software, or investigating the effects of data quality on eye movement measures. Such initiatives are invaluable for the advance of eye-movement research.

8.4.2 Analysing fixation locations: Saliency models

Gaze control has been typically investigated through the study of two main aspects: the location where fixations land (fixation position or location) and the time-lapse during which the eyes remain still at a given location (fixation duration). In infants, a number of studies have investigated fixation locations (e.g., Bronson, 1990, 1994; Frank et al., 2014; Frank, Vul, & Saxe, 2011; Hunnius & Geuze, 2004a). Only few studies, however, have analysed fixation durations (but see Bronson, 1990; Hunnius & Geuze, 2004b; Papageorgiou et al., 2014; Wass & Smith, 2014). With the view to fill this gap the present work has focused on fixation durations. Still, there are a number of questions related to fixation locations that could be answered with further analyses on the data collected for this thesis and the use of saliency models. For instance, how do long and short fixations relate to the region of the stimuli being viewed? Are long fixation durations more likely to fall on areas with more semantic or visual content? How fast do infants react to motion in a complex naturalistic scene? And in an abstract scene?

8.4.3 CRISP Disengagement component

As mentioned earlier, the CRISP model was not able to capture the long tails of the very skewed distributions. I proposed that future extensions of the model could include a disengagement component able to simulate the very long fixation durations that are most likely consequence of problems disengaging (see Chapter 5).

In Chapter 5 I discussed whether the disengagement component should affect the labile or the non-labile stage of saccade programming. I concluded that both stages should be influenced by this component, since disengagement abilities are related to both difficulties triggering a saccade (in young infants) and the visual stimuli currently being processed. Nevertheless, more research needs to be done to determine the functioning of this new component, and the extent to which it would affect the labile and non-labile stages in different participants.

8.4.4 Using the model to generate novel predictions

In this thesis I have used computational modelling to explain and interpret fixation durations data from infants. Another use for computational models is to generate novel predictions based

on preliminary data. This approach has the potential of predicting how infants that were tested at a certain age will perform later on. Such a model could be extremely advantageous for the study of atypical development, where the study of fixation durations is gaining popularity. For instance, recent studies have reported that 6 to 9 months infants at familial risk of autism spectrum disorders (ASD) with shorter fixation durations are more likely to receive an ASD diagnosis later on (Wass et al., 2015). The data from the longitudinal studies described in this thesis could be used as a basis for developing a predictive model of typically developing infants that could later be extended for atypically developing infants.

8.5 Conclusion

The main goal of my thesis work was to investigate the development of fixation durations and saccadic control over the first year of life in a series of cross-sectional and longitudinal studies where I used eye-tracking as well as computational methods.

The eye-tracking studies highlighted how infant fixation durations are influenced by the low-level visual features from the viewing material (bottom-up control) from a very early age. It was not until the 5 months visit that infants also showed an increase in top-down control. Fixation durations decreased with age for all viewing conditions but not at the same rate. Furthermore, individual differences between long and short lookers were found across visits and viewing conditions.

The data and the results from the eye-tracking studies constituted the basis for the simulation studies where I modelled fixation data using the CRISP model of fixation durations (Nuthmann et al., 2010). While the simulation studies were grounded on assumptions and the conclusions that were taken from them were mainly based in simulated data, these results provided a feasible interpretation of the data and contributed to understanding the mechanisms underlying fixation durations and saccadic control in infancy. For the first time the two-step notion of saccade programming was introduced in infants, which states that saccades are programmed in two consecutive stages: the labile stage, during which the saccade program can still be cancelled, and the subsequent non-labile stage, where the saccade program can no longer be cancelled. Furthermore, model simulations suggested that at least until 10 months, infant fixation durations are influenced by both the developmental state of the visuo-motor system and

information processing speed. The simulations of the data from short and long lookers highlighted the huge developmental differences that can be found across infants.

On a methodological level, I was able to develop GraFIX, a new method and software to detect fixations in low and high quality eye-tracking data. As the validation analyses presented in Chapter 3 demonstrated, this method provides a better alternative for detecting fixations compared to previous automatic and manual approaches.

In conclusion, the present work provided valuable insights into the development of the mechanisms underlying fixation durations and saccadic control in infancy, an area of investigation that has been underrepresented in the research literature mainly due to the many practical and technical constraints that testing infants entail. Furthermore, my thesis demonstrates the benefits of mixing behavioural and computational methods to investigate realms of research not easily accessible through experimentation, such as oculomotor control in infancy.

Bibliography

Addyman, C., & Mareschal, D. (2010). The perceptual origins of the abstract same/different concept in human infants. *Animal Cognition*, *13*(6), 817–833.

Adolph, K. E., & Berger, S. E. (2006). Motor development. In W. Damon, R. Lerner, D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology: Vol 2: Cognition, perception and language* (6th ed., pp. 161–213). New York: Wiley.

Althoff, R., & Cohen, N. J. (1999). Eye-movement-based memory effect: a reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 997–1010.

Andersen, R. A., Batista, A. P., Snyder, L. H., Buneo, C. A., & Cohen, Y. E. (2000). Programming to look and reach in the posterior parietal cortex. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 515–524). Cambridge, MA: MIT press.

Anderson, A. J., & Johnson, C. A. (2006). Comparison of the ASA, MOBS, and ZEST threshold methods. *Vision Research*, *46*(15), 2403–2411.

Andrews, T. J., & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, *39*(17), 2947–2953.

Anliker, J. (1976). Eye movements: On-line measurement, analysis, and control. In R. S. Monty & J. W. Senders (Eds.), *Eye Movements and Psychological Processes* (pp. 185–199). Hillsdale, NJ: Lawrence Erlbaum Associates.

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*(1), 62.

Appel, M. A., & Campos, J. J. (1977). Binocular disparity as a discriminable stimulus parameter for young infants. *Journal of Experimental Child Psychology*, *23*(1), 47–56.

Aslin, R. N. (1981). Development of smooth pursuit in human infants. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 31–51). Hillsdale, NJ: Eribaum.

Aslin, R. N. (1987). Visual and auditory development in infancy. In J. D. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 5–97). New York: Wiley.

Aslin, R. N. (1993). Infant accommodation and convergence. In K. Simons (Ed.), *Early visual development: Normal and abnormal* (pp. 30–38). New York: National Research Council.

Aslin, R. N. (2007). What's in a Look? *Developmental Science*, *10*(1), 48–53.

Aslin, R. N. (2012). Infant Eyes: A Window on Cognitive Development. *Infancy*, *17*(1), 126–140.

Aslin, R. N., & McMurray, B. (2004). Automated corneal-reflection eye tracking in infancy: Methodological developments and applications to cognition. *Infancy*, *6*(2), 155–163. doi:10.1207/s15327078in0602_1

Aslin, R. N., & Salapatek, P. (1975). Saccadic localization of visual targets by the very young human infant. *Perception & Psychophysics*, *17*(3), 293–302.

Atkinson, J. (2000). *The Developing Visual Brain*. Oxford: Oxford University Press.

Atkinson, J., Hood, B., Wattam-Bell, J., & Braddick, O. (1992). Changes in infants' ability to switch visual attention in the first three months of life. *Perception*, *21*, 643–653.

Baddeley, R. J., & Tatler, B. W. (2006). High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis. *Vision Research*, *46*(18), 2824–2833.

Baillargeon, R. (1987). Young infants' reasoning about the physical and spatial properties of a hidden object. *Cognitive Development*, *2*(3), 179–200.

Baillargeon, R., & Graber, M. (1988). Evidence of location memory in 8-month-old infants in a nonsearch AB task. *Developmental Psychology*, *24*(4), 502.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, 1278–1289.

Baldwin, D. A., & Moses, L. J. (1996). The ontogeny of social information gathering. *Child Development*, *67*(5), 1915–1939.

Balkenius, C. (2000). Attention, habituation and conditioning: Toward a computational model. *Cognitive Science Quarterly*, *1*(2), 171–214.

Banks, M. S., & Shannon, E. (1993). Spatial and chromatic visual efficiency in human neonates. In C. E. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 1–46). Hillsdale, NJ: Erlbaum.

Bayley, N. (1993). *Bayley scales of infant development: Manual*. Psychological Corporation.

Becker, W., & Jurgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, *19*(9), 967–983.

Bell, M. A. (1998). Frontal lobe function during infancy: Implications for the development of cognition and attention. In J. E. Richards (Ed.), *Cognitive neuroscience of attention: a developmental perspective* (pp. 287–316). Mahwah, NJ: Erlbaum.

Benson, V., Castelhano, M. S., Au-Yeung, S. K., & Rayner, K. (2012). Eye movements reveal no immediate "WOW"("which one's weird") effect in autism spectrum disorder. *The Quarterly Journal of Experimental Psychology*, *65*(6), 1139–1150.

Berg, D. J., Boehnke, S. E., Marino, R. A., Munoz, D. P., & Itti, L. (2009). Free viewing of dynamic stimuli by humans and monkeys. *Journal of Vision*, *9*(5), 19.

Blaga, O. M., & Colombo, J. (2006). Visual processing and infant ocular latencies in the overlap paradigm. *Developmental Psychology*, *42*(6), 1069–1076.

Blignaut, P. (2009). Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, *71*(4), 881–895.

Blignaut, P., & Wium, D. (2014). Eye-tracking data quality as affected by ethnicity and experimental design. *Behavior Research Methods*, *46*(1), 67–80.

Boccignone, G., & Ferraro, M. (2004). Modelling gaze shift as a constrained random walk. *Physica A: Statistical Mechanics and Its Applications*, *331*(1-2), 207–218.

Borji, A., & Itti, L. (2013). State-of-the-Art in Visual Attention Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 185–207.

Bornstein, M. H., & Sigman, M. D. (1986). Continuity in mental development from infancy. *Child Development*, *57*, 251–274.

Braddick, O., & Atkinson, J. (2011). Vision Research. *Vision Research*, *51*(13), 1588–1609.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge: MIT press.

Bronson, G. W. (1972). Infants' reactions to unfamiliar persons and novel objects. *Monographs of the Society for Research in Child Development*, *37*(3, Serial No. 148).

Bronson, G. W. (1974). The postnatal growth of visual capacity. *Child Development*, *45*, 873–890.

Bronson, G. W. (1990). Changes in infants' visual scanning across the 2-to 14-week age period. *Journal of Experimental Child Psychology*, *49*(1), 101–125.

Bronson, G. W. (1994). Infants' Transitions toward Adult-like Scanning. *Child Development*, *65*(5), 1243–1261.

Bugajska, M., Trafton, J. G., Fransen, B., & Harrison, A. M. (2009). An Embodied Model of Infant Gaze-Following. In *International Conference on Cognitive Modeling*.

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature*, *371*(6497), 511–513.

Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art.* Chicago: University of Chicago Press.

Butcher, P. R., Kalverboer, A. F., & Geuze, R. H. (2000). Infants' shifts of gaze from a central to a peripheral stimulus: A longitudinal study of development between 6 and 26 weeks. *Infant Behavior and Development*, *23*(1), 3–21.

Butterworth, G. (2004). Joint visual attention in infancy. In A. Slater & G. Bremner (Eds.), *Theories of infant development* (pp. 317–354). Oxford, UK: Wiley–Blackwell.

Canfield, R. L., & Ceci, S. J. (1992). Incorporating learning into a theory of intellectual development. In R. J. Sternberg & C. A. Berg (Eds.), *Intellectual develop- ment.* New York: Cambridge Univ Press.

Canfield, R. L., Wilken, J., Schmerl, L., & Smith, E. G. (1995). Age-related change and stability of individual differences in infant saccade reaction time. *Infant Behavior and Development*, *18*(3), 351–358.

Carlson, E., & Triesch, J. (2004). A computational model of the emergence of gaze following. *Progress in Neural Processing*, *15*, 105–114.

Carpenter, R. H. S. (1988). *Movements of the eyes.* London: Pion Limited.

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, *377*(6544), 59–62.

Castelhano, M. S., & Henderson, J. M. (2008). Stable individual differences across images in human saccadic eye movements. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *62*(1), 1–14.

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6.

Choi, J. E. S., & Vaswani, P. A. (2014). Vigor of movements and the cost of time in decision making. *The Journal of Neuroscience*, *34*(4), 1212–1223.

Cohen, L. B., & Gelber, E. R. (1975). Infant visual memory. In P. Salapatek (Ed.), *Infant perception: From sensation to cognition. Volume I: Basic visual processes* (Vol. 1, pp. 347–403). New York: Academic Press, Inc.

Colombo, J. (1993). *Infant cognition: Predicting later intellectual functioning* (Vol. 5). Sage publications.

Colombo, J. (1995). On the neural mechanisms underlying developmental and individual differences in visual fixation in infancy: Two hypotheses. *Developmental Review*, *15*(2), 97–135.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*(1), 337–367.

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior*, *34*, 283–322.

Colombo, J., Freeseman, L. J., Coldren, J. T., & Frick, J. E. (1995). Individual differences in infant fixation duration: Dominance of global versus local stimulus properties. *Cognitive Development*, *10*(2), 271–285.

Colombo, J., & Mitchell, D. W. (1990). Individual differences in early visual attention: Fixation time and information processing. In J. Colombo & J. Fagen (Eds.), *Individual differences in infancy: Reliability, stability, prediction* (pp. 193–227). Hillsdale, NY: Lawrence Erlbaum Associates.

Colombo, J., Mitchell, D. W., Coldren, J. T., & Freeseman, L. J. (1991). Individual differences in infant visual attention: Are short lookers faster processors or feature processors? *Child Development*, *62*(6), 1247–1257.

Colombo, J., Mitchell, D. W., O'Brien, M., & Horowitz, F. D. (1987). The stability of visual habituation during the first year of life. *Child Development*, *58*, 474–487.

Colombo, J., Richman, W. a, Shaddy, D. J., Greenhoot, a F., & Maikranz, J. M. (2001). Heart rate-defined phases of attention, look duration, and infant performance in the paired-comparison paradigm. *Child Development*, *72*(6), 1605–16.

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: bottom-up versus top-down. *Current Biology*, *14*(19), R850–R852.

Corbetta, D., Guan, Y., & Williams, J. L. (2011). Infant Eye-Tracking in the Context of Goal-Directed Actions. *Infancy*, *17*(1), 102–125.

Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development*, *77*(3), 680–695.

Csibra, G., Tucker, L. A., & Johnson, M. H. (1998). Neural correlates of saccade planning in infants: A high-density ERP study. *International Journal of Psychophysiology*, *29*(2), 201–215.

Csibra, G., Tucker, L. A., & Johnson, M. H. (2001). Differential frontal cortex activation before anticipatory and reactive saccades in infants. *Infancy*, *2*(2), 159–174.

Cuevas, K., & Bell, M. A. (2013). Infant attention and early childhood executive function. *Child Development*, *85*(2), 397–404.

Delabarre, E. B. (1898). A method of recording eye-movements. *The American Journal of Psychology*, *9*(4), 572–574.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development*, *56*, 868–883.

Diamond, A., & Gilbert, J. (1989). Development as progressive inhibitory control of action: Retrieval of a contiguous object. *Cognitive Development*, *4*(3), 223–249.

Diamond, A., & Goldman-Rakic, P. S. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: Evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research*, *74*(1), 24–40.

Ditchburn, R. W. (1973). *Eye-movements and visual perception.* (pp. 97–98). Oxford: Clarendon Press.

Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review*, *8*(2), 145.

Domsch, H., Lohaus, A., & Thomas, H. (2010). Influences of information processing and disengagement in infants' looking behaviour. *Infant and Child Development*, *19*(2), 161–174.

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice* (Vol. 373). Springer.

Durand, F., & Dorsey, J. (2002). Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics*, *21*(3), 257–266.

Ellis, S. R., & Stark, L. (1986). Statistical dependency in visual scanning. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *28*(4), 421–438.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*(1), 71–99.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development* (Vol. 10). Cambridge: MIT press.

Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T., Johnson, M. H., & Team, T. B. A. S. of I. S. (2013). Disengagement of Visual Attention in Infancy is Associated with Emerging Autism in Toddlerhood. *Biological Psychiatry*, *74*(3), 189–194.

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., … Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: disengagement and facilitation. *Journal of Child Psychology and Psychiatry*, *50*(5), 637–642.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*(5), 621–636.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review*, *112*(4), 777–813.

Enright, J. T. (1998). Estimating peak velocity of rapid eye movements from video recordings. *Behavior Research Methods, Instruments, & Computers*, *30*(2), 349–353.

Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review*, *109*(3), 545.

Evdokimidis, I., Smyrnis, N., Constantinidis, T., Stefanis, N., Avramopoulos, D., Paximadis, C., … Stefanis, C. (2002). The antisaccade task in a sample of 2,006 young men. *Experimental Brain Research*, *147*(1), 45–52.

Fagan III, J. F. (1984). The relationship of novelty preferences during infancy to later intelligence and later recognition memory. *Intelligence*, *8*(4), 339–346.

Fantz, R. L. (1956). A method for studying early visual development. *Perceptual and Motor Skills*, *6*(g), 13–15.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science*, *140*(3564), 296–297.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, *146*(3644), 668–670.

Fantz, R. L., & Miranda, S. B. (1975). Newborn infant attention to form of contour. *Child Development*, *46*(1), 224–228.

Farroni, T., Simion, F., Umiltà, C., & Barba, B. D. (1999). The gap effect in newborns. *Developmental Science*, *2*(2), 174–186.

Feng, G. (2011). Eye Tracking: A Brief Guide for Developmental Researchers. *Journal of Cognition and Development*, *12*(1), 1–11.

Ferrera, V. P., Nealey, T. A., & Maunsell, J. H. R. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature*, *358*(6389), 756–758.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford, UK: Oxford University Press.

Findlay, J. M., & Harris, L. R. (1984). Small saccades to double-stepped targets moving in two dimensions. *Advances in Psychology*, *22*, 71–78.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, *22*(04), 661–674.

Fischer, B., & Boch, R. (1983). Saccadic eye movements after extremely short reaction times in the monkey. *Brain Research*, *260*(1), 21–26.

Fischer, B., & Breitmeyer, B. (1987). Mechanisms of visual attention revealed by saccadic eye movements. *Neuropsychologia*, *25*(1), 73–83.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, *57*(1), 191–195.

Fischer, B., & Weber, H. (1993). Express saccades and visual attention. *Behavioral and Brain Sciences*, *16*(03), 553–567.

Fox, N., Kagan, J., & Weiskopf, S. (1979). The growth of memory during infancy. *Genetic Psychology Monographs*, *99*, 91–130.

Franchak, J. M., & Adolph, K. E. (2010). Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. *Vision Research*, *50*(24), 2766–2774.

Franchak, J. M., Kretch, K. S., Soska, K. C., Babcock, J. S., & Adolph, K. E. (2010). Head-mounted eye-tracking of infants' natural interactions: A new method. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 21–27). ACM.

Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology*, *118*, 13–26.

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–170.

Frank, M. C., Vul, E., & Saxe, R. (2011). Measuring the Development of Social Attention Using Free-Viewing. *Infancy*, *17*(4), 355–375.

Frick, J. E., Colombo, J., & Saxon, T. F. (1999). Individual and developmental differences in disengagement of fixation in early infancy. *Child Development*, *70*(3), 537–48.

Friedburg, C., Allen, C. P., Mason, P. J., & Lamb, T. D. (2004). Contribution of cone photoreceptors and post-receptoral mechanisms to the human photopic electroretinogram. *The Journal of Physiology*, *556*(3), 819–834.

Frintrop, S., Rome, E., & Christensen, H. I. (2010). Computational visual attention systems and their cognitive foundations: A survey. *ACM Transactions on Applied Perception (TAP)*, *7*(1), 6.

Gilmore, R. O., & Johnson, M. H. (1997). Body-centered representations for visually-guided action emerge during early infancy. *Cognition*, *65*(1), B1–B10.

Glaholt, M., Rayner, K., & Reingold, E. (n.d.). A rapid and immediate effect of stimulus quality on the duration of individual fixations during reading. *Visual Cognition*.

Gopnik, A., & Tenenbaum, J. B. (2007). Bayesian networks, Bayesian learning and cognitive development. *Developmental Science*, *10*(3), 281–287.

Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye Tracking in Infancy Research. *Developmental Neuropsychology*, *35*(1), 1–19.

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research*, *58*(3), 455–472.

Hainline, L., Turkel, J., Abramov, I., Lemerise, E., & Harris, C. M. (1984). Characteristics of saccades in human infants. *Vision Research*, *24*(12), 1771–1780.

Haith, M. M. (1969). Infrared television recording and measurement of ocular behavior in the human infant. *American Psychologist*, *24*(3), 279.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23.

Hammoud, R. I. (2008). *Passive eye monitoring: Algorithms, applications and experiments*. New York: Springer-Verlag.

Harris, C. M., Hainline, L., Abramov, I., Lemerise, E., & Camenzuli, C. (1988). The distribution of fixation durations in infants and naive adults. *Vision Research*, *28*(3), 419–432.

Heide, W., Blankenburg, M., Zimmermann, E., & Kömpf, D. (1995). Cortical control of double-step saccades: implications for spatial orientation. *Annals of Neurology*, *38*(5), 739–748.

Heinke, D., & Humphreys, G. W. (2005). Computational models of visual selective attention: A review. In G. Houghton (Ed.), *Connectionist models in cognitive psychology* (Vol. 1, pp. 273–312). London: Psychology Press.

Held, R. (1993). Development of binocular vision revisited. In M. H. Johnson (Ed.), *Brain Development and Cognition: A Reader*. Oxford: Basil Blackwell.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford, UK: Elsevier.

Henderson, J. M., Chanceaux, M., & Smith, T. J. (2009). The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision*, *9*(1), 32.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*(1), 243–271.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review*, *16*(5), 850–856.

Henderson, J. M., Nuthmann, A., & Luke, S. G. (2013). Eye movement control during scene viewing: Immediate effects of scene luminance on fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 318.

Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin {&} Review*, *15*(3), 566–573.

Henderson, J. M., & Smith, T. J. (2009). How are eye fixation durations controlled during scene viewing? Further evidence from a scene onset delay paradigm. *Visual Cognition*, *17*(6-7), 1055–1082.

Henderson, J. M., Weeks Jr, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 210.

Henson, D. B. (1979). Investigation into corrective saccadic eye movements for refixation amplitudes of 10 degrees and below. *Vision Research*, *19*(1), 57–61.

Hirsch, J., & Curcio, C. A. (1989). The spatial resolution capacity of human foveal retina. *Vision Research*, *29*(9), 1095–1101.

Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 598.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye Tracking. A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: what it is and how to measure it. In *Proceedings of the symposium on eye tracking research and applications* (pp. 45–52). ACM.

Hood, B. M. (1993). Inhibition of return produced by covert shifts of visual attention in 6-month-old infants. *Infant Behavior and Development*, *16*(2), 245–254.

Hood, B. M., & Atkinson, J. (1990). Sensory visual loss and cognitive deficits in the selective attentional system of normal infants and neurologically impaired children. *Developmental Medicine & Child Neurology*, *32*(12), 1067–1077.

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development*, *16*(4), 405–422.

Hu, Y., & Walker, R. (2011). The neural basis of parallel saccade programming: An fmri study. *Journal of Cognitive Neuroscience*, *23*(11), 3669–3680.

Huey, E. B. (1898). Preliminary experiments in the physiology and psychology of reading. *The American Journal of Psychology*, *9*(4), 575–586.

Hunnius, S., & Geuze, R. H. (2004a). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy*, *6*(2), 231–255.

Hunnius, S., & Geuze, R. H. (2004b). Gaze shifting in infancy: a longitudinal study using dynamic faces and abstract stimuli. *Infant Behavior and Development*, *27*(3), 397–416.

Hunnius, S., Geuze, R. H., & van Geert, P. (2006). Associations between the developmental trajectories of visual scanning and disengagement of attention in infants. *Infant Behavior and Development*, *29*(1), 108–125.

Hunter, S. K., & Richards, J. E. (2011). Characteristics of Eye Movements to Sesame Street from 8 to 26 Weeks of Age. *Manuscript Submitted for Publication*.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29–53). Oxford: Elsevier Science Ltd. doi:10.1016/B978-008043361-5/50003-1

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10), 1489–1506.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203.

Jacobson, S. W., Jacobson, J. L., O'Neill, J. M., Padgett, R., Frankowski, J. J., & Bihun, J. T. (1992). Visual expectation and dimensions of infant information processing. *Child Development*, *63*(3), 711–724.

Jankowski, J. J., & Rose, S. A. (1997). The distribution of visual attention in infants. *Journal of Experimental Child Psychology*, *65*(2), 127–140.

Johnson, M. H. (1990). Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience*, *2*(2), 81–95.

Johnson, M. H. (1995a). The development of visual attention: A cognitive neuroscience perspective. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 735–747). Cambridge, MA: MIT press.

Johnson, M. H. (1995b). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology*, *28*(5), 281–291.

Johnson, M. H. (2003). Perceptual development is an acquired process. In E. E. Smith, R. L. Atkinson, & E. R. Hilgard (Eds.), *Atkinson & Hilgard's introduction to psychology*. Wadworth/Thomson Learning.

Johnson, M. H. (2011). *Developmental cognitive Neuroscience* (Third edit.). Oxford: Wiley-Blackwell.

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, *3*(4), 335–344.

Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where Infants Look Determines How They See: Eye Movements and Object Perception Performance in 3-Month-Olds. *Infancy*, *6*(2), 185–201.

Jost, T., Ouerhani, N., Wartburg, R. v, Müri, R., & Hügli, H. (2005). Assessing the contribution of color in visual attention. *Computer Vision and Image Understanding*, *100*(1), 107–123.

Kagan, J., McCall, R. B., Reppucci, N. D., Jordan, J., Levine, J., & Minton, C. (1971). *Change and continuity in infancy.* New York: Wiley.

Kalesnykas, R. P., & Hallett, P. E. (1994). Retinal eccentricity and the latency of eye saccades. *Vision Research*, *34*(4), 517–531.

Kano, F., Hirata, S., Call, J., & Tomonaga, M. (2011). The visual strategy specific to humans among hominids: A study using the gap-overlap paradigm. *Vision Research*, *51*(23-24), 2348–2355.

Kano, F., & Tomonaga, M. (2011a). Perceptual mechanism underlying gaze guidance in chimpanzees and humans. *Animal Cognition*, *14*(3), 377–386.

Kano, F., & Tomonaga, M. (2011b). Species difference in the timing of gaze movement between chimpanzees and humans. *Animal Cognition*, *14*(6), 879–892.

Kano, F., & Tomonaga, M. (2013). Head-mounted eye tracking of a chimpanzee under naturalistic conditions. *PloS One*, *8*(3), e59785.

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review*, *27*(3), 283–348.

Kaufman, J., Gilmore, R., & Johnson, M. H. (2006). Frames of reference for anticipatory action in 4-month-old infants. *Infant Behavior and Development*, *29*(3), 322–333.

Kavšek, M. (2004). Predicting later IQ from infant visual habituation and dishabituation: A meta-analysis. *Journal of Applied Developmental Psychology*, *25*(3), 369–393.

Keller, H., & Gauda, G. (1987). Eye contact in the first months of life and its developmental consequences. *Advances in Psychology*, *46*, 129–143.

Kelly, D. J., Walker, R., & Norbury, C. F. (2013). Deficits in volitional oculomotor control align with language status in autism spectrum disorders. *Developmental Science*, *16*(1), 56–66.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*(3), 307–321.

Kikuchi, Y., Senju, A., Akechi, H., Tojo, Y., Osanai, H., & Hasegawa, T. (2011). Atypical disengagement from faces and its modulation by the control of eye fixation in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *41*(5), 629–645.

Klahr, D., Langley, P., & Neches, R. (1984). *Production system models of learning and development*. Cambridge: MIT press.

Klahr, D., & MacWhinney, B. (1998). Information processing. *Handbook of Child Psychology*, *2*, 631–678.

Klahr, D., & Siegler, R. S. (1978). The representation of children's knowledge. *Advances in Child Development and Behavior*, *12*, 61–116.

Kolakowski, S. M., & Pelz, J. B. (2006). Compensating for eye tracker camera movement. In *Proceedings of the 2006 symposium on Eye tracking research & applications* (pp. 79–85). ACM.

Komogortsev, O. V, Gobert, D. V, Jayarathna, S., & Gowda, S. M. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *Biomedical Engineering, IEEE Transactions on*, *57*(11), 2635–2645.

König, S. D., & Buffalo, E. A. (2014). A nonparametric method for detecting fixations and saccades using cluster analysis: Removing the need for arbitrary thresholds. *Journal of Neuroscience Methods*, *227*, 121–131.

Kontsevich, L. L., & Tyler, C. W. (2004). What makes Mona Lisa smile? *Vision Research*, *44*(13), 1493–1498.

Kopecz, K. (1995). Saccadic reaction times in gap/overlap paradigms: A model based on integration of intentional and visual information on neural, dynamic fields. *Vision Research*, *35*(20), 2911–2925.

Kowler, E. (2011). Eye movements: The past 25years. *Vision Research*, *51*(13), 1457–1483.

Kurtzberg, D., & Vaughan, H. G. (1982). Topographic analysis of human cortical potentials preceding self-initiated and visually triggered saccades. *Brain Research*, *243*(1), 1–9.

Lacouture, Y., & Cousineau, D. (2008). How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Tutorials in Quantitative Methods for Psychology*, *4*(1), 35–45.

Lamb, M. E., Morrison, D. C., & Malkin, C. M. (1987). The development of infant social expectations in face-to-face interaction: A longitudinal study. *Merrill-Palmer Quarterly*, *33*, 241–254.

Land, M. F. (1999). Motion and vision: why animals move their eyes. *Journal of Comparative Physiology A*, *185*(4), 341–352.

Land, M. F. (2009). *Looking and acting: vision and eye movements in natural behaviour*. (B. W. Tatler, Ed.). Oxford: Oxford University Press.

Land, M. F., Furneaux, S. M., & Gilchrist, I. D. (2002). The organization of visually mediated actions in a subject without eye movements. *Neurocase*, *8*(1), 80–87.

Land, M. F., & Lee, D. N. (1994). Where do we look when we steer. *Nature*, *369*, 742–3.

Land, M. F., & Tatler, B. W. (2001). Steering with the head: The visual strategy of a racing driver. *Current Biology*, *11*(15), 1215–1220.

Larsson, L., Nyström, M., & Stridh, M. (2013). Detection of Saccades and Postsaccadic Oscillations in the Presence of Smooth Pursuit. *Biomedical Engineering, IEEE Transactions on*, *60*(9), 2484–2493.

Lawson, K. R., & Ruff, H. A. (2004). Early focused attention predicts outcome for children born prematurely. *Journal of Developmental and Behavioral Pediatrics*, *25*(6), 399–406.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265–288.

Lewandowsky, S., & Farrell, S. (2011). *Computational modeling in cognition: Principles and practice*. Thousand Oaks. CA: Sage.

Lewis, M., Goldberg, S., & Campbell, H. (1969). A developmental study of information processing within the first three years of life: Response decrement to a redundant signal. *Monographs of the Society for Research in Child Development*, *34*(9, Serial No. 133).

Lewis, T. L., & Maurer, D. (1992). The development of the temporal and nasal visual fields during infancy. *Vision Research*, *32*(5), 903–911.

Liu, S., Quinn, P. C., Wheeler, A., Xiao, N., Ge, L., & Lee, K. (2011). Similarity and difference in the processing of same-and other-race faces as revealed by eye tracking in 4-to 9-month-olds. *Journal of Experimental Child Psychology*, *108*(1), 180–189.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *The Journal of Neuroscience*, *7*(11), 3416–3468.

Loftus, G. R. (1985). Picture perception: Effects of luminance on available information and information-extraction rate. *Journal of Experimental Psychology: General*, *114*(3), 342.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 565.

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, *10*(2), 1–11.

Mallin, B. M., & Richards, J. E. (2012). Peripheral Stimulus Localization by Infants of Moving Stimuli on Complex Backgrounds. *Infancy*, *17*(6), 692–714.

Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1997). Fixation sequences made during visual examination of briefly presented 2D images. *Spatial Vision*, *11*(2), 157–178.

Mareschal, D. (2010). Computational perspectives on cognitive development. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(5), 696–708.

Mareschal, D., & French, R. M. (1997). A connectionist account of interference effects in early infant memory and categorization. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the 19th Annual Cognitive Science Society Conference* (pp. 484–489). London: Eribaum.

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M., Thomas, M. S. C., & Westermann, G. (2007). Neuroconstructivism-I: How the brain constructs cognition.

Mareschal, D., & Johnson, S. P. (2002). Learning to perceive object unity: A connectionist account. *Developmental Science*, *5*(2), 151–172.

Mareschal, D., Plunkett, K., & Harris, P. (1999). A computational and neuropsychological account of object-oriented behaviours in infancy. *Developmental Science*, *2*(3), 306–317.

Mareschal, D., & Shultz, T. R. (1996). Generative connectionist networks and constructivist cognitive development. *Cognitive Development*, *11*(4), 571–603.

Martinez-Conde, S. (2005). The Role of Eye Movements During Visual Fixation. *Barrow Quarterly*, *21*(3), 44–48.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, *5*(3), 229–240.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychological Bulletin*, *81*(12), 899.

Matsuzawa, M., & Shimojo, S. (1997). Infants' fast saccades in the gap paradigm and development of visual attention. *Infant Behavior and Development*, *20*(4), 449–455.

Mayes, L. C., & Kessen, W. (1989). Maturational changes in measures of habituation. *Infant Behavior and Development*, *12*(4), 437–450.

McAuley, J. H., Rothwell, J. C., & Marsden, C. D. (1999). Human anticipatory eye movements may reflect rhythmic central nervous activity. *Neuroscience*, *94*(2), 339–350.

McCall, R. B., Hogarty, P. S., Hamilton, J. S., & Vincent, J. H. (1973). Habituation rate and the infant's response to visual discrepancies. *Child Development*, *44*, 280–287.

McConkie, G. W. (1981). Evaluating and reporting data quality in eye movement research. *Behavior Research Methods {&} Instrumentation*, *13*(2), 97–106.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*(6), 578–586.

McMurray, B., Aslin, R. N., & Toscano, J. C. (2009). Statistical learning of phonetic categories: insights from a computational approach. *Developmental Science*, *12*(3), 369–378.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831.

McSorley, E., Haggard, P., & Walker, R. (2009). The spatial and temporal shape of oculomotor inhibition. *Vision Research*, *49*(6), 608–614.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, *16*(1), 369–402.

Mesulam, M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology*, *10*(4), 309–325.

Miles, F. A. (1998). The neural processing of 3-D visual information: evidence from eye movements. *European Journal of Neuroscience*, *10*(3), 811–822.

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *The Journal of Neuroscience*, *16*(16), 5154–5167.

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2010). Clustering of Gaze During Dynamic Scene Viewing is Predicted by Motion. *Cognitive Computation*, *3*(1), 5–24.

Mohler, C. W., & Wurtz, R. H. (1977). Role of striate cortex and superior colliculus in visual guidance of saccadic eye movements in monkeys. *Journal of Neurophysiology*, *40*(1), 74–94.

Mokler, A., & Fischer, B. (1999). The recognition and correction of involuntary prosaccades in an antisaccade task. *Experimental Brain Research*, *125*(4), 511–516.

Moore, C., & Corkum, V. (1998). Infant gaze following based on eye direction. *British Journal of Developmental Psychology*, *16*(4), 495–503.

Morris, T., Blenkhorn, P., & Zaidi, F. (2002). Blink detection for real-time eye tracking. *Journal of Network and Computer Applications*, *25*(2), 129–143.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 667.

Nakagawa, A., & Sukigara, M. (2007). Infant eye and head movements toward the side opposite the cue in the anti-saccade paradigm. *Behavioral and Brain Functions*, *3*(1), 5.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, *45*(2), 205–231.

Navalpakkam, V., & Itti, L. (2006). An integrated model of top-down and bottom-up attention for optimizing detection speed. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* (Vol. 2, pp. 2049–2056).

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* San Francisco: Freeman.

Nelson, C. A. (1995). The ontogeny of human memory: A cognitive neuroscience perspective. *Developmental Psychology*, *31*(5), 723.

Newell, A. (1994). *Unified theories of cognition*. Harvard University Press.

Noton, D., & Stark, L. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*(9), 929–IN8.

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research*, *49*(3), 322–336.

Nuthmann, A., & Henderson, J. M. (2012). Using CRISP to model global characteristics of fixation durations in scene viewing and reading with a common mechanism. *Visual Cognition*, *20*(4-5), 457–494.

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A Computational Model of Fixation Durations in Scene Viewing. *Psychological Review*, *117*(2), 382–405.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, *42*(1), 188–204.

Oakes, L. M. (2010). Infancy Guidelines for Publishing Eye-Tracking Data. *Infancy*, *15*(1), 1–5.

Oakes, L. M. (2012). Advances in Eye Tracking in Infancy Research. *Infancy*, *17*(1), 1–8.

Ogawa, T., & Komatsu, H. (2004). Target selection in area V4 during a multidimensional visual search task. *The Journal of Neuroscience*, *24*(28), 6371–6382.

Papageorgiou, K., Smith, T. J., Wu, R., Johnson, M. H., Kirkham, N. Z., & Ronald, A. (2014). Individual Differences in Infant Fixation Duration relate to Attention and Behavioral Control in Childhood. *Psychological Science*, *25*(7), 1371–1379. doi:10.1177/0956797614531295

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107–123.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*(3), 302–321.

Posner, M. I. (1988). Structures and function of selective attention. In T. Boll & B. Bryant (Eds.), *Clinical neuropsychology and brain function: Research, measures and practice*. Washington, DC: American Psychological Association.

Quinn, P. C., & Eimas, P. D. (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of Experimental Child Psychology*, *63*(1), 189–211.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*(3), 190.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372.

Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2009). Eye movements and visual encoding during scene perception. *Psychological Science*, *20*(1), 6–10.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*(4), 445–476.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the EZ Reader model. *Psychological Science*, *17*(9), 742–746.

Reuter-Lorenz, P. A., Hughes, H. C., & Fendrich, R. (1991). The reduction of saccadic latency by prior offset of the fixation point: an analysis of the gap effect. *Perception & Psychophysics*, *49*(2), 167–175.

Reuter-Lorenz, P. A., Oonk, H. M., Barnes, L. L., & Hughes, H. C. (1995). Effects of warning signals and fixation point offsets on the latencies of pro-versus antisaccades: implications for an interpretation of the gap effect. *Experimental Brain Research*, *103*(2), 287–293.

Reynolds, G. D., & Richards, J. E. (2007). Infant heart rate: A developmental psychophysiological perspective. In L. Schmidt & S. Segalowitz (Eds.), *Developmental Psychophysiology: Theory, Systems, and Applications*. Cambridge Univ Press.

Richards, J. E., & Holley, F. B. (1999). Infant attention and the development of smooth pursuit tracking. *Developmental Psychology*, *35*(3), 856–867.

Richardson, D. C., Dale, R., & Spivey, M. J. (2007). Eye movements in language and cognition: A brief introduction. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (Vol. 18, pp. 323–344). Amsterdam: John Benjamins.

Richardson, D. C., & Johnson, S. P. (2008). Eye tracking research in infants and adults. In V. Sloutsky, B. Love, & K. McRae (Eds.), *30th Annual meeting of the cognitive science society: CogSci 2008.* (pp. 23–26). Washington, DC.

Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: evidence from eye tracking. *Developmental Science*, *12*(4), 549–556.

Robertson, S. S. (1982). Intrinsic temporal patterning in the spontaneous movement of awake neonates. *Child Development*, *53*, 1016–1021.

Rolfs, M. (2009). Microsaccades: small steps on a long way. *Vision Research*, *49*(20), 2415–2441.

Rommelse, N. N. J., der Stigchel, S., & Sergeant, J. A. (2008). A review on eye movement studies in childhood and adolescent psychiatry. *Brain and Cognition*, *68*(3), 391–414.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2002). Processing speed in the 1st year of life: a longitudinal study of preterm and full-term infants. *Developmental Psychology*, *38*(6), 895.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Caro, D. M. (2002). A longitudinal study of visual expectation and reaction time in the first year of life. *Child Development*, *73*(1), 47–61.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences*, *24*(2), 113–121.

Ross-sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*(6), 1807–1822.

Rothbart, M. K. (1988). Temperament and the development of inhibited approach. *Child Development*, *59*, 1241–1250.

Rothbart, M. K., Posner, M. I., & Rosicky, J. (1994). Orienting in normal and pathological development. *Development and Psychopathology*, *6*(04), 635–652.

Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S. J., & Passingham, R. E. (2000). The prefrontal cortex: response selection or maintenance within working memory? *Science*, *288*(5471), 1656–1660.

Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development: Themes and variations*. Oxford: Oxford University Press.

Ruff, H. A., & Saltarelli, L. M. (1993). Exploratory play with objects: Basic cognitive processes and individual differences. *New Directions for Child and Adolescent Development*, *1993*(59), 5–16.

Ruff, H. A., Saltarelli, L. M., Capozzoli, M., & Dubiner, K. (1992). The differentiation of activity in infants' exploration of objects. *Developmental Psychology*, *28*(5), 851.

Ruff, H. A., & Turkewitz, G. (1975). Developmental changes in the effectiveness of stimulus intensity on infant visual attention. *Developmental Psychology*, *11*(6), 705.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning internal representations by error propagation. In *Parallel distributed processing: Explorations in the Microstructure of Cognition no. 2: Psychological and biological models* (pp. 318–362). Cambridge: MIT press.

Rumelhart, D. E., & McClelland, J. L. (1985). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & P. Research group (Eds.), *Parallel distributed processing: Explorations in the Microstructure of Cognition no. 2: Psychological and biological models*. Cambridge: MIT press.

Saez de Urabain, I. R., Johnson, M. H., & Smith, T. J. (2014). GraFIX: A semiautomatic approach for parsing low- and high-quality eye-tracking data. *Behavior Research Methods*, *47*(1), 53–72. doi:10.3758/s13428-014-0456-0

Salapatek, P. (1975). Pattern perception in early infancy. *Infant Perception: From Sensation to Cognition*, *1*, 133–248.

Salapatek, P., Bechtold, A. G., & Bushnell, E. W. (1976). Infant Visual Acuity as a Function of Viewing Distance. *Child Development*, *47*, 860–863.

Salapatek, P., & Kessen, W. (1966). Visual scanning of triangles by the human newborn. *Journal of Experimental Child Psychology*, *3*(2), 155–167.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications* (pp. 71–78). ACM.

Sanders, A. F. (1963). *The selective process in the functional visual field*. Assen: Van Gorcum.

Saslow, M. G. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *JOSA*, *57*(8), 1024–1029.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. CRC press.

Schall, J. D. (1991). Neuronal basis of saccadic eye movements. *Vision and Visual Dysfunction*, *4*, 388–442.

Schall, J. D. (1995). Neural basis of saccade target selection. *Reviews in the Neurosciences*, *6*(1), 63–85.

Schiller, P. H. (1985). A model for the generation of visually guided saccadic eye movements. In D. Rose & V. G. Dobson (Eds.), *Models of the visual cortex* (pp. 62–70). Chicester: Wiley.

Schiller, P. H. (1998). The neural control of visually guided eye movements. In J. E. Richards (Ed.), *Cognitive neuroscience of attention* (pp. 3–50). Engle-wood Cliffs,NJ: Lawrence Erlbaum.

Schiller, P. H., Logothetis, N. K., & Charles, E. R. (1991). Parallel pathways in the visual system: their role in perception at isoluminance. *Neuropsychologia*, *29*(6), 433–441.

Schlesinger, M., Amso, D., & Johnson, S. P. (2007). The Neural Basis for Visual Selective Attention in Young Infants: A Computational Account. *Adaptive Behavior*, *15*(2), 135–148.

Schlesinger, M., & McMurray, B. (2012). The past, present, and future of computational models of cognitive development. *Cognitive Development*, *27*(4), 326–348.

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *2*(4), 419–436.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–671.

Senju, A., Kikuchi, Y., Akechi, H., Hasegawa, T., Tojo, Y., Osanai, H., & Johnson, M. H. (2011). Atypical modulation of face-elicited saccades in autism spectrum disorder in a double-step saccade paradigm. *Research in Autism Spectrum Disorders*, *5*(3), 1264–1269.

Shaddy, D. J., & Colombo, J. (2004). Developmental changes in infant attention to dynamic and static stimuli. *Infancy*, *5*(3), 355–365.

Shic, F., Scassellati, B., & Chawarska, K. (2008). The incomplete fixation measure. In *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 111–114).

Shukla, M., Wen, J., White, K. S., & Aslin, R. N. (2011). SMART-T: A system for novel fully automated anticipatory eye-tracking paradigms. *Behavior Research Methods*, *43*(2), 384–398.

Shultz, T. R., Mareschal, D., & Schmidt, W. C. (1994). Modeling cognitive development on balance scale phenomena. *Machine Learning*, *16*(1-2), 57–86.

Sigman, M. D., Cohen, S. E., Beckwith, L., & Parmelee, A. H. (1986). Infant attention in relation to intellectual abilities in childhood. *Developmental Psychology*, *22*(6), 788.

Sirois, S., & Mareschal, D. (2002). Models of habituation in infancy. *Trends in Cognitive Sciences*, *6*(7), 293–298.

Slater, A. (1995). Individual differences in infancy and later IQ. *Journal of Child Psychology and Psychiatry*, *36*(1), 69–112.

Smeets, J. B. J., & Hooge, I. T. C. (2003). Nature of variability in saccades. *Journal of Neurophysiology*, *90*(1), 12–20.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, *33*(1), 5–42.

Smith, T. J., & Henderson, J. M. (2009). Facilitation of return during scene viewing. *Visual Cognition*, *17*(6-7), 1083–1108.

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology*, *21*(1), 195.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*(1), 29–56.

Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers*, *25*(2), 137–142.

Stanley, J. C. (1976). Computer simulation of a model of habituation. *Nature*, *261*, 146–148.

Staub, A., & Benatar, A. (2013). Individual differences in fixation duration distributions in reading. *Psychonomic Bulletin & Review*.

Stoesz, B. M., & Jakobson, L. S. (2014). Developmental changes in attention to faces and bodies in static and dynamic scenes. *Frontiers in Psychology*, *5*.

Stratton, G. M. (1902). Eye-Movements and the Aesthetics of Visual Form. *Philosophische Studien*, *14*, 82–96.

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, *45*(5), 643–659.

Tatler, B. W., Gilchrist, I. D., & Land, M. F. (2005). Visual memory for objects in natural scenes: From fixations to object files. *The Quarterly Journal of Experimental Psychology Section A*, *58*(5), 931–960.

Thomas, M. S. C. (2003). Multiple causality in developmental disorders: Methodological implications from computational modelling. *Developmental Science*, *6*(5), 537–556.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological Review*, *113*(4), 766.

Triesch, J., Teuscher, C., Deák, G. O., & Carlson, E. (2006). Gaze following: why (not) learn it? *Developmental Science*, *9*(2), 125–147.

Trukenbrod, H. A., & Engbert, R. (2014). ICAT: A computational model for the adaptive control of fixation durations. *Psychonomic Bulletin & Review*, 1–28.

Tummeltshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants track the reliability of potential informants. *Psychological Science*, *25*(9), 1730–1738.

Tyrrell, R. A., & Owens, D. A. (1988). A rapid technique to assess the resting states of the eyes and other threshold phenomena: the modified binary search (MOBS). *Behavior Research Methods*, *20*(2), 137–141.

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *The Quarterly Journal of Experimental Psychology*, *59*(11), 1931–1949.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. Ingle, M. Goodale, & R. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT press.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, *16*(3), 219–242.

Van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, *18*(6), 681–694.

Võ, M. L. H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, *9*(3), 24.

Walden, T. A., & Ogan, T. A. (1988). The development of social referencing. *Child Development*, *59*, 1230–1240.

Walker, R., & Benson, V. (2013). Remote distractor effects and saccadic inhibition: Spatial and temporal modulation. *Journal of Vision*, *13*(11), 9.

Walker, R., Husain, M., Hodgson, T. L., Harrison, J., & Kennard, C. (1998). Saccadic eye movement and working memory deficits following damage to human prefrontal cortex. *Neuropsychologia*, *36*(11), 1141–1159.

Walker, R., & McSorley, E. (2006). The parallel programming of voluntary and reflexive saccades. *Vision Research*, *46*(13), 2082–2093.

Walshe, R. C., & Nuthmann, A. (2013). Programming of saccades to double-step targets in scene viewing: A test of assumptions present in the CRISP model. In *In Proceedings of the 35th Annual Meeting of the Cognitive Science Society, Berlin, Germany* (pp. 1569–1574).

Wass, S. V, Jones, E. J. H., Gliga, T., Smith, T. J., Charman, T., Johnson, M. H., & others. (2015). Shorter spontaneous fixation durations in infants with later emerging autism. *Scientific Reports*, *5*.

Wass, S. V, Porayska-Pomsta, K., & Johnson, M. H. (2011). Training Attentional Control in Infancy. *Current Biology*, *21*(18), 1543–1547.

Wass, S. V, & Smith, T. J. (2014). Individual Differences in Infant Oculomotor Behavior During the Viewing of Complex Naturalistic Scenes. *Infancy*, *19*(4), 352–384.

Wass, S. V, Smith, T. J., & Johnson, M. H. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, *45*(1), 229–250.

West, B. T. (2009). Analyzing longitudinal data with the linear mixed models procedure in SPSS. *Evaluation {&} the Health Professions*, *32*(3), 207–228.

Westermann, G., Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., & Thomas, M. S. C. (2007). Neuroconstructivism. *Developmental Science*, *10*(1), 75–83.

Westermann, G., & Ruh, N. (2012). A neuroconstructivist model of past tense development and processing. *Psychological Review*, *119*(3), 649.

Westermann, G., Ruh, N., & Plunkett, K. (2009). Connectionist approaches to language learning. *Linguistics*, *47*(2), 413–452.

Westheimer, G. (1954). Eye movement responses to a horizontally moving visual stimulus. *AMA Archives of Ophthalmology*, *52*(6), 932–941.

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., & Vogeley, K. (2010). It's in your eyes--using gaze-contingent stimulus to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, *5*(1), 98–107.

Wolff, P. H. (1987). *The development of behavioral states and the expression of emotions in early infancy: New proposals for investigation*. University of Chicago Press.

Wolff, P. H. (1991). Endogenous motor rhythms in young infants. *Advances in Psychology*, *81*, 119–133.

Wurtz, R. H., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology*, *107*(2), 118–136.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15–B25.

Yarbus, A. L. (1967). *Eye movements and vision* (Vol. 2). Plenum press New York.

Yonas, A., Arterberry, M. E., & Granrud, C. E. (1987). Four-month-old infants' sensitivity to binocular and kinetic information for three-dimensional-object shape. *Child Development*, *58*(4), 910–917.

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science*, *12*(5), 798–814.

Yuan, Y. C. (2010). Multiple imputation for missing data: Concepts and new development (Version 9.0). *SAS Institute Inc, Rockville, MD*.

Yuodelis, C., & Hendrickson, A. (1986). A qualitative and quantitative analysis of the human fovea during development. *Vision Research*, *26*(6), 847–855.

Appendix A

Pseudo-code of the proposed interpolation algorithm

1. Calculate the velocities for each point

2. Iteratively find velocity peaks and flag those over the velocity threshold as saccades

3. Find next gap

    3.1. If the gap is longer than the interpolation latency:

        3.1.1. Calculate the Euclidean distances from a central point for the fixations preceding and following the gap.

        3.1.2. If the difference between both Euclidean distances is lower than the maximum displacement threshold:

            3.1.2.1. Interpolate.

Pseudo-code of the proposed velocity algorithm and the post-doc validation

1. Calculate the velocities for each point

2. Iteratively find velocity peaks and flag those over the velocity threshold as saccades

3. The data points in between two saccades are grouped and flagged as fixations

4. Post-hoc validation: Merge adjacent fixations with similar location

    4.1. If this post-hoc validation criterion is selected

        4.1.1. Find next two adjacent fixations

        4.1.2. If the gap between the two fixations is shorter then 50 ms

            4.1.2.1. Calculate the distance in degrees between the locations for the two fixations

            4.1.2.2. If the distance is shorter than the maximum displacement threshold

                4.1.2.2.1. Merge both fixations

5. Post-hoc validation: The root mean square (RMS) of inter-sample distances per fixations does not exceed a threshold

    5.1. If this post-hoc validation criterion is selected

        5.1.1. Find all fixations with a RMS over the threshold

        5.1.2. Delete them

6. Post-hoc validation: Minimum fixation duration

    6.1. If this post-hoc validation criterion is selected

        6.1.1. Find all fixations with a duration over the threshold

6.1.2. Delete them

Appendix B

Detailed descriptions for the naturalistic videos from the cross-sectional study defined in Chapter 4.

Table Appendix 1 Descriptions for the naturalistic videos described in Chapter 4.

Video	Active state	Inactive state
1	The character is dancing and playing with a transparent bag full of colourful balls.	The character does not move.
2	The character is holding a red balloon with both hands and moving it around.	The character occludes her face with the balloon.
3	The character faces the camera and moves a yellow octopus that she is holding around.	The character has her back turned to the viewer.
4	The character is holding two balloons and is dancing with them.	The character does not move.
5	The character is looking at the viewing trying to attract his or her attention.	The character is occluding her face.
6	The character is dancing and moving a yellow toy around.	The character does not move.
7	The character is dancing behind a pink object.	The character is hidden behind the pink object.
8	The character is dancing and moving a red toy around.	The character does not move.
9	The character is dancing behind a pink object (different actresses as in video 7).	The character is hidden behind the pink object.
10	The character is dancing and playing with one balloon that she is holding with one hand.	The character does not move.

Figure Appendix 1 Naturalistic videos for the cross-sectional study described in chapter 4.

Appendix C

Table Appendix 2 Descriptions for the naturalistic videos described in chapter 6.

Video	Active state	Inactive state
1	The character is dancing and playing with a transparent bag full of colourful balls.	The character does not move.
2	The character is holding a red balloon with both hands and moving it around.	The character occludes her face with the balloon.
3	The character faces the camera and moves a yellow octopus that she is holding around.	The character has her back turned to the viewer.
4	The character is holding two balloons and is dancing with them.	The character does not move.
5	The character is looking to the front trying to attract his or her attention.	The character is occluding her face.
6	The character is dancing and moving a yellow toy around.	The character does not move.
7	The character is dancing behind a pink object.	The character is hidden behind the pink object.
8	The character is dancing and moving a red toy around.	The character does not move.
9	The character is dancing and playing with a transparent bag full of colourful balls.	The character does not move.
10	The character is dancing and playing with a transparent bag full of colourful balls.	The character does not move.
11	The character is holding two balloons and is dancing with them.	The character does not move.
12	The character is dancing and moving a yellow toy around.	The character does not move.

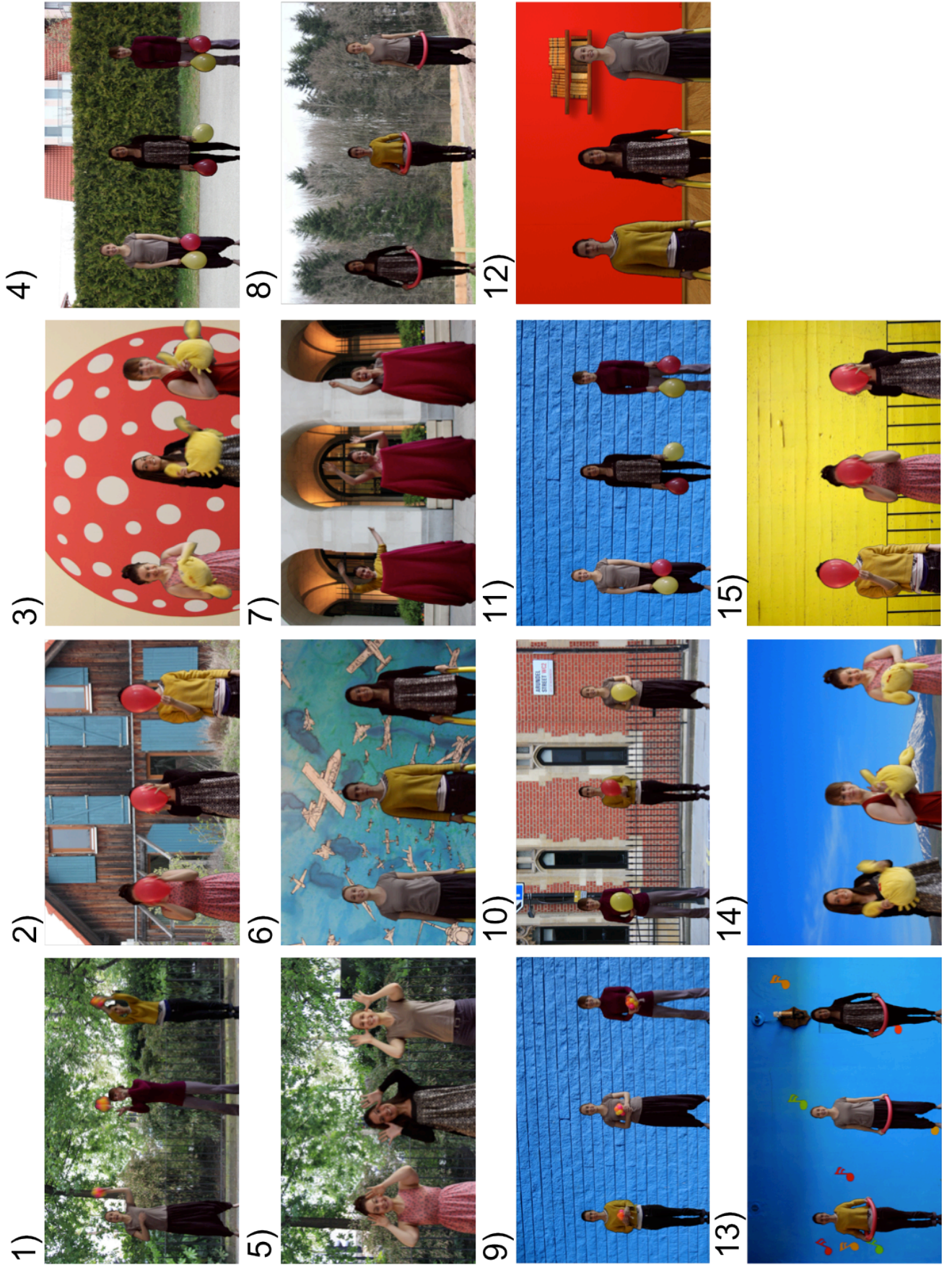| 13	The character is dancing and moving a red toy around.	The character does not move.
14	The character faces the camera and moves a yellow octopus that she is holding around.	The character has her back turned to the viewer.
15	The character is holding a red balloon with both hands and moving it around.	The character occludes her face with the balloon.

Figure Appendix 2 Naturalistic videos for the longitudinal studies described in chapter 6.
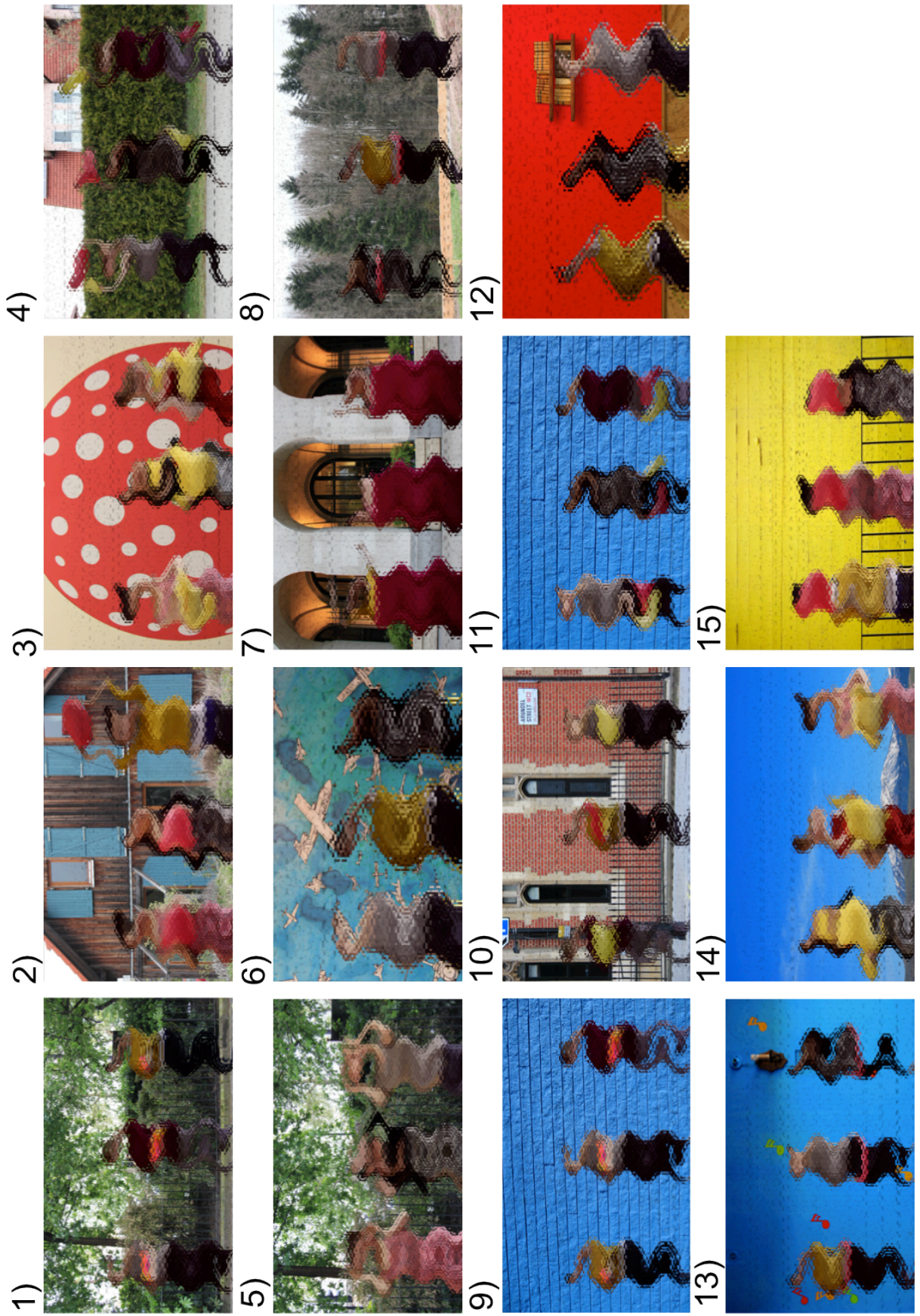
Figure Appendix 3 Abstract videos for the longitudinal studies described in chapter 6.
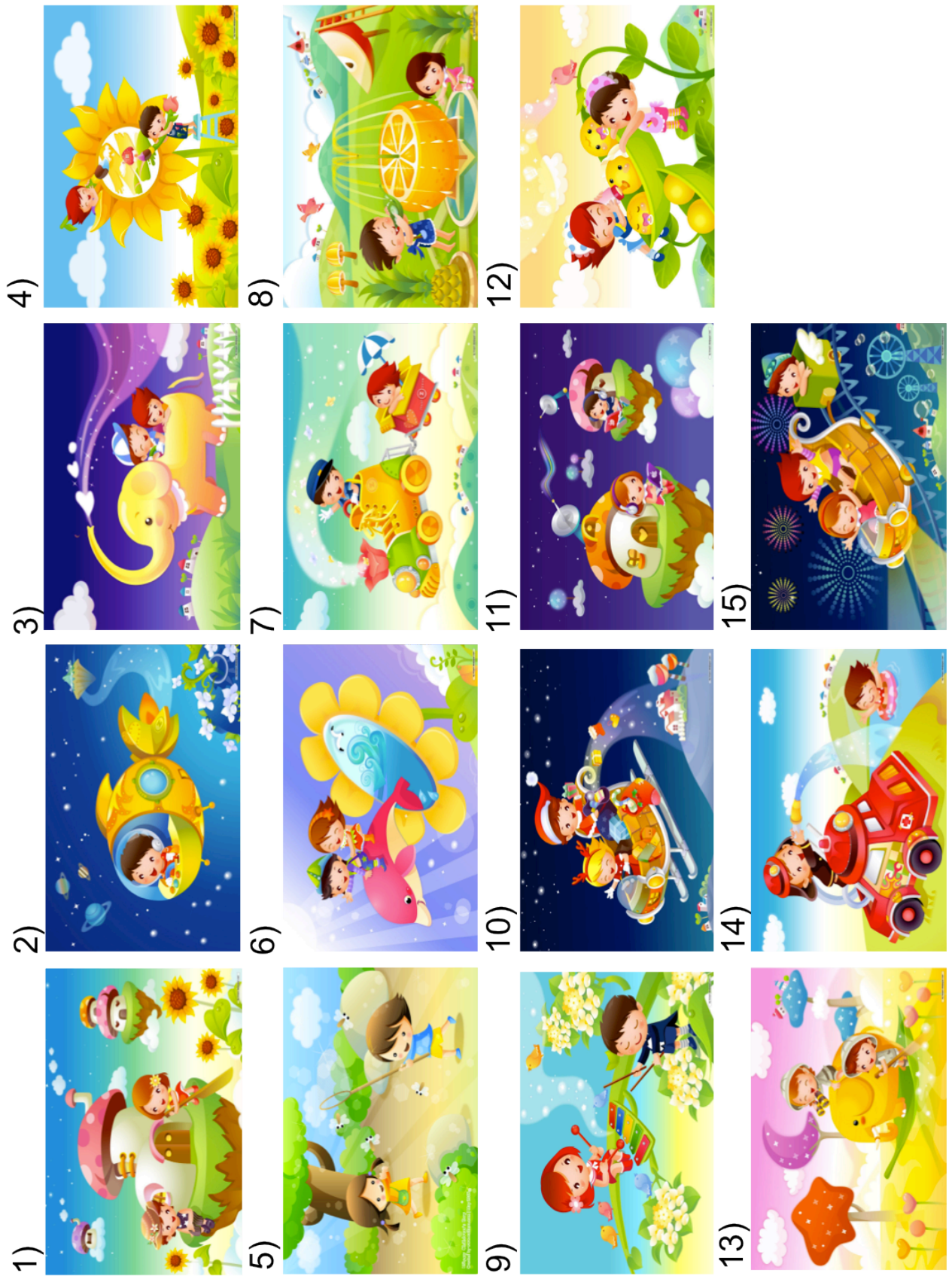
Figure Appendix 4 Static images for the longitudinal studies described in Chapter 6.

Appendix D

In Chapter 6, the data from the longitudinal study testing infants from 6 to 12 months was analysed using the MCMC multiple imputation technique. First, ten plausible synthetic values or imputations were generated for each missing data point, leading to 10 different data sets. Secondly, each data set was analysed with the corresponding ANOVA, which generated 10 slightly different and plausible results. Finally, the 10 estimates were pooled into a single estimate by calculating their means, combining the variations within and across the 10 imputed data sets.

Tables 3, 4, 5 and 6 show the descriptive statistics for mean fixation durations and the ex-Gaussian components (μ, σ, and τ) for the original data set, the data sets generated with the MCMC multiple imputation technique, and the pooled data.

Table Appendix 3 Descriptive statistics for mean fixation durations in naturalistic videos, abstract videos and static images from the original and imputed data sets.

	Imputation Number	Naturalistic videos			Abstract videos			Static images		
		N	Mean	Std.	N	Mean	Std.	N	Mean	Std.
Visit 1	Original data	18	.516	.053	18	.800	.219	18	.325	.039
	1	19	.512	.055	19	.764	.264	19	.324	.039
	2	19	.510	.057	19	.761	.273	19	.328	.040
	3	19	.512	.055	19	.793	.215	19	.330	.044
	4	19	.519	.053	19	.799	.213	19	.329	.041
	5	19	.515	.052	19	.829	.248	19	.328	.039
	6	19	.513	.053	19	.783	.225	19	.328	.040
	7	19	.517	.052	19	.808	.215	19	.327	.039
	8	19	.519	.053	19	.808	.216	19	.328	.041
	9	19	.506	.068	19	.796	.213	19	.332	.049
	10	19	.518	.052	19	.803	.213	19	.329	.041

		N			N			N		
	Pooled	19	.514		19	.794		19	.328	
Visit 2	Original data	16	.468	.045	16	.743	.187	16	.317	.039
	1	19	.467	.048	19	.775	.228	19	.318	.041
	2	19	.472	.045	19	.719	.227	19	.320	.040
	3	19	.471	.051	19	.739	.257	19	.320	.040
	4	19	.473	.044	19	.765	.219	19	.322	.043
	5	19	.489	.081	19	.814	.332	19	.319	.038
	6	19	.474	.045	19	.761	.282	19	.320	.040
	7	19	.481	.053	19	.805	.290	19	.319	.037
	8	19	.481	.054	19	.820	.285	19	.320	.038
	9	19	.490	.092	19	.839	.387	19	.316	.037
	10	19	.475	.047	19	.766	.253	19	.320	.039
	Pooled	19	.477		19	.780		19	.319	
Visit 3	Original data	18	.452	.046	18	.666	.161	18	.322	.041
	1	19	.440	.068	19	.635	.208	19	.321	.040
	2	19	.448	.048	19	.669	.157	19	.322	.040
	3	19	.455	.046	19	.650	.172	19	.323	.041
	4	19	.447	.050	19	.669	.157	19	.322	.040
	5	19	.444	.057	19	.648	.176	19	.322	.040
	6	19	.448	.049	19	.658	.160	19	.322	.040
	7	19	.455	.046	19	.678	.165	19	.322	.040
	8	19	.453	.045	19	.686	.178	19	.322	.040
	9	19	.453	.045	19	.665	.157	19	.322	.040
	10	19	.445	.055	19	.652	.168	19	.322	.040
	Pooled	19	.449		19	.661		19	.322	
Visit 4	Original data	19	.446	.057	19	.699	.157	19	.334	.045
	1	19	.446	.057	19	.699	.157	19	.334	.045

304

		N	Mean	Std.	N	Mean	Std.	N	Mean	Std.
	2	19	.446	.057	19	.699	.157	19	.334	.045
	3	19	.446	.057	19	.699	.157	19	.334	.045
	4	19	.446	.057	19	.699	.157	19	.334	.045
	5	19	.446	.057	19	.699	.157	19	.334	.045
	6	19	.446	.057	19	.699	.157	19	.334	.045
	7	19	.446	.057	19	.699	.157	19	.334	.045
	8	19	.446	.057	19	.699	.157	19	.334	.045
	9	19	.446	.057	19	.699	.157	19	.334	.045
	10	19	.446	.057	19	.699	.157	19	.334	.045
	Pooled	19	.446		19	.699		19	.334	

Table Appendix 4 Descriptive statistics for μ fixation durations in naturalistic videos, abstract videos and static images from the original and imputed data sets.

	Imputation Number	Naturalistic videos			Abstract videos			Static images		
		N	Mean	Std.	N	Mean	Std.	N	Mean	Std.
Visit 1	Original data	18	.248	.044	18	.214	.055	18	.227	.041
	1	19	.250	.044	19	.218	.056	19	.228	.040
	2	19	.247	.043	19	.212	.054	19	.230	.042
	3	19	.243	.047	19	.207	.063	19	.228	.040
	4	19	.270	.106	19	.230	.086	19	.250	.108
	5	19	.242	.049	19	.208	.060	19	.223	.042
	6	19	.247	.043	19	.218	.055	19	.232	.045
	7	19	.244	.045	19	.216	.053	19	.228	.040
	8	19	.246	.043	19	.214	.053	19	.231	.043
	9	19	.252	.047	19	.218	.055	19	.248	.100

	10	19	.255	.054	19	.213	.053	19	.244	.084
	Pooled	19	.250		19	.215		19	.234	
Visit 2	Original data	16	.218	.025	16	.202	.042	16	.214	.027
	1	19	.220	.027	19	.205	.044	19	.215	.026
	2	19	.216	.025	19	.196	.044	19	.212	.025
	3	19	.213	.028	19	.190	.053	19	.212	.027
	4	19	.234	.077	19	.221	.096	19	.234	.078
	5	19	.224	.033	19	.201	.041	19	.217	.028
	6	19	.222	.026	19	.209	.042	19	.213	.031
	7	19	.218	.025	19	.204	.048	19	.214	.026
	8	19	.216	.026	19	.196	.048	19	.212	.029
	9	19	.222	.059	19	.200	.063	19	.215	.046
	10	19	.228	.047	19	.207	.047	19	.219	.037
	Pooled	19	.221		19	.203		19	.216	
Visit 3	Original data	18	.220	.030	18	.224	.040	18	.219	.034
	1	19	.218	.031	19	.218	.045	19	.216	.037
	2	19	.221	.030	19	.225	.039	19	.221	.034
	3	19	.231	.054	19	.216	.051	19	.219	.033
	4	19	.221	.030	19	.226	.040	19	.219	.033
	5	19	.217	.034	19	.232	.054	19	.211	.049
	6	19	.225	.035	19	.223	.039	19	.216	.036
	7	19	.220	.030	19	.225	.039	19	.218	.033
	8	19	.220	.030	19	.230	.047	19	.220	.033
	9	19	.223	.032	19	.222	.039	19	.214	.041
	10	19	.228	.043	19	.223	.039	19	.225	.042
	Pooled	19	.222		19	.224		19	.218	
Visit4	Original data	19	.204	.020	19	.207	.032	19	.221	.046

		N	Mean	Std.	N	Mean	Std.	N	Mean	Std.
	1	19	.204	.020	19	.207	.032	19	.221	.046
	2	19	.204	.020	19	.207	.032	19	.221	.046
	3	19	.204	.020	19	.207	.032	19	.221	.046
	4	19	.204	.020	19	.207	.032	19	.221	.046
	5	19	.204	.020	19	.207	.032	19	.221	.046
	6	19	.204	.020	19	.207	.032	19	.221	.046
	7	19	.204	.020	19	.207	.032	19	.221	.046
	8	19	.204	.020	19	.207	.032	19	.221	.046
	9	19	.204	.020	19	.207	.032	19	.221	.046
	10	19	.204	.020	19	.207	.032	19	.221	.046
	Pooled	19	.204		19	.207		19	.221	

Table Appendix 5 Descriptive statistics for σ fixation durations in naturalistic videos, abstract videos and static images from the original and imputed data sets.

	Imputation Number	Naturalistic videos			Abstract videos			Static images		
		N	Mean	Std.	N	Mean	Std.	N	Mean	Std.
	Original data	18	.070	.027	18	.053	.040	18	.054	.019
	1	19	.074	.031	19	.053	.039	19	.056	.021
	2	19	.076	.036	19	.061	.053	19	.055	.020
	3	19	.070	.026	19	.054	.039	19	.054	.019
Visit 1	4	19	.070	.026	19	.054	.039	19	.054	.019
	5	19	.072	.027	19	.057	.043	19	.055	.020
	6	19	.071	.027	19	.054	.039	19	.053	.019
	7	19	.070	.026	19	.056	.041	19	.054	.019
	8	19	.072	.028	19	.051	.040	19	.054	.019

		N			N			N		
	9	19	.073	.030	19	.057	.042	19	.055	.019
	10	19	.069	.026	19	.054	.039	19	.052	.020
	Pooled	19	.072		19	.055		19	.054	
Visit 2	Original data	16	.046	.019	16	.044	.027	16	.043	.014
	1	19	.046	.018	19	.044	.025	19	.044	.017
	2	19	.051	.022	19	.052	.036	19	.043	.017
	3	19	.047	.019	19	.045	.029	19	.042	.014
	4	19	.048	.018	19	.036	.036	19	.040	.019
	5	19	.049	.020	19	.041	.043	19	.045	.018
	6	19	.050	.020	19	.048	.027	19	.042	.013
	7	19	.045	.017	19	.045	.031	19	.041	.015
	8	19	.051	.022	19	.050	.030	19	.042	.018
	9	19	.043	.021	19	.038	.043	19	.041	.017
	10	19	.051	.025	19	.046	.026	19	.040	.018
	Pooled	19	.048		19	.044		19	.042	
Visit 3	Original data	18	.046	.018	18	.050	.027	18	.044	.020
	1	19	.047	.017	19	.048	.027	19	.042	.022
	2	19	.047	.017	19	.050	.026	19	.045	.020
	3	19	.046	.017	19	.049	.027	19	.042	.021
	4	19	.045	.017	19	.050	.026	19	.048	.026
	5	19	.050	.024	19	.049	.026	19	.046	.021
	6	19	.046	.017	19	.049	.026	19	.044	.019
	7	19	.044	.018	19	.050	.026	19	.041	.023
	8	19	.047	.018	19	.049	.026	19	.043	.020
	9	19	.047	.018	19	.050	.026	19	.044	.019
	10	19	.046	.017	19	.049	.026	19	.042	.021
	Pooled	19	.047		19	.049		19	.044	

		N	Mean	Std.	N	Mean	Std.	N	Mean	Std.
Visit4	Original data	19	.038	.014	19	.033	.021	19	.050	.027
	1	19	.038	.014	19	.033	.021	19	.050	.027
	2	19	.038	.014	19	.033	.021	19	.050	.027
	3	19	.038	.014	19	.033	.021	19	.050	.027
	4	19	.038	.014	19	.033	.021	19	.050	.027
	5	19	.038	.014	19	.033	.021	19	.050	.027
	6	19	.038	.014	19	.033	.021	19	.050	.027
	7	19	.038	.014	19	.033	.021	19	.050	.027
	8	19	.038	.014	19	.033	.021	19	.050	.027
	9	19	.038	.014	19	.033	.021	19	.050	.027
	10	19	.038	.014	19	.033	.021	19	.050	.027
	Pooled	19	.038		19	.033		19	.050	

Table Appendix 6 Descriptive statistics for т fixation durations in naturalistic videos, abstract videos and static images from the original and imputed data sets.

Imputation Number		Naturalistic videos			Abstract videos			Static images		
		N	Mean	Std.	N	Mean	Std.	N	Mean	Std.
Visit 1	Original data	18	.268	.051	18	.586	.230	18	.098	.026
	1	19	.266	.050	19	.588	.223	19	.096	.027
	2	19	.276	.060	19	.632	.301	19	.103	.033
	3	19	.271	.051	19	.589	.224	19	.092	.036
	4	19	.264	.053	19	.554	.261	19	.098	.025
	5	19	.267	.050	19	.585	.223	19	.097	.026
	6	19	.256	.073	19	.618	.265	19	.095	.028
	7	19	.262	.056	19	.560	.249	19	.100	.025

		N			N			N		
	8	19	.276	.061	19	.619	.267	19	.095	.029
	9	19	.279	.069	19	.620	.270	19	.104	.035
	10	19	.275	.058	19	.607	.242	19	.105	.039
	Pooled	19	.269		19	.597		19	.099	
Visit 2	Original data	16	.250	.042	16	.541	.192	16	.103	.029
	1	19	.252	.052	19	.548	.178	19	.109	.035
	2	19	.254	.047	19	.440	.346	19	.118	.056
	3	19	.253	.043	19	.521	.204	19	.093	.044
	4	19	.247	.054	19	.553	.219	19	.107	.031
	5	19	.251	.044	19	.615	.299	19	.103	.028
	6	19	.254	.057	19	.656	.379	19	.105	.027
	7	19	.260	.073	19	.632	.390	19	.115	.041
	8	19	.272	.071	19	.539	.261	19	.107	.029
	9	19	.264	.053	19	.451	.283	19	.125	.061
	10	19	.279	.112	19	.659	.578	19	.115	.047
	Pooled	19	.259		19	.561		19	.110	
Visit 3	Original data	18	.232	.049	18	.443	.163	18	.102	.034
	1	19	.242	.066	19	.449	.161	19	.106	.036
	2	19	.234	.048	19	.462	.179	19	.105	.035
	3	19	.238	.054	19	.459	.173	19	.103	.033
	4	19	.224	.060	19	.416	.196	19	.103	.033
	5	19	.231	.048	19	.447	.160	19	.105	.034
	6	19	.236	.051	19	.445	.159	19	.101	.033
	7	19	.232	.048	19	.443	.159	19	.105	.035
	8	19	.240	.058	19	.456	.169	19	.104	.034
	9	19	.239	.056	19	.459	.174	19	.106	.037
	10	19	.232	.048	19	.444	.159	19	.102	.033

	Pooled	19	.235		19	.448		19	.104	
Visit4	Original data	19	.242	.055	19	.502	.169	19	.113	.024
	1	19	.242	.055	19	.502	.169	19	.113	.024
	2	19	.242	.055	19	.502	.169	19	.113	.024
	3	19	.242	.055	19	.502	.169	19	.113	.024
	4	19	.242	.055	19	.502	.169	19	.113	.024
	5	19	.242	.055	19	.502	.169	19	.113	.024
	6	19	.242	.055	19	.502	.169	19	.113	.024
	7	19	.242	.055	19	.502	.169	19	.113	.024
	8	19	.242	.055	19	.502	.169	19	.113	.024
	9	19	.242	.055	19	.502	.169	19	.113	.024
	10	19	.242	.055	19	.502	.169	19	.113	.024
	Pooled	19	.242		19	.502		19	.113	