How children make sense of the world

de Bordes, Pieter

DOI:
10.33612/diss.157936930

How children make sense of the world

A perceptual learning account

P.F. de Bordes

Cover Herbert Bayer, 1964. *The art of progress is to preserve order amid change and to preserve change amid order. Alfred North Whitehead (1861-1947).*
© Smithsonian American Art Museum, Washington.
Layout Renate Siebes | Proefschrift.nu
Printed by ProefschriftMaken, De Bilt

university of groningen

How children make sense of the world

A perceptual learning account

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Monday 1 March 2021 at 12.45 hours

by

Pieter Ferdinand de Bordes

born on 28 March 1985
in Amersfoort

Supervisor
Prof. P. van Geert

Co-supervisors
Dr. R.F.A. Cox
Dr. F. Hasselman

Assessment committee
Prof. G. Grëdeback
Prof. E. Kunnen
Prof. M. Raijmakers

Voor Rafaël en Coco

Contents

Chapter 1

Introduction

Primer: making sense of infant behaviour

The diaper changing ritual

Whenever young Rafaël (14 months old) needs a diaper change, his father places him on the changing mat where their diaper changing ritual can start: after his father untightened his dirty diaper, Rafaël usually readily raises his legs up in the air, allowing his father to easily clean his bum and remove the diaper. During this time, Rafaël sometimes likes to play a little game in which he grasps the power cord next to him, waits until his father shifts his gaze from Rafaël's eyes to the cord, and pulls it whenever his father is looking at it, in a somewhat angry and ironic way, upon which he bursts into laughter because being naughty is the funniest thing ever for an infant. After his bum is cleaned and the clean diaper is well attached, Rafaël's father can proceed with closing the bodysuit using the three buttons at the bottom. His father counts each button he closes out loud until he arrives at the third one. After the third button is closed, Rafaël usually shouts out something that can be understood as 'three' just before his dad can finish, for which he is praised, and which marks the end of the diaper changing ritual.

In this short episode of dyadic interaction, young Rafaël is displaying a lot of interesting advanced social behaviours and abstract skills such as imitation, gaze following, facial expression recognition, and mathematical ability. These behaviours are typically studied by developmental psychologists in order to understand how and when they are instigated, developed and applied (see e.g., Siegler & Alibali, 2005). Considering the complexity of human behaviour, it is not surprising that several psychological perspectives exist that fundamentally differ in their descriptions of these behaviours and how they develop (see e.g., Haith, 1998; Spelke, 1998).

In the following subsections of this primer, Rafaël's developing skills with respect to imitation, gaze following, facial expression recognition and mathematical ability will be briefly described and explained from two different theoretical perspectives. Firstly, the cognitivist perspective, which has dominated the field of developmental science for the past 70 years, and secondly, the ecological perspective, which is more recent, less well-known, but gaining in popularity. The aim of these subsections is to provide the reader with an intuitive sense of how these perspectives operate, and how they differ in describing and explaining developing social and non-social behaviours in a concrete setting as described above. In the subsequent section, each of the perspectives are discussed more generally and abstractly, in terms of their fundamental assumptions about human behaviour and development. These perspectives are then compared in order to elucidate on what grounds these perspectives differ and to argue why the ecological perspective can be a potential alternative to the cognitivist perspective in describing and explaining the development. In the final section an outline of the rest of the thesis is presented.

Imitating intentions

Rafaël is able to raise his legs on the right moment, namely after his diaper was untightened so his father can clean his bum. Instead of strictly imitating his father's arbitrary acts, which would have the rather absurd implication that he would lift up the legs of another infant being changed, he is able to re-establish the intended goal of his father's action which is having his legs raised so his father can proceed cleaning his bum easily. This can be regarded as imitation (Hopper, 2010; Meltzoff, 1995). According to some cognitive psychologists, Rafaël's behaviour displays an understanding of others as psychological beings just like himself (e.g., Meltzoff, 2007). This understanding includes an interpretation of behaviour of others as representing mental states such as beliefs, emotions, goals and intentions. Therefore, it could be that at some point, Rafaël understood the intended goal of his father, and established or imitated it accordingly (Baron-Cohen, 1995; Meltzoff, 1995). This seems an efficient manner by which the behaviour of others can be interpreted in a structured and meaningful way in order to respond to it adaptively. According to Meltzoff (1995, p. 838), "failure to attribute mental states to people confronts one with a bewildering series of movements, a jumble of behaviour that is difficult to predict and even harder to explain".

From an ecological view, Rafaël might have learned to flex his legs independently, causing them to be in upright position, whenever he feels his diaper is untightened, and he is laying on the changing mat. Rafaël's repeated experience with the same behavioural sequence of his father during the diaper change, allows Rafaël to perceive his body to be in the same sequence of specific positions over and over again. In other words, Rafaël has experienced and learned the inevitability of his legs being raised at a certain point within the perceived action sequence that takes place within the dyadic interaction of the diaper change ritual (see Evans & Porter, 2009). More specifically, repeated experience of coregulating haptic and optic information in conjunction with bodily actions can mould a perception-action coupling in which specific perceptual information functionally integrates with a specific motor response (see Fogel, 1992; Gibson & Pick, 2000; Hunnius & Bekkering, 2014; Ruffman, 2014; Smith, Jayaraman, Clerkin, & Yu, 2018). In this manner, infants efficiently learn to perform certain actions upon perceiving specific (social) information, without contemplating what the information stands for in terms of for instance the underlying intentions of others (see Chapter 3 of this dissertation).

Following eye gaze of others

Another interesting feature is that Rafaël seems to notice his father's gaze shift from his eyes to the power cord, and that he is able to follow his father's gaze towards the cord. Making eye contact can be considered as a sign that a person wants to communicate something (Csibra, 2010). According to some cognitive scientists, infants as young as six months of age are able to detect this communicative intention of others through eye contact, leading infants to follow the subsequent shift of gaze of the other person to look at what the other person wants the infant to see (Senju & Csibra, 2008). This might explain Rafaël's ability to follow his father's eye gaze after they established eye contact.

However, from an ecological view, Rafaël's shift of attention from his father's eyes towards the power cord can be accounted for differently. For instance, humans – from birth on – prefer to focus on upright and frontally presented faces which are visually symmetrical in comparison to averted faces that are visually asymmetrical (Macchi Cassia, Turati, & Simion, 2004). In addition, they tend to visually follow lateral movements of faces and extend this movement beyond them (Farroni, Mansfield, Lai, & Johnson, 2003). Taken these two factors into account, gaze following could be instigated by attending to an upright and frontally presented face which subsequently moves towards a certain direction, making the face less interesting and stimulating the infant to look at the same direction the face has moved to. This reasoning does not imply that Rafaël needs to understand that his father wants to communicate something or even that his father perceives the power cord (Gredebäck, Astor, & Fawcett, 2018; Chapter 2 and Chapter 3 of this dissertation).

Recognition of facial expressions

The somewhat angry and ironic facial expression of Rafaël's father seems to be of significance to Rafaël as it makes him laugh (being the naughty boy he is). This suggests that he is able to recognize his father's facial expression. Cognitive scientists have described the mental schema in which humans can contrast perceived and experienced emotions by mentally differentiating them according to two orthogonal dimensions. First, a valence continuum that runs from pleasant such as 'happy' or 'pleased' to unpleasant such as 'anger' or 'disgust', and, second, an arousal continuum that can be high such as 'scared' or 'surprised' or low such as 'calm' or 'bored' (Russell & Bullock, l986; Widen & Russell, 2003). Accordingly, perceived emotional expressions are interpreted by means of placing them along these two dimensions by which meaning is attached to them (Russell, 1980). Young infants initially rely on the valence dimension, and later develop the use of the arousal dimension in order to better mentally differentiate observed emotional facial expressions (Widen & Russel, 2008).

From the ecological viewpoint, the ability of Rafaël to respond to his father's expression of anger the way he does, could have emerged from the experience with context-dependent functional relations between the perceived expressions and the possible actions that are adaptive in response to them (see Camras, 2000). From this point of view, Rafaël's reaction is not based on interpreting his father's expression in terms of 'pleasant' or 'unpleasant'. Instead, it is based on the functional integration of perceiving the expression of anger, and the action of laughing in response to it that took place over the repeated experiences with this particular expression in this particular context. In addition, the experience Rafaël has with the effects of his behaviour in many different contexts in which his father shows an emotional appraisal enables him to flexibly learn the connection between the observed expression to naughty behaviours in one context (e.g., pulling a power cord) that can lead to laughter, and forbidden behaviours in another context (e.g., throwing food) that can lead to punishment (see Chapter 4 of this dissertation).

Mathematical understanding

Finally, Rafaël has showed the ability to finish his father's counting sequence that matches the number of buttons closed on his bodysuit. From a cognitivist view, children around his age are believed to have a mathematical understanding in which they are able to make a rudimentary mental connection between numbers as words and quantities, leading them to mentally present approximate cardinalities from which they can build further understanding of mathematics (see Feigenson, Dehaene, & Spelke, 2004; Wang & Feigenson, 2019). Although he might not fully understand the exact meaning of each of the counting words, he is able to understand that number words are about numerosity, making it likely that he can utter a number word after hearing other numbers words, and specifically a number word that corresponds to the number of buttons closed (i.e., 'three').

From an ecological view, his ability to correctly utter the word 'three' at the right moment within the sequence of counting might be considered as resulting from a perceptual learning process in which information from different sense modalities gets organized over time (Smith & Gasser, 2005). For instance, across the hundreds of times his diaper has been changed up to this point, he might have seen his father's mouth moving while he heard his father utter the number words in a specific sequence that occurred at the same time and in the same order as the physical feeling of the buttons of his bodysuit being closed. Repeated exposure to the covariation of multimodal perceptual information as exemplified above can lead to perceptual attunement to the structure in which this information covaries (Fogel, 1992; Smith et al., 2018). In this structure, basic information can be nested such as specific context-dependent utterances, timing and sequence but also more abstract information regarding for example cardinality and numerosity (e.g., Bruineberg, Chemero, & Rietveld, 2019). Attunement to this nested information gets more efficient over experience, allowing children to perceptually access the abstract principles that govern logical systems such as mathematics more readily (see Chapter 5 of this dissertation). Therefore, Rafaël's behaviour does not necessarily reflect his understanding of abstract principles such as cardinality but instead, can be seen as a steppingstone towards it.

Contrasting views on human development

A cognitivist view on human behaviour

> "*To construct is the essence of vision. Dispense with construction and you dispense with vision. Everything you experience by sight is your construction.*"
> (Hoffman, 2000, p. 10)

Cognitive scientists investigate internal processes (i.e., mental processes) of the human mind such as attention, memory and reasoning, and the manner in which these internal processes relate to the input from and output to the outside world. From a classical cognitivist view, the workings of the human mind can be compared to the workings of a computer (see Fodor, 1985) as they both seem to operate according to a three-step process. (1) They both receive discrete *input* of stimuli which is subsequently (2) *internally processed* according to a set of preprogramed rules and innate capacities in order to (3) generate a behavioural *output* in the form of a response. In addition, non-physical mental or software processes are distinguished dualistically from the physical body or hardware respectively in which they take place (see Neisser, 1963). Consequently, stimuli input is translated into discrete (mental) code or symbols that represent or stand for the input in the form of a mental representation that can be stored in memory and onto which computations (i.e., mental processes) can be made in order to formulate a response to the input as output (i.e., behaviour). Hence, the input is received passively and processed actively. For example, 2-D images enter the retinas of the two eyes passively after which they are actively combined into a mental representation in order to produce 3-D images that can be used to guide navigation in the 3-D world (Marr, 1976; Stevens, 2012; Welchman, 2016). In this way, the input is enriched, structured, and attains meaning as the stimulus is thought to be meaningless (see Palmer, 1999; Saffran & Kirkham, 2018; also see the 'Poverty of the stimulus' argument by Chomsky, 1980). Hence, what is processed or computed in the mind is not a mere copy of the external world, but rather a meaningful interpretation of it, that is dependent on, and constructed by memory (i.e., stored symbols) and the contingencies of its own set of axioms and syntax (see Fodor, 1985; Neisser, 1967; Palmer, 1999). The axioms can be conceived as theories, core knowledge or assumptions humans have which are used to attach meaning to input and are believed to be innate (Frith, 2013; Spelke & Kinzler, 2007). Examples of such axioms are the ability to see differences in quantities, known as 'number sense' (Piazza, Pinel, Le Bihan, & Dehaene, 2007), and the ability to consider others as psychological beings and having mental states like beliefs, intentions and goals, known as 'Theory of Mind' (Baron-Cohen, 1995; Ruffman, 2014; Spelke, 1998). The syntax refers to the set of rules, principles, and processes that govern the structure of thought. Examples of the syntax of the human mind are decision heuristics, biases (Tversky & Kahneman, 1974), short-term memory capacity, and selective attention (Baddeley & Hitch, 1974; Downing, 2000). Together, memory, axioms, and syntax of the mind have the capacity to enrich and structure perception in order to form mental representations, which are then assumed to guide behaviour. For example, the perceived behaviour of other people is interpreted as goal-directed by which it becomes predictable (Biro & Leslie, 2007; Tomasello, 1999), and two-dimensional patterns of light are transformed into stable three-dimensional perceptions of objects, which enables successful grasping (Palmer, 1999). Consequently, from a cognitivist view, our behaviour is driven by thought processes in which we actively, mentally construct of what we passively perceive.

A cognitivist view on learning

From a cognitivist view, there seem to be many different theories on how learning and development takes place, each having different assumptions on the content and structures of knowledge, axioms, syntax, mental representations, and the interactions between them (see e.g., Demetriou, Mouyi, & Spanoudis, 2010; Malmberg, Raaijmakers, & Shiffrin, 2019). In this subsection, an attempt will be made to briefly describe those theories on learning and development that are widely supported within the cognitivist tradition, namely Piaget's theory on cognitive development (Piaget & Cook, 1952) and theories on working memory and its maturation (Baddeley & Hitch, 1974; Cowan, 2016).

According to the cognitivist perspective on Piaget's theory on development, the mind is updated by either adding knowledge and mental representations, by changing existing knowledge, or by changing the manner it is mentally represented[1]. For example, perceptual input can be refined upon computation using the current syntax, axiom configurations, and

[1] Originally, Piaget used the term 'scheme' instead of 'mental representation' to denote (characteristics of) an activity repertoire that can be changed by processes of (assimilation and) accommodation. Importantly, Piaget never considered a scheme as purely mentalistic in nature and instead, rather conceived it as an action repertoire in a broad sense that can include motor actions and perceptual input (see Boom, 2009). Here, however, following the cognitivist tradition, mental representation is used as a synonym to Piaget's 'scheme' (see e.g., Kibler, 2011; Woolfolk, Hoy, Hughes, & Walkup, 2007).

mental representations by which the input is translated into symbols that can be stored in a memory system in a pre-existing mental representation. This form of learning is known as '*assimilation*' (Piaget & Cook, 1952). For example, upon perception, children might learn to recognize a racing bicycle as being a bicycle, since they have previously learned that all bicycles have two wheels. Alternatively, mental representations themselves can also be adaptively updated and refined in response to novel input or the perceived difference between the intended output and the actual output, known as '*accommodation*' (Piaget & Cook, 1952; also see Netti & Nusantara, 2016). For example, children might notice that motor cycles have engines whereas bicycles do not, leading them to form a new mental representation that sets motor cycles apart from bicycles based on this feature. In this manner, knowledge is accrued and refined internally which can be used to enrich and disambiguate our perceptual input in order to guide our behaviour and attention in an increasingly meaningful and productive manner.

Apart from the updating processes described above, cognitive development can also take place in terms of maturation of the syntax components by which the speed and capacity to process and store increases with age. Cognitive scientists have identified several syntax components that operate in concert, in order to create, change, and store mental representations of which the most popular ones are the working memory and the long-term memory that each have their own subcomponents (Atkinson & Shiffrin, 1968; Baddeley, 2003; Baddeley & Hitch, 1974; Malmberg et al., 2019). Working memory seems to be a central point where perceptual input is selectively encoded into symbols that can be temporally stored in order to be structurally related to other symbols within mental representations that are either created within the working memory, retrieved from the long-term memory, or both. Importantly, working memory capacity increases with age, leading to an increased speed and power with which knowledge can be comprehended. This in turn leads children to develop their ability to execute complex tasks (simultaneously), and to memorize complex knowledge (Cowan, 2016). For example, maturation of the working memory during childhood has been linked to increased mathematical skills (Raghubar, Barnes, & Hecht, 2010). Taken together, cognitive development seems to be comprised of accumulation and (re)organisation of knowledge within the mind on the one hand, and an increased capacity and speed of the mind to do so on the other hand.

An ecological view on human behaviour

> "*You cannot step twice into the same stream*" (Heraclitus, in Beris & Giacomin, 2014).

Ecological psychologists investigate human action and perception as a unitary, emerging, and synergic product of their capacities, and their environment as "animal and environment make an inseparable pair" (Gibson, 1979, p. 8). From an ecological point of view, the Cartesian dualism that treats the immaterial mind and the material body as ontologically distinct (i.e., '*the ghost in the machine*') is rejected in favour of neutral monism. Neutral monism refers to the idea that thoughts, behaviour, body, and environment are ontologically inseparable, and have a common ground that can be simply regarded to as 'experience' (see Dewey & Bentley, 1946; James, 1895; Lobo, Heras-Escribano, & Travieso, 2018). Accordingly, experience cannot exist in an environment without an organism just like experience of an organism cannot exist without an environment. Dewey and Bentley (1946) described the phenomenology of events and experience as 'transactional' as opposed to 'interactional', stating that:

> "*If inter-action is inquiry of a type in which events enter under the presumption that they have been adequately described prior to the formulation of inquiry into their connections, then trans-action is inquiry of a type in which existing descriptions of events are accepted only as tentative and preliminary, so that new descriptions of the aspects and phases of events, whether in widened or narrowed form, may freely be made at any and all stages of the inquiry.*" (Dewey & Bentley, 1946, p. 535)

The emergence of behaviour

Their treatment of the term 'transaction' bears a lot of similarities with the term 'emergent properties', which is more commonly used nowadays to describe experience and behaviour of complex systems such as organisms (see Bar-Yam, 2004; van Dijk, 2020). Behaviour that emerges within an environment does not depend on the individual parts or entities that make up organism-environment system. Instead, it depends on how these parts or entities within the system relate to each other and reorganize to form something novel (e.g., Den Hartigh, Cox, & van Geert, 2017; van Geert, 2000, 2019; Steenbeek & van Geert, 2020). Behaviour as emerging from being in an environment is not reducible or even causally linked to disembodied or symbolic operations (as cognitive psychologists would claim), or a world with physical entities that relate to our physical bodies (as behaviourists would claim). An example of emergent behaviour is a flash of insight children display while solving a mathematical problem. This behaviour cannot be described in terms of the constituents of the event (i.e., reductionism), as it is rather a product of how these constituents dynamically interact (i.e., emergence). For example, neither the capacity of children for understanding mathematics, nor the math problem they are confronted with can fully account for the idea that they can attain a new strategy to solve a math problem at a given time. Instead, the experience of trying to solve the math problem might lead them to attend to (higher-order or abstract) relations between parts of the math problem in such a way that they can distil an abstract structure therein that leads to the solution (see Stephen, Boncoddo, Magnuson, & Dixon, 2009; Chapter 5 of this dissertation).

Emergent properties do not arise as a function of the laws of physics but are instead constrained by it (see Bar-Yam, 2004; Jacobs & Michaels, 2007; Stöckler, 1991). This makes them predictable to some extent, and thus in reach for empiricism when its constituents are described on the proper level. For example, the laws of physics constrain the specific movements young infants can make towards their mother such as crawling, considering for example the strength of their muscles and the surface they locomote on (see Adolph, 2019; Adolph & Berger, 2007). The fact that infants have the tendency to locomote towards their mother, however, cannot be described and predicted at the level of physics, and requires the laws of another descriptive level such as attachment theory (Bowlby, 1958) on the psychological level. Moreover, laws on the psychological level can ultimately pose constrains on the laws of physics making it for example unlikely that infants will physically propel away from her mother. This paves the way for stability of behaviour (i.e., 'attractor states') on one hand (Guevara, Cox, van Dijk, & van Geert, 2017), but also allows for adaptive flexibility of behaviour on the other hand (Adolph, Joh, Franchak, Ishak, & Gill, 2009). This stability and flexibility both make up for the system's state space (i.e., behavioural repertoire) that comprises of all possible system configurations (Chow, Davids, Hristovski, Araujo, & Passos, 2011). Simple systems like a collection of water molecules in a container can only be configured in a few kinds of collective states (e.g., solid, gas, or liquid) whereas the more complex biological adaptive systems that produce human behaviour can configure in a huge number of collective states (e.g., running, crying, writing, etc.). That is because behaviour is emerging from organism-environment couplings comprising an open and multicausal system. This means that humans can use multiple and different sorts of information to manifest the same behaviour but can also display different behaviours based on the same information (see Thelen, Schöner, Scheier, & Smith, 2001; Yu & Smith, 2017; Chapter 2 and Chapter 5 of this dissertation). This is because all humans are different; no single display of one's behaviour is exactly the same as we are continuously changing (i.e., our experience is incremental), just like the environment is continuously changing. This is nicely captured in the phrase "*Panta Rhei*" by Heraclitus, which means "everything flows" (Beris & Giacomin, 2014). It is in the midst of things and the continuous flow of everything that behaviour emerges and continues to exist (Fischer & Bidell, 2006). Altogether, this line of thought has led researchers to conclude that the natural variability of human behaviour and development should be the central focus of the scientific endeavour of psychology (van Geert, 2000, 2019;

van Geert & van Dijk, 2002; Siegler, 2006; see Chapter 5 of this dissertation), and "research must be designed to deal with variability, or it is doomed to fail to provide an adequate analysis of development" (Fischer & Bidell, 2006, p. 347).

Affordances and invariance detection

The continuity of behaviour within the environment can be segmented into affordances that are "both physical and psychical, yet neither" (Gibson, 1979, p. 121), and are meaningful for the individual as the action possibilities an environment has to offer, considering the individual's capacity to perceive and act (Gibson, 1979; Gibson & Pick, 2000; Lobo et al., 2018). Affordances are not only limited to objects but can also consist of the layout of surfaces (e.g., having corners or holes) and events, including social ones like a funeral or an angry face that all pertain to action (Gibson & Pick, 2000). "We perceive affordances of the ground to be walked on, of the cup to be drunk from, of the noises, fumes, and onrush of a truck in our path to be avoided" (Gibson, 1988, p. 4).

The foundation of perceiving affordances lies in ecological information as a set of structures and regularities that are invariant (e.g., Bruineberg, Chemero, & Rietveld, 2019; Lobo et al., 2018). For example, physical information such as the relative height of staircase steps with regard to the length of legs informs an individual of its 'climbability' as affordance, based on the invariant ratio of the height of staircase steps relative to the length of the legs of an individual (Konczak, Meeuwsen, & Cress, 1992; Warren, 1984). Invariants can be either structural or transformational (Hellendoorn, Wijnroks, & Leseman, 2015). Structural invariants remain the same over changing conditions such as the stability of the size and the shape of a ball despite its perceptual differences when seen from varying distances and viewpoints. For instance, these structural invariants can inform an individual on whether the ball is in reach, and can be grasped with one hand. Transformational invariants specify the structure of change as it unfolds over time. Examples are the changing spatial configuration of a human body that denotes a predictable walking pattern or the changing optical size of a thrown ball that allows humans to predict the moment when it can be caught (see Fajen, 2008; Fajen, Riley, & Turvey, 2008). Invariants can also be nested in other invariants or combined with other invariants, including those in our own behaviours and socio-cultural practices, and give rise to abstract understanding of our surroundings (Bruineberg et al., 2019; see Gibson, 1979). For instance, when solving a physics problem, one can attain a more efficient solving strategy by detecting certain key regularities (i.e., specifying variables) in the execution of the current strategy that lead to the correct result more efficiently (see Chapter 5 of this dissertation). Most humans seem to be endowed with the capacity to detect order amid change, and to detect change amid order in the tremendous and chaotic sea of perceptual information we are constantly surrounded with (Adolph & Kretch, 2015; Gibson & Pick, 2000; Hellendoorn, Wijnroks, & Leseman, 2015).

An ecological view on learning

Already from birth on, new-borns readily explore their surroundings through which they can pick up invariants in this world (Adolph & Kretch, 2015; Gibson, 1988). Although in the beginning, "the baby, assailed by eyes, ears, nose, skin, and entrails at once, feels it all as one great blooming, buzzing confusion" (James, 1890 p. 488), they are already able to perceptually and selectively attune to the invariants within their surroundings and the invariants in the relations they maintain with their surroundings (Gibson & Pick, 2000; also see Von Hofsten, 2007). For instance, already within the first months of life, infants detect the invariant co-occurrence of hearing and seeing their caretaker. Over time and with experience, perception (e.g., hearing the sounds of the caretaker) and action (moving the head and eyes until (s)he is in sight) reorganizes to form a tight coupling between the two, ultimately leading infants to readily visually orient towards the sound of their caretaker (see Smith & Gasser, 2005). The attunement of perception and action that leads to the detection of invariants is called 'differentiation', and is the central mechanism of perceptual learning (e.g., Adolph & Kretch, 2015; Gibson & Gibson, 1955; Goldstone, 1998). Differentiation occurs in the service of affordances that have adaptive value and to that end, differentiation is selective. For example, as E. J. Gibson highlighted (1988, p. 23):

> "*Color receptivity is mature well before six months, but color does not appear to be an important factor in defining affordances of objects at this stage and was not differentiated as specifying anything important. Indeed, when one considers the action repertory of a six-month-old, what could color signify? Finding and securing something warm to the touch is a different matter. This does not mean that visual information is not important--optical specification of substance, shape, and where something is located certainly is important. Perception is selective at six months, but not in purely sensory respects; exploratory activity is geared to affordances of objects.*"

In this manner, perceptual learning is geared by exploration of the environment, leading to responses to perceptions not previously responded to (Gibson, 1988; Gibson & Gibson, 1955). That is because explorative behaviours, unlike for instance repetitive behaviours, impose variability within the person-environment transaction which can optimally reveal the invariants that specify affordances. With experience, attention becomes more attuned to the relevant information that allows adaptive behaviour to emerge from it efficiently.

This makes it likely that infants are inclined to pay attention to the shapes of an object rather than its colour if the intention is to grasp the object. With the experience of grasping objects and failing to grasp objects, infants might selectively search for parts of the object that can be grasped (i.e., specifying variable), while ignoring other aspects of the object. Over time and with experience, attention becomes attuned towards the information that optimally and efficiently reveals (i.e., '*differentiates*') affordances. In effect, this leads to an increasingly efficient integration of perception and action, and allows for the development of skilled behaviours and abstract understanding (Bruineberg et al., 2019; Huet et al., 2011; Winkler, Mueller, Friederici, & Männel, 2018). For instance, adults need very little perceptual information to identify another person as walking, and to identify the gender and mood of another person by the style of walking (try it yourself, see www.biomotionlab.ca/html5-bml-walker/). The increased selectivity by which humans attune to information within their environment that effectively specifies affordances, and leads to the integration of perception and action, is known as 'the education of attention' (Gibson, 1979; Jacobs & Michaels, 2007).

Where the two accounts meet: enrichment versus differentiation

How do humans depart from the 'great blooming, buzzing confusion' when we enter this world, and develop into adults that interact with their environment in an increasingly differentiated manner? The cognitivist perspective and the ecological perspective discussed in this chapter each seems to have a distinct answer to this question, and they seem to be in stark contrast with each other when it comes to human learning and epistemology (i.e., the origin of knowledge), as pointed out by many scholars over the last century (e.g., Bauchau, 2006; Costall, 2011; Haith, 1998; Holt, 1914). In their influential paper of 1955, the Gibsons phrased this question eloquently in the title of their paper and in terms of the theoretical ramifications in respect to the views addressed in this chapter: "Perceptual Learning: Differentiation or Enrichment?" (Gibson & Gibson, 1955). From a cognitivist view, they viewed learning as a process of enrichment, whereas from an ecological view they viewed learning as a process of differentiation, as will be explained below.

The cognitivist view supposes that the physical environment is meaningless (impoverished), and requires interpretation in order to become structured and meaningful. Hence, humans passively receive meaningless information through the sensory organs. Perception is then processed mentally using different sorts of inborn core knowledge and stored memory by which it can be recognized as meaningful in order to respond to it adaptively. In this manner, newly received information is *enriched* with core knowledge and what was previously learned, and the result of it integrates with existing mental constructs for development to occur (Gibson & Gibson, 1955). Consequently, these mental constructs *correspond progressively* *less* with the meaningless received information, and correspond more with the (structure of) knowledge that served adaptive behaviour until that point (Gibson & Gibson, 1955).

Alternatively, the ecological view asserts that the ecological environment is rich in terms of information that reveals the possibilities for meaningful action and perception, and simply requires exploration through which these possibilities can be picked up (Gibson, 1988). Hence, perception is an active process in which sources of variation, such as structural changes in dimensions and features of (objects in) the environment, are progressively more perceptually differentiated, and increasingly coupled to specific actions whenever these actions serve adaption meaningfully. In this view, the mental realm is obscure because development entails that perception gets richer in terms of responses to it (i.e., differentiation), and not in terms of mental constructions we make of it. Therefore, according to the ecological view, *perception corresponds progressively more* with the physical objects and properties in the environment whenever it serves adaptive behaviour (Gibson & Gibson, 1955).

The notable point of the ecological view is that meaning is perceived directly and is not a product of mentally processing whatever is perceived passively. For example, humans do not need to process a staircase in terms of its defining features such as its material, dimensions and colour in order to assess its utility for going up or down. Instead, they perceive its utility for action and its features (e.g., whether it is made of paper or wood) only when it is deemed relevant for action; a staircase of paper would probably collapse under the weight of a human being (see Gibson, 1979; Jacobs & Michaels, 2007; Read & Szokolszky, 2018; Szokolszky, Read, Palatinus, & Palatinus, 2019). By means of the education of attention (i.e., differentiation), actions become more stringent upon specific perceived information, such that the steepness of a staircase might lead people to go down backwards when the staircase is steep, or forwards when it is not (see Gibson, 1979; van der Kamp, Oudejans, & Savelsbergh, 2003). Education of attention is not confined to the relations humans maintain with their physical world alone but could also extend to more abstract matters like social relations (see Chapter 2, Chapter 3, and Chapter 4), and understanding how physical mechanisms work (see Chapter 5). Therefore, inasmuch humans are endowed with the capacity to detect invariants in their environment based on their explorative behaviours therein, the ecological view has the potential to describe the full range of human behaviour and development, including skilful behaviours, and its development for which the cognitivist view normally requires innate knowledge, mental constructs, and mental processes to be at work. The ecological alternative to cognition overcomes dualism in the cognitivist view that separates thinking from behaviour and humans from their environment, and instead introduces a world view that leaves no room for reductionism in which one part of a duality such as behaviour is explained by its counter side, such as thinking (Costall, 1995; van Dijk & Withagen, 2014).

Considering the parsimony of assumptions (i.e., ability to detect invariants and to explore that leads to the discovery of affordances in the service of adaption action) of the ecological approach, together with its potential explanatory value, it seems worth it to explore its potential further within research on how humans are able to develop abstract skills that make them so unique and advanced in comparison to other animals. Therefore, this dissertation is an attempt to show the potential of ecological psychology in describing the development of skills such as social behaviour and understanding of physical mechanisms that are typically addressed from a cognitivist view. The main hypothesis of this dissertation is:

> *Children develop social and non-social skills by means of continuous perception-action attunement to the relevant information in the environment.*

The organization of this dissertation

The chapters in this dissertation gradually depart from the static descriptions of how infants learn social skills (see Chapter 2 and Chapter 3) towards more dynamic descriptions in terms of how children develop social skills (see Chapter 4) and non-social skills (see Chapter 5).

Chapter 2 reports a study on the ability of infants to follow the direction of sight of another person (i.e., gaze following). This study partly replicated an experiment by Senju and Csibra (2008) in which they showed that infants were more likely to follow the gaze direction of an adult toward an object whenever it was preceded with eye contact. They reasoned that eye contact can convey the intention to communicate something, and that infants perceive eye contact as such, giving them a reason to follow the gaze of the person. In the reported study in Chapter 2, an experimental condition was added to the ones used by Senju and Csibra (2008) in which the attention of infants was drawn towards the eye region of an adult without making eye contact. We reasoned that if infants were likely to follow the gaze of the adult in this condition, it would render the reasoning of Senju and Csibra (2008) obsolete and opens up the possibility that gaze following skills might be rather a product of perceptual attunement and attentional modulation instead.

Chapter 3 reports a study on a special case of imitation, namely the ability of infants to successfully re-enact a failed attempt by an adult to combine two objects in a specific manner. This study was a variation of a study reported by Meltzoff (1995) in which it was claimed that infants are likely to imitate the intentions of another person because infants are endowed with the ability to detect the intention behind the perceived behaviour, even when the perceived behaviour is a failed attempt of what was intended. In this chapter, we offer an attunement-based explanation. Specifically, we suggested that the success by infants to combine the two objects might be a result of stimulus enhancement. This refers to the idea that observing someone performing actions onto objects in a specific manner can lead the observer to perceptually attune to the key object manipulations and object features that reveal the proper object affordance. In addition, we reasoned that if infants would follow the gaze of another person that acts out the failed attempt, it could further assist them in allocating their attention towards finding the proper object affordance. Across several conditions, we varied the salience of the eye region of the person before she acted out a failed attempt to combine objects in a specific manner. We predicted that making the eyes of the person salient (either ostensively or non-ostensively) would draw the attention of infants towards them, leading infants to follow the gaze of that person during the failed attempt demonstration. In order to test these claims, we measured the success infants had with acting out the proper affordances with the objects. In addition, a spatiotemporal analysis of infants' looking behaviour was performed in order to reveal differences in successful and unsuccessful attempts. This was done in order to assess whether the gaze and the object movements of the person facilitated attunement and affordance learning by directing the attention of infants towards important object-directed actions on crucial moments during the failed attempt demonstrations.

Chapter 4 reports a cross-sectional study on the ability of five-to-nine-year-old children in identifying facial expressions according to the context in which they are normally perceived. In this chapter, we argued that the perception of different facial expressions emerges from context-dependent functional relations between the perceived expressions and the possible actions that are adaptive in response to them. As with increasing age, children gain experience in perceiving facial expressions in relation to contextual information, we predicted that facial expressions gradually become more differentiated in terms of the contexts in which they are typically perceived. To investigate this idea, we presented children prototypical contexts for different emotion categories, and subsequently asked them to identify whether different kinds of facial expressions belonged to the previously presented prototypical context, or not, using a two-alternative force-choice task. Correct and incorrect identifications as belonging to an emotion category were quantified using Signal Detection Theory into a single index, representing their ability to differentiate each of the universal facial expressions. In addition, we calculated the diversity of incorrect categorizations for each facial expression per age group as a supplementary index of their ability to differentiate facial expressions. We predicted that with increasing age, children would correctly identify more facial expressions as belonging to their prototypical contexts, and that they would identify less kinds of facial expressions as belonging to a context in which they are unlikely to occur. This would suggest that they become better at differentiating facial expressions in terms of the contexts in which they are normally perceived.

In Chapter 5, a study is reported in which we observed children's strategy during a task in which they had to find the rotation direction of the last gear in a series of connected gear chains, given the rotation direction of the first gear. From a perceptual learning account, we reasoned that children would differentiate more efficient strategies (i.e., affordances) based on detecting invariants that are nested within the application of less efficient strategies on the task. Therefore, we predicted that children generally would develop new strategies while working on this task in a specific order, namely from unskilled sensorimotor strategies to more abstract strategies. However, we assumed that this learning process would be non-linear, meaning that during the task execution, some children might make forward and backward transitions between strategy use, and even occasional transitions that skip certain strategies in the predicted order. This would reflect the natural variability within the person-environment transaction that optimizes the differentiation process, leading to the detection of the invariants that specify more efficient strategies to solve the task. These predictions were assessed using a Dynamic Overlapping Waves Model.

In Chapter 6, final remarks and caveats are presented regarding the conclusions of each chapter. In addition, suggestions are made for future research, education, clinical implications, and theory building. This chapter concludes with an epilogue in which final remarks are given on how the field of developmental psychology can move forwards by leaving the cognitivist view behind in favour of adopting an ecological view.

Chapter 2

Infants' gaze following through attention modulation: Intention is in the eye of the beholder

P. F. de Bordes
R. F. A. Cox
F. Hasselman
A. H. N. Cillessen

Based on:
De Bordes, P. F., Cox, R. F. A., Hasselman, F., & Cillessen, A. H. N. (2013). Toddlers' gaze following through attention modulation: Intention is in the eye of the beholder. *Journal of Experimental Child Psychology*, *116*(2), 443-452. https://doi.org/10.1016/j.jecp.2012.09.008

Abstract

We investigated 20-month-olds' ($N = 56$) gaze following by presenting infants with a female model that displayed either ostensive or no ostensive cues before shifting her gaze laterally toward an object. The results indicated that infants reliably followed the model's gaze redirection after mutual eye contact was established but did so equally reliably after the model's eyes had been made salient nonostensively. Moreover, both conditions elicited gaze following more prominently than when infants' attention was initially directed away from the eyes either by specifically accentuating the mouth or by covering the entire face before the model redirected her eyes laterally. These findings suggest that gaze following by infants is more likely to be driven by general attention mechanisms than by their appreciation of somebody else's communicative intent through perceiving eye contact.

Introduction

It is a fundamental issue in the study of social development how infants learn to perceive the referential nature of other people's nonverbal behaviors such as facial expressions, pointing, and (in particular) eye contact and gaze (re)direction. First of all, within minutes from birth, infants are more likely to attend to face-like stimuli than scrambled or random patterns (Goren, Sarty, & Wu, 1975). Newborns also prefer to look at face-like stimuli that show the eyes compared with stimuli that do not (Batki, Baron-Cohen, Connellan, & Ahluwalia, 2000). Moreover, newborns show a preference for faces with direct (mutual) gaze as opposed to averted gaze (Farroni, Csibra, Simion, & Johnson, 2002). To elaborate, within the first year of life, infants are able to use an adult's eyes as an informative source, following their gaze direction toward objects or events that indicate where the adult's visual attention is directed (Carpenter, Nagell, & Tomasello, 1998). This is known as joint attention, where an infant joins in the attention of another person toward an entity. Unlike dyadic interactions such as protoconversations, joint attention is typically triadic in the sense that it involves both the coordination of the infant's interactions with another person and an entity or event to which they share attention, resulting in a referential triangle of child, adult, and entity/event (Tomasello, 1999). The significance of gaze following as an important cornerstone of social development is generally recognized (e.g., Baron-Cohen, Campbell, Karmiloff-Smith, Grant, & Walker, 1995; Carpenter et al., 1998; Flom, Lee, & Muir, 2006; Slaughter & McConnell, 2003; Tomasello, 1999). By following gaze, infants are able to redirect their own attention toward whatever is relevant to another person. By doing this, infants can discover what might drive that person's current and future actions.

An interesting finding that has been replicated consistently is that infants are most likely to follow the gaze of others after mutual eye contact has been established (Farroni, Johnson, Brockbank, & Simion, 2000; Gredebäck, Örnkloo, & Von Hofsten, 2006; Hood, Willen, & Driver, 1998; Senju & Csibra, 2008). Because mutual gaze can be interpreted as a communicative or ostensive signal, some researchers take this finding to reflect infants' understanding of others' communicative intent and expectation of a more active communicative role from the information source (e.g., Csibra, 2010; Csibra & Gergely, 2009; Grossmann, Parise, & Friederici, 2010; Hoehl et al., 2009; Senju & Csibra, 2008; Shepherd, 2010). Within this perspective, it is suggested that in order to establish joint attention between an infant and an adult by means of gaze following, the infant is required to detect a communicative intent of the adult provided through eye contact (Bruinsma, Koegel, & Koegel, 2004; Farroni et al., 2002; Mundy & Newell, 2007; Senju & Csibra, 2008). Subsequent behavior of the adult will then be interpreted as communicative, as in joint attention. In other words, it is the appreciation of communicative intent by which infants modulate their interpretation for what they see next.

The current study questioned this interpretation of the role of mutual eye contact in gaze following and investigated whether an attention modulation mechanism is sufficient to explain infant gaze following. First, we hypothesized that once an infant's attention has been drawn toward a model's eyes, the infant is more likely to follow subsequent gaze redirections than when attention has not been drawn toward the eyes. This demonstrates that the basic result of gaze-following research is replicated and, by that, emphasizes the importance of attention toward the eye region (e.g., Farroni et al., 2000; Farroni, Mansfield, Lai, & Johnson, 2003; Senju & Hasegawa, 2006). Second, we hypothesized that this effect would be present even without establishing actual eye contact between the infant and the model. This means that eye contact does not need to be established as a prerequisite for the infant to follow subsequent gaze redirection of the model. Such a result will weaken any account of gaze following that makes reference to some sort of mental interpretation of eye contact on the part of the infant as a necessity for gaze following to occur. No additional mentalistic reference or interpretation by the infant is needed as a result of the eye contact in order to follow the gaze subsequently. In the remainder of this Introduction, we review studies that show under which conditions infants are likely to follow gaze. Specifically, the study by Senju and Csibra (2008) is reviewed, whereby the attention modulation and interpretation modulation accounts are contrasted. Finally, the current study is introduced and its hypotheses are formulated.

Conditions for gaze following in young infants

Infant gaze following has been studied extensively in 4-month-olds by Farroni et al. (2000, 2003). They focused on the basic sufficient conditions for infant gaze following to occur. They found two such conditions using spatial cueing paradigms in which a sequence of three pictures was shown: one face frontally presented to the infant, another face turned away laterally from the infant either left or right, and finally a stimulus at either the left or right visual field. The first condition concerns frontal face presentation. Infant gaze following was most efficient when the perceived face was presented in an upright orientation with frontal gaze before the averted gazing face was shown (Farroni et al., 2003; Senju & Johnson, 2009). The infant's preference for faces presented in an upright and frontal orientation has been formalized in a model named CONSPEC. According to this model, the infant's attention is drawn to stimuli that share the basic properties of a face (Morton & Johnson, 1991). Macchi Cassia, Turati, and Simion (2004) extended this model by showing that it deals with the configurational properties in terms of symmetry and distribution of elements (top-heavy hypothesis) rather than the face-like features of the stimuli per se. Specifically, infants show a preference for looking at stimuli that comply with a configuration of elements that are laterally symmetric and have more elements distributed on the upper part.

The second condition was that gaze following occurred only when infants could see a lateral motion of facial properties after frontal face presentation (Farroni et al., 2000). Using a spatial cueing paradigm, it was found that 4-month-olds followed the lateral movement of the head while the eyes remained frontally fixated (Farroni et al., 2000, Experiment 2). In addition, these infants did not follow eye gaze when the motion from frontal to averted eye gaze was not perceivable due to an eye blink of 1 s (Farroni et al., 2000, Experiment 3). Finally, 4-month-olds followed the lateral movement of the eyes even when the lateral movement was from averted to central orientation (Farroni et al., 2003, Experiment 3). That is, if the perceived eyes were first centrally fixated with respect to the infants, then averted to the left, and finally moved to the center again, infants were faster to respond to a stimulus onset in their right visual field than in their left visual field. This indicates that they tend to follow the last lateral motion they perceive even if that last motion resulted in frontally positioned eyes. In sum, 4-month-olds may be seen as biased to look at specific stimulus configurations such as a face and tend to follow the last lateral movement they perceive (Farroni et al., 2003). If these conditions are met, we may expect to observe gaze following according to these data.

An attention-based account of gaze following

An important remark, however, with respect to the studies conducted by Farroni and colleagues (2000, 2003) is that the face was never present at the same time as the target object. Hood and colleagues (1998) showed that gaze following in 4-month-olds decreased drastically when the face with averted eyes was shown together with the target object. These infants tended to remain fixated on the face if this was present during target presentation, demonstrating their difficulty with disengaging attention from the configurational properties of the face. Together with the finding that direct gaze is more effectively detected by both infants and adults when the face is presented frontally than when it is averted (e.g., Senju & Hasegawa, 2006; Senju, Hasegawa, & Tojo, 2005), this elucidates the prominent but restrictive nature of the top-heavy hypothesis with respect to 4-month-olds' "attention grabbing" and gaze following (Macchi Cassia et al., 2004). That is, on the one hand, direct gaze is most effectively detected when looking at frontal (i.e., symmetric) faces; on the other hand, frontal faces (even with averted gaze) continue to draw infants' attention, restraining them from reallocating their attention to a peripheral target (but see Hains & Muir, 1996).

The same contextual cautions apply to the second identified necessary condition of lateral movement following. It has been found that 4-month-olds follow the lateral direction of the movement of the eyes, but not the lateral movement of other equally small facial properties such as the tongue (Hood et al., 1998). Because various studies have shown that, when presented with an upright frontal face, infants and adults alike mainly tend to scan the eye region of the face (Haith, Bergman, & Moore, 1977; Jones, Carr, & Klin, 2008; Klin & Jones, 2008; Walker-Smith, Gale, & Findlay, 1977), this might just as well be an attentional mechanism. In the case of a tongue movement, for instance, attention is not likely to be directed at the relevant region of the face where the lateral movement occurs (i.e., the mouth region). So, although eyes are certainly highly attractive and informative—they do move a lot—infants probably do not follow the tongue movement simply because their attention was directed to the eyes already (Hood et al., 1998). This by no means implies that eyes are somehow more important for additional reasons.

As a final argument, consider the study by Vecera and Johnson (1995), who reported that when the eyes were presented in a context of a scrambled face, infants were not interested enough to maintain their attention toward it and, as a consequence, they did not perceive the subsequent motion of the eyes. This finding suggests that what are important are the configurational properties of the face, that is, the elements in relation to each other. This is what seems to drive infants to direct their attention to the eyes and not the features of the face, that is, not the elements by themselves (e.g., the eyes, the mouth).

An interpretation-based account of gaze following

The account of gaze following as described above is somewhat different from the hypothesis in a recent study by Senju and Csibra (2008), where the perception of eye contact in the context of an upright frontal face is suggested to serve as a signal by which expectancy of an upcoming interesting event is communicated to a child. In their study, they presented the head of a female model in upright frontal orientation to 6-month-olds. Then either the model performed an eyebrow flash and established eye contact with the infant before averting her face toward an object (Initial Eye Contact [EIC] condition) or an animation was presented onto her face before her face averted toward the object (No Eye Contact [NEC] condition). Note that in this study, gazes were cued by a lateral movement of the head instead of the eyes only, as was done in the studies mentioned above. It was found that infants reliably followed the gaze of the model only in the IEC condition and not in the NEC condition. From this, the authors concluded that the difference was due to the infants' understanding of communicative intent conveyed by eye contact of the female model, something that was absent in the NEC condition.

However, in line with the attentional account of gaze following, these results could also be due to the fact that in the NEC condition infants' tendency to scan faces and direct attention to the eye region was blocked because the head and eyes were not visible during the initial phase of the trials. As a result, the salience of the lateral motion of the face and eyes was reduced because, much like the failure to follow lateral tongue movement, infants' initial attention was not directed to the eye region. In addition, infants' allocation of attention toward a peripheral target in the IEC condition could be prompted by the fact that the perceived averted face does not comply with the preferred configurational properties of a face (i.e., symmetrical top-heavy configuration) anymore, thereby making this region less interesting and so increasing the probability of a shift of attention to the peripheral target to which the face is moving. This could also explain results in other studies where it has been shown that infants between 6 and 18 months of age are able to reliably follow someone's gaze based on head turns (Butterworth & Jarrett, 1991; Carpenter et al., 1998; Deák, Flom, & Pick, 2000; Flom & Pick, 2005; Mumme & Fernald, 2003; Scaife & Bruner, 1975). Therefore, we think a plausible alternative hypothesis is that infants follow gaze by modulation of attention rather than appreciating someone's communicative intent (interpretation modulation) in situations where they perceive an upright frontally presented face or face-like stimulus.

The current study

In this study, we investigated whether gaze following of infants is due to the fact that they perceive upright frontal faces with eye contact as a primer for communication (i.e., interpretation modulation) or that gaze following of infants is facilitated by an attentional bias toward looking at the eyes of a face combined with their tendency to follow lateral movements (i.e., attention modulation). Instead of head turns as indicators of where the gaze is directed, we used eye gaze alteration while the head remains in a frontal and upright position during target presentation because we are interested in the manipulations of attention direction in the biased context described by the top-heavy hypothesis (Macchi Cassia et al., 2004) on the probability of lateral eye movement detection. This, of course, has consequences for the age of the participants; by including older infants in our study (20 months of age instead of 4–6 months of age as used in previously discussed research), we ensured that our participants would be able to perceive the small lateral movement of the eyes because at this age infants are better able to detect high-frequency spatial changes (see, e.g., Hainline, 1998). Furthermore, infants of this age are better able to disengage their attention from a centrally presented and interesting stimulus such as a face (Hood et al., 1998; Atkinson, Hood, Wattam-Bell, & Braddick, 1992). This was evidenced in a study by Corkum and Moore (1995). In comparing different age groups on gaze following, they found that only from 18 months of age onward did infants begin to follow eye movements while the head remains frontal and in sight during target presentation (see also Doherty, 2006).

Using participants in this age group allowed us to investigate their gaze-following skills in a more ecologically valid setting, that is, with a frontal positioned face continuously present during the gaze-following phase, unlike the spatial cueing studies mentioned above (Hood et al., 1998; Farroni et al., 2000, 2003). In addition, if we were to find evidence that corroborated our attention modulation account in these older infants, this would strengthen our argument because it can be expected that they already are more advanced in perceiving communicative cues than younger infants (Gredebäck, Theuring, Hauf, & Kenward, 2008).

To avoid conflation of the perception of an upright frontal face and perceiving this as a communicative act, we performed two additional variations of Senju and Csibra's (2008) setup. In the Eye Salience (ES) condition, moving and flashing dots were presented over the eyes of an upright frontally presented face before the eyes became visible and averted laterally. In this version, eye contact and so a communicative act is not established, whereas the attention is still drawn toward the eyes region, a situation that would otherwise also occur naturally (Senju & Johnson, 2009). In a fourth condition, a moving and flashing dot was presented over the mouth of the model while her face was presented in an upright and frontal manner (Mouth Salience [MS] condition). We included this last condition to assess whether infants' gaze following would be reduced when their attention was drawn away from the eyes even though the eyes were still visible (i.e., eye contact could still be established). With these conditions in addition to the IEC and NEC conditions, we predicted that gaze following would occur with equal probability in the IEC and the ES conditions and was more likely to occur in these conditions than in the NEC and MS conditions.

Methods

Participants

Infants were recruited through written invitation to their home address. Addresses of 20-month-olds in Nijmegen, The Netherlands, were provided by local government authorities. Of the 510 invitations sent, 61 parents (12%) responded with a confirmation to participate. This led to study completion of 61 infants (31 boys and 30 girls) with an average age of 20 months 2 days (*SD* = 17 days). All participants were native Dutch. Infants were excluded from the analysis if they did not look at the screen during one or more phases in two or more trials. Brief glances away from the screen were allowed unless the infants looked away during eye movement of the model. Using this criterion, 5 infants were excluded from the analysis (2 in the No Eye Contact [NEC] condition, 2 in the Eye Salience [ES] condition, and 1 in the Mouth Salience [MS] condition). This resulted in having 14 infants in the Initial Eye Contact (IEC) condition, 17 in the NEC condition, 15 in the ES condition, and 10 in the MS condition.

Apparatus

During the experiment, infants sat on their parents' lap in a room with dimmed light, approximately 60 cm away from a screen (17-inch TFT monitor, 60 Hz) with a built-in remote eye-tracking system having a sample rate of 60 Hz (Tobii T120, Tobii Technology, Danderyd, Sweden). The 31.2 by 22.9-cm screen and the distance from it created a 30° by 21° visual angle with a screen resolution of 1280 × 1024. The eye tracker was linked to a laptop on which software (Clearview Version 2.7, Tobii Technology) was used to calibrate the eyes of the infants and to present stimuli and record eye movements and fixations of the infants directed on the screen. In addition, a small webcam (Sonix SN9C201, Taiwan) was placed on top of the Tobii monitor in order to record infants' behavior during presentation of the stimuli.

Procedure

On arrival, the infants and their parents were brought into the testing room. After a brief warm-up period, each parent was seated in front of the Tobii eye tracker with the infant on her lap. Then a five-point calibration procedure was performed in which expanding and contracting black and white circles, accompanied by bleeping sounds, were presented in the middle and corners of the screen on a white background. The calibration procedure was repeated until all five points were calibrated, after which the experiment was executed. During the experiment, the experimenter sat out of sight behind the infant. During the experiment, six trials were presented to the infant, with each trial being preceded by an "attention grabber". At the end, the parent(s) and child were offered either a 10-euro gift certificate or a infants' book, after which they were debriefed and thanked for their cooperation.

Stimulus material

The stimuli and procedure replicated those of Senju and Csibra (2008), but with some adjustments. In each trial, the torso and head of a female model was presented centrally on the Tobii screen, with a colorful object placed at equal distances on her left and right (see Figure 2.1). The objects were similar in terms of their spatial dimensions and color properties. At a viewing distance of 60 cm, the 6.1 by 7.7-cm female head subtended a 5.8° by 7.3° visual angle and the 7.3 by 6.0-cm objects (average size) subtended a 7.0° by 7.3° visual angle. Each trial consisted of a movie clip that can be divided into three phases. During Phase 1, a model was presented on the screen, directing her face downward for 2 s. During Phase 2, the model looked up for 2 s, having her face in an upright and frontal position. The manipulation of this phase created the four conditions to be applied between participants, meaning that the infants were either in the IEC condition, the ES condition, the MS condition, or the NEC condition. In the IEC condition, the model looked up into the camera. In the ES condition, the model looked up into the camera while moving and blinking dots were covering her eyes, thereby preventing the eyes from being visible during this phase. In the NEC condition, the model looked up into the camera while a colorful animation of a flower was placed over her head. In the MS condition, the model looked up into the camera while a moving and blinking dot was placed over her mouth. Finally, during Phase 3, the model turned her eyes toward an object to her right or left and fixated her eyes on it for 5 s. For each condition, the same video recording of the model was used, meaning that the differences among conditions were brought about merely by adding dots (in the ES and MS conditions) or adding a flower animation (in the NEC condition). Therefore, the duration of each phase was similar across conditions, and salient features such as eye gaze shift were visible in each condition and presented at the same onset across conditions. Six trials were presented to the infants in this way. Each trial was preceded by an attention grabber displaying colorful animations with bleeping sounds until the infants attended to the center of the screen. The side that the model gazed toward during Phase 3 varied across trials in ABBABA order (with “A” being either left or right in the first trial and “B” being the other side). The order of presentation was randomized across participants.

Figure 2.1. Schematic overview of the procedure and different conditions in the experiment. Phase 1 (top row) is the looking down phase (duration: 2 s). Phase 2 (middle row) is the manipulation phase during which the four conditions are presented to the infants (duration: 2 s). Phase 3 (bottom row) is the looking phase where the model shifts her gaze toward one of the two objects (duration: 5 s). During the second phase, the four types of ostensive and nonostensive signals are presented: IEC = initial eye contact; ES = eye salience; MS = mouth salience; NEC = no eye contact. Note that in the ES condition (as well as in the NEC condition), none of the eyes was visible at any point during the manipulation phase.

Scoring and data reduction

The principal measure of infant gaze following was whether the child's first saccade from head to object (head–object saccade) was directed at the object that the model gazed at (congruent) or the other object (incongruent) during Phase 3. As is common in this kind of paradigm, difference scores were created by subtracting the number of incongruent head–object saccades from the number of congruent head–object saccades and dividing the outcome by the total number of head–object saccades (see Gredebäck et al., 2008; Senju & Csibra, 2008). The obtained difference scores reflect the same information as proportional scores, which have a range of -1 (all first saccades are directed to the incongruent object) to +1 (all first saccades are directed to the congruent object), where 0 indicates that there is no directional preference relative to where the model gazed. After the data collection, Phase 3 of the gaze replay movie clips was exported with a temporal resolution of 60 frames per second (fps). During scoring, however, gaze replay movies were coded at 15 fps. For each trial, raters scored whether the saccades made were congruent, incongruent, or neither, meaning the head–object saccade did not occur. One third of the movie files were rescored by a coder who was unaware of the study purposes and hypotheses. The intercoder agreement was high (96% agreement, Cohen's $j = .90$). In addition, we determined the total amount of time the infants fixated on each object (congruent and incongruent) during Phase 3 from the eye-tracking data by means of an area of interest (AOI) analysis (see Gredebäck et al., 2008). Similar to Gredebäck and colleagues (2008), we determined a fixation as looking at a certain AOI for at least 200 ms. No post hoc data smoothing or interpolation was applied. We created standard difference scores for this measure, meaning that we subtracted the total amount of looking time to the incongruent object from the total amount of looking time to the congruent object and divided this number by the total amount of looking time to the objects.

An analysis of variance (ANOVA) indicated that across the four conditions (IEC, ES, MS, and NEC), the average amount of time looking at the model's head during Phase 2 (manipulation phase) did not differ, $F(3, 55) = 0.45$, $p = .72$. This indicates that the head region was evenly attractive across conditions. In addition, t tests indicated that infants in the IEC and ES conditions looked at the eye region significantly above chance level, $t(13) = 4.19$, $p = .001$, and $t(14) = 11.97$, $p < .001$, respectively. This was not the case for the infants in the MS condition $t(8) = 1.54$, $p = .16$. This reveals that the manipulation worked as intended; attention of the infants in the MS condition was effectively drawn away from the eyes, whereas this was not the case for infants in the IEC and ES conditions. Then, we checked standard difference scores of the first head–object saccade for normality and equal variances across conditions. The data were normally distributed, having equal variances across groups by condition.

Results

A four-way (Condition) ANOVA was conducted on the standard difference scores of the first head– object saccade. A main effect of condition was found, $F(3, 55) = 3.39$, $p = .025$. Post hoc least significant difference (LSD) tests indicated that infants in the ES condition ($M = 0.20$, $SD = 0.56$) and IEC condition ($M = 0.26$, $SD = 0.38$) did not differ from one another and made significantly more first head–object saccades toward the congruent object than to the incongruent object than infants in the MS condition ($M = -0.19$, $SD = 0.59$) and NEC condition ($M = -0.15$, $SD = 0.38$), which also did not differ from each other. Using standard multiple regression analysis, we checked whether factors other than condition (i.e., age, gender, and order) could account for differences on the standard difference scores. There were no effects of age, gender, or order on these differences, revealing that the differences on the standard difference scores of the first head–object saccade were indeed due to the manipulation, which differed between participants.

In addition, when pooling the standard difference scores of the first head–object saccade of the ES and IEC conditions, a t test revealed that participants in these conditions made the first head–object saccades to the gazed at object significantly above chance level (i.e., different from zero), $t(28) = 2.60$, $p = .014$. This was not the case for the group consisting of infants in the MS and NEC conditions, $t(26) = -1.93$, $p = .065$.

The same ANOVA was conducted on the difference score for total fixation time. Here, the effect of condition was not significant, $F(3, 55) = 0.31$, $p = .82$. This indicates that infants spent equal amounts of time looking at congruent and incongruent objects.

Discussion

The results of the standard difference scores of the first head–object saccade indicate that infants are more likely to follow gaze when initially the eyes either were made salient (ES condition) or were presented frontally (IEC condition) compared with when an animation was presented onto the whole head region (NEC condition) or the mouth was made salient (MS condition). This difference was not due to differential attractiveness of the head across conditions or to other factors such as gender and presentation order. We also found that the infants' first head–object saccade occurred above chance level to the congruent object in the IEC and ES conditions, which was not the case in the NEC and MS conditions. Importantly, the IEC and ES conditions did not differ in this respect. Together, this means that infants were able to reliably follow gaze as long as their attention was drawn toward the eyes either by their (natural) tendency to scan the eye region or by (artificially) attracting their attention toward the eyes. Remarkably, this was corroborated by the results of the novel MS condition in which the eyes of the model were clearly visible. In this condition, the animation placed on the mouth region pulled infants' attention away from the eyes, effectively preventing them from noticing a change in eye gaze.

We also examined the amount of looking time directed at the congruent and incongruent objects across conditions, and we found no difference in looking time across conditions. Several reasons might account for this result. For instance, using a highly similar setup, Gredebäck and colleagues (2008) found differences in looking time to the objects in infants up to 9 months of age but not at 12 months of age. This means that 12-month-olds no longer looked more at the congruent object than at the incongruent object. The authors argued that for this age group, the extent of the infants' explorative (visual) search might be larger. This suggests that as age increases, higher variability in fixations is to be expected. In addition, the objects used in the current study might have been insufficiently attractive for the infants to capture their gaze for more extensive periods of time. The objects were rather plain geometrical shapes instead of attractive toys (e.g., with faces on them), which are commonly used in this type of research (e.g., Gredebäck et al., 2008; Moore & Povinelli, 2007; Senju & Csibra, 2008).

The results of this study emphasize the importance of attention modulation in understanding 20-month-olds' gaze following at the expense of interpretation modulation even when mutual eye contact is established (cf. Gredebäck et al., 2008; Senju & Csibra, 2008). If interpretation modulation were the case, infants must have understood the mutual eye contact in the IEC condition as a communicative intent, which must have resulted in better performance with respect to gaze following compared with the ES condition. This, however, was not found. Therefore, we interpret the results as indicative of an attention mechanism through which initial eye contact modulates the attention of infants in terms of gaze following rather than a mechanism of interpretation of mutual eye contact in terms of forming referential expectations (Senju & Csibra, 2008; Topál, Gergely, Miklósi, Erdo}hegyi, & Csibra, 2008).

We used 20-month-olds in the current study instead of younger infants as was done in many of the previous studies mentioned in the Introduction. This enabled us to investigate specifically the role of the eyes while keeping the face present in an upright and frontal orientation during object presentation. Many researchers have investigated infants' gaze-following skills at ages ranging up to 18 months (e.g., Brooks & Meltzoff, 2002, 2005; Moll & Tomasello, 2004), but few have applied eye movement (as opposed to head movement) as a cue of gaze redirection when the frontal and upright positioned face and target are visually accessible at the same time (Corkum & Moore, 1995). Despite the numerous developmental changes occurring between, for example, 6 months and 20 months of age, we argue that the findings presented here are nonetheless very relevant for our understanding of gaze following in younger infants. The rationale for this is simply that if infants of this age do not appear to use the suggested ostensive communicative cues in this task, is there reason to doubt whether less developed 4-month-olds would do so?

In addition, the current study emphasizes the important role of the top-heavy hypothesis (Macchi Cassia et al., 2004) in drawing the attention of infants toward the face and predominantly the eyes. Although the configurational properties that draw the attention of infants have been investigated extensively, it still remains unclear how the perception of motion of the eyes leads infants to extend that motion in the specified direction. So, although this study highlights that paying attention to the eyes is an important factor in initiating subsequent gaze following, it does not explain why infants continue their gaze in the direction that the model's eyes have moved (but see Doherty, 2006). This is also an issue that clearly deserves further investigation (Farroni et al., 2000).

On a wider scope, this study addresses a fundamental issue with respect to our explanations of infant behavior, such as gaze following, in terms of sprouting social cognition. Often, findings are interpreted in a rich way, that is, by assuming that infants can read other people's intentions and act accordingly (Senju & Csibra, 2008). However, more often than not, it seems that other more basic processes of perception and action play a crucial role in these abilities and, in fact, are closer to the actual mechanisms at hand (see also Gibson & Pick, 2000; Paulus, 2011, 2012; Smith & Thelen, 2003; Spencer, Dineva, & Smith, 2009; Triesch, Teuscher, Deák, & Carlson, 2006). Moreover, it does not mean that these mechanisms actually reflect the sociocultural origins of infant behavior (i.e., that they understand or read intentions in the strict sense that is often referred to). It is highly probable that they merely reflect the way in which we tend to adopt a sociocultural context in our explanations of infant behavior (see, e.g., Churchland, 1981; Haith, 1998).

Chapter 3

Attunement and affordance learning in infants

P. F. de Bordes
F. Hasselman
R. F. A. Cox

Published as:
De Bordes, P. F., Hasselman, F., & Cox, R. F. A. (2019). Attunement and affordance learning in infants. *Journal of Cognition and Development*, *20*(4), 534-554.
https://doi.org/10.1080/15248372.2019.1626398

Abstract

From a perceptual learning perspective, infants use social information (like gaze direction) in a similar way as other information in our physical environment (like object movements) to specify action possibilities. In the current study, we assumed that infants are able to learn an affordance upon observing an adult failing to act out that affordance, without appreciating object-directed intentions, or, communicative intent towards the infant. Using a variation of the Re-enactment procedure, we found that when the attention of infants ($N = 46$, $M_{age} = 20$ months) was drawn towards the eyes of the person before she acted out the failed attempt, either by ostensive cues or non-ostensive cues, infants achieved more affordances than 15 when their attention was not directed towards the eyes. As directing the attention of infants to the eyes of another person frequently results in gaze following, this suggests that infants use the gaze direction of another person in order to learn what affordance that other person is trying to realize. In addition, the results of a spatiotemporal analysis on the eye movements of infants suggest that the gaze and the object movements of the person facilitate learning by directing the attention of infants towards important object-directed actions on crucial moments during the failed attempt demonstrations. These results are discussed in terms of perceptual attunement and affordance learning.

Introduction

Adults tend to actively facilitate learning contexts for infants such that it allows them to acquire skills like the use of tools, language and cultural practices (Tomasello, 1999). These learning contexts are commonly facilitated by directing the attention of infants towards objects of interest (by means of, e.g., pointing and gaze) and by demonstrating the way how certain objects or tools can be used (e.g., Gergely, Bekkering, & Király, 2002; Hopper, 2010; Király, 2009; Meltzoff, 1995; Nielsen, 2006; Southgate, Chevallier, & Csibra, 2009). A longstanding debate is whether infants need to understand or interpret these ostensive signals as communicative or infer someone's intentions or goals while that person demonstrates how certain objects can be used, in order to learn from that person effectively (Csibra & Gergely, 2009; de Bordes, Cox, Hasselman, & Cillessen, 2013; Doherty, 2006; Haith, 1998; Heyes, 2012, 2016; Hopper, 2010; Paulus, 2011; Want & Harris, 2002). In developmental research, learning by infants in non-social situations has been frequently explained by what aspects of the task or the task materials infants need to understand in order to produce a certain goal (e.g., Dejonckheere, Smitsman, & Verhofstadt-Denève, 2007; Huang & Charman, 2005; Thelen, Schöner, Scheier, & Smith, 2001). Conversely, superior learning outcomes in social settings are usually explained in terms of what infants understand of the (social) behavior of others (e.g., Topál, Gergely, Miklósi, Erdőhegyi, & Csibra, 2008; see Hopper, 2010 for an overview). Although it seems straightforward that infants do not need to understand the (social) behavior of others when they successfully learn something in non-social contexts, relatively few studies have assessed how infants can learn from others without them having mentalistic assumptions about that other person (Goldstein & Schwade, 2008; also see Hellendoorn, 2014; Heyes, 2012). Therefore, the current study sought to identify a non-mentalistic and perceptual basis for how infants can learn certain affordances when an adult draws their attention towards objects and demonstrates how they can be used. An affordance in this sense is defined as an action possibility provided by the social and/or material environment given the action capabilities of the infant (also see Gibson, 1979, 1988; Gibson & Pick, 2000; Ishak, Franchak, & Adolph, 2014; Michaels & Carello, 1981).

In Meltzoff (1995) study, 18-month-olds viewed a woman acting out either a failed attempt or a successful attempt in combining several object pairs after which they tended to combine the object pairs successfully in both conditions at similar rates. Meltzoff reasoned that as the infants had never seen the fulfilled object-directed actions in the failed attempt demonstration, they must have been able to detect the model's intended goal that they subsequently imitated (Meltzoff, 1995; for similar findings, see Bellagamba & Tomasello, 1999; Danish & Russell, 2007; Johnson, Booth, & O'Hearn, 2001). As an alternative to this interpretation of the data, Huang and Charman (2005) have suggested that the object movements in the demonstrations might contain sufficient perceptual information for infants to realize the affordance as an action possibility (i.e., the goal or endstate of the task). Huang and Charman (2005, experiment 3) introduced a non-social 'ghost' condition of the task used by Meltzoff (1995) in which they showed infants a demonstration of the failed attempt where the actor handling the objects was not visible. It was found that the success rate of infants realizing the affordance after having watched the object movement demonstration was similar to the success rate after watching a failed attempt demonstration in which a model acted out the object-directed actions. They concluded, therefore, that mental state attribution or interpreting social cues is not necessary for infants to realize the target affordance under these conditions (for similar findings, see Horne, Erjavec, & Lovett, 2009; Huang, Heyes, & Charman, 2002; Shneidman, Todd, & Woodward, 2014; Thompson & Russell, 2004). According to Huang and Charman (2005), a perceptual mechanism that might enable infants to realize the target affordance after seeing the failed attempt demonstration or the 'ghost' failed attempt demonstration could be stimulus enhancement (also see Charman & Huang, 2002).

Stimulus enhancement is a concept used in both developmental and comparative research and refers to the process in which the attention of a person or animal is drawn towards an object or object part, simply because someone or something is interacting with that object, increasing the chance that the observer will interact with that object or object part thereafter (e.g., Fritz, Bisenberger, & Kotrschal, 2000; Heyes, Ray, Mitchell, & Nokes, 2000; Hopper, 2010; Hoppitt & Laland, 2013; Horne et al., 2009; Huang & Charman, 2005; Huang et al., 2002; Want & Harris, 2002; Zentall, 2001). In Huang and Charman's (2005) non-social 'ghost' condition, the object movements and the way the object pairs were brought in close proximity apparently specified the appropriate action possibilities for the infants (i.e., how they can be combined in specific ways). This enabled most of them (but not all) to realize the target affordance. This suggests that stimulus enhancement can specify affordances for those objects without the necessity to perceive the completion of the actions afforded by those objects. Interestingly, most infants in the study by Meltzoff (1995) realized the target affordance after seeing the failed attempt demonstration, but not after they viewed someone handling the objects in a non-specific way during the adult manipulation demonstration. During the failed attempt demonstration, relevant parts of the objects were manipulated with the relevant actions (e.g., conveying a stick towards a hole in which it can be placed), whereas this was not the case during the adult manipulation demonstration (e.g., picking up the stick and placing it on the table), leading infants in the former but not the latter condition to realize the target affordances (Charman & Huang, 2002). This suggests that affordance learning can occur through stimulus enhancement as long as it specifies a certain action possibility to the observer in relation to the stimulus as opposed to guiding the observer's attention towards an object (part) in a more general (i.e., non-specifying) way. In this sense, the way stimulus enhancement leads to affordance learning closely resembles to what Gibson (1979) defined as 'education of attention' in which "the perceptual system is attuned to 'picking up' critical features of the environment" (Ingold, 2001, p. 137).

In our view, there are many ways in which attention can be attuned to critical features of the environment from which affordances can be learned. These can be social such as pointing and gaze (de Bordes et al., 2013; Goldin-Meadow, 2007), non-social such as handling or dropping an object (Fritz et al., 2000; Zentall, 2001), and inanimate such as mechanical movements (Dejonckheere et al., 2007) or animations (Grant & Spivey, 2003). In natural social situations, infants ranging between 12 and 18 months of age tend to look at the hand movements and the gaze of the caregiver from which they can learn about what the caregiver is doing and more general leads to joint attention (Yu & Smith, 2013, 2017). Infants in this age group (12 to 18 months of age) tend to follow the hand movements to a greater extent than gaze direction (Yu & Smith, 2013; also see Deák, Krasno, Jasso, & Triesch, 2018, for similar results across the age range of 3 to 11 months of age). They might follow hand movements to a greater extent than someone's gaze because only from 18 months of age, infants can follow the direction of eye gaze reliably (Corkum & Moore, 1995; de Bordes et al., 2013). In addition, following hand movements seem more economic as that might be more indicative of what affordances someone is trying to realize rather than following gaze alone (see Yu & Smith, 2013, 2017). However, because gaze and hand movements are coupled (e.g., Adam, Buetti, & Kerzel, 2012), observing the gaze of another person can possibly provide additional information on what affordances that person is trying to realize than observing hand movements alone (also see Sebanz, Bekkering, & Knoblich, 2006). For instance, infants can use gaze direction in order to anticipate object-directed actions (Paulus, 2011), potentially leading them to observe (parts of) objects and locations on moments that are crucial for understanding the affordance that is being realized. Therefore, we suggest that observing the gaze of a person could support the perceptual system of infants in attuning to critical features of the environment, including certain object movements, by which affordances can be learned from that person, even when the fulfilled affordance is not observable, as is the case when demonstrating a failed attempt.

Gaze following can be instigated by attracting the attention of the infant towards the eye region of another person (e.g., de Bordes et al., 2013). Specifically, research has shown that whereas 6-month-olds can reliably follow head turns after their attention has been drawn towards the head of another person (Gredebäck, Astor, & Fawcett, 2018; Senju & Csibra, 2008; Szufnarowska, Rohlfing, Fawcett, & Gredebäck, 2014), 18 to 20-month-olds can use the affordance of follow eye movements of another person when their attention has been drawn towards the eye region, rather than the face in general (Corkum & Moore, 1995; de Bordes et al., 2013). In case of a failed attempt demonstration, in which the target affordance is not realized, attracting the attention of infants towards the eyes prior to a failed attempt demonstration could lead infants from 18 months of age to use the gaze of that person. This could direct their field of view towards the actions performed by the observed person in a more specifying manner than if they would look at the object movements alone, as suggested above. In the re-enactment procedure, for instance, observing someone looking towards the hole to which that person is conveying a stick might indicate that the stick can be placed in the hole.

In order to test the idea that gaze following promotes affordance learning when a model demonstrates certain object-directed actions, we used Meltzoff's (1995) failed attempt condition which does not include the end-state of the object-directed actions. We did not include the end-state of the object-directed actions in order to specifically assess affordance learning in this situation and not the mere reproduction of the goal as would be the case in goal emulation (Hopper, 2010). We varied the likelihood of infants following the gaze direction of the mode during the failed attempt demonstration by manipulating the eye region of the model immediately prior to the demonstration, similar to de Bordes et al. (2013).

Specifically, we attracted infants' attention to the eyes of the model by leaving them either visible (Eye Contact condition) or by presenting moving and flashing dots on top of the eyes (Eye Salience condition). We added a third condition in which the models' face was covered completely with a flower animation (No Eye Contact condition), drawing attention to the facial region rather than the eyes alone. Following from our reasoning that gaze following assists infants in discovering the object affordance, we expected infants to have a bigger chance to realize the target affordance after their attention was drawn towards the model's eye region prior to the failed attempt demonstration than if their attention was directed towards the head region in a non-specific way. This would lead to higher success rates in the Eye Contact and Eye Salience conditions compared to the No Eye Contact condition. In addition, we expected no difference in gaze following and subsequent affordance realization between the Eye Contact and Eyes Salience conditions, which would replicate earlier findings of the attention modulating nature of gaze following (de Bordes et al., 2013; also see Gredebäck et al., 2018) and generalize it to the Re-enactment context.

An additional aim of this study was to explore the spatiotemporal structure of infant looking behavior, and how this is related to the specification of the target affordance, as a function of condition contrasts and whether they successfully realized the target affordance or not. To this end, we have included eye-tracking measures of the infants while they watched the failed attempt demonstrations. We speculated that infants who were successful at realizing the target affordance might be inclined to look at the object movements during times in which those object movements could specify the object affordance whereas infants who were unsuccessful would look at the object movements in a more unspecific manner. Together, this will help us to better understand how hand movements and the gaze behavior of adults can specify affordances to infants in this context and how it relates to affordance learning in general.

Methods

Participants

Addresses of 510 parents of 20-month-old infants in the city of Nijmegen, the Netherlands, were provided by local government authorities. A total of 61 parents (12%) responded to the invitation with a confirmation and signed informed consent. All 61 infants were included in the study (M_{age} = 20 months and 2 days, *SD* = 17 days; 30 girls and 31 boys).

Stimulus material & procedure

Upon arrival, infants and parent(s) were brought to the testing room. After a brief warming-up period, infants sat on their parents' lap in a room with dimmed light, approximately 60 cm away from a 17-in. TFT monitor (60 Hz) with a built-in remote eye-tracking system (Tobii T120, Tobii Technology, Danderyd, Sweden). The 31.2 by 22.9-cm screen and the distance from it created a 30° by 21° visual angle with a screen resolution of 1280 × 1024. We used Clearview software (Version 2.7, Tobii Technology) to calibrate the eyes of the infants (on a nine-point calibration program), to present stimuli, and record eye movements of the infants. Before the experiment started, infants participated in another experiment reported elsewhere (de Bordes et al., 2013). Each trial was preceded by an attention grabber displaying colorful animations with bleeping sounds until the infants attended to the center of the screen. Trials consisted of videos of failed attempt demonstrations with a female model manipulating five different object pairs identical to those used in Meltzoff's (1995) study. The objects were 1) a dumbbell-shaped toy that could be pulled apart and put together again ('Dumbbell'), 2) a box with a hole in which a wooden stick could fit to activate a buzzer ('Box-Stick'), 3) a loop that could be draped over a horizontal prong that was attached to a vertically standing board ('Loop-Prong'), 4) a chain of beads that could be placed in a cylinder-shaped container ('Beads') and 5) a wooden square with a cylinder on top on which a plastic square with a round hole could be placed ('Square-Dowel'). A detailed description of the geometrical properties of these objects is given in Meltzoff (1995).

Each failed attempt demonstration was shown in three sequential phases (Figure 3.1). In the first phase, the model's face oriented downward for two seconds. In the second phase (manipulation phase), she looked up and straight ahead for two seconds. In this second phase, the same basic video was used for all three conditions but with modifications between them, creating different conditions across infants. In the Eye Contact condition (EC) the model moved her head up and looked straight ahead for two seconds (unmodified). In the Eyes Salience condition (ES) colorful blinking and moving dots overlaid the model's eyes during the time she looked up and straight ahead. This rendered the eyes themselves invisible but the eye region salient. In the No Eye Contact condition (NEC) a colorful animation of a flower overlaid her head of the model while she looked up and straight ahead, obscuring the whole face region. The third phase was identical across conditions and consisted of the model acting out three failed attempts demonstrations in succession while having her hands, arms, torso and head in full display. This typically consisted of placing an object (i.e., dynamic object) closely to another object (i.e., static object) as will be explained below. After each demonstration (lasting for approximately seven seconds), the objects were restored to their initial position before the next demonstration begun (also called the 'retrieval phase') until three demonstrations were finished.

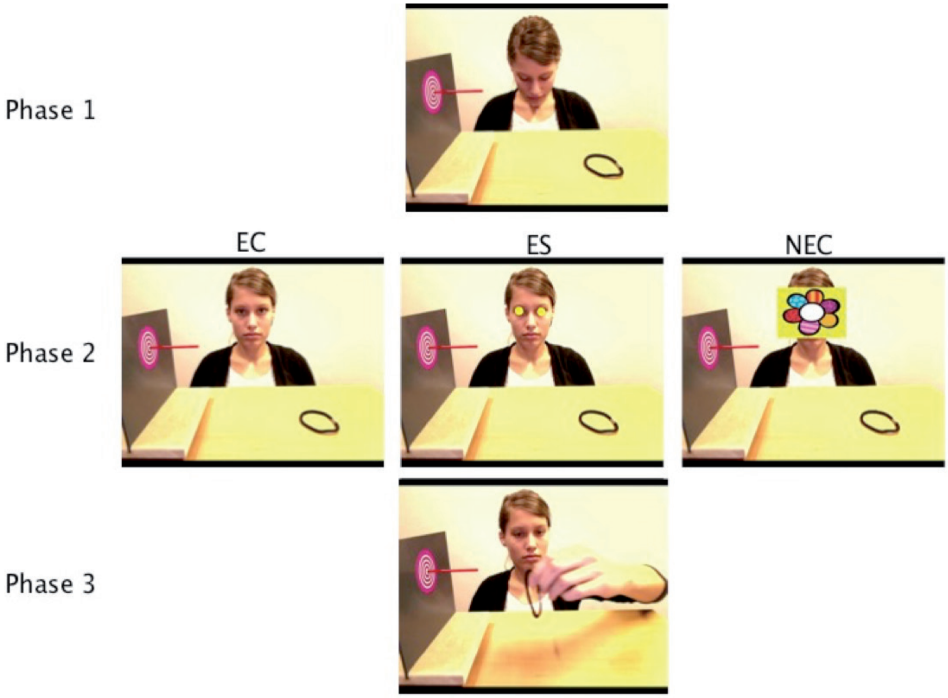

Figure 3.1. The manipulation phases for each of the conditions EC, ES and NEC. Conditions across children only differed in phase 2.

For the Dumbbell, the failed attempt consisted of the model picking up the object with both hands and attempting to pull one of the wooden blocks sideward. However, instead of the wooden block being detached from the stick this resulted in slipping of the fingers over the block and a lateral movement of only the hand in that direction. Slippage was alternated from the left to the right, and back to the left again.

For the Box-Stick, the model would pick up the stick and move it towards the hole in the box, where it would land right next to the hole. In the first failed attempt, she placed the stick to the left side of the hole, in the second on the right side, and in the third-failed attempt, just above the hole.

For the Loop-Prong, the model picked up the nylon loop, moved it towards the prong where she would drop the loop on the table. She first moved the loop to the left of the prong, then to the right, and finally below the prong.

For the Beads, the model picked up the chain of beads and moved it towards the cylinder where she dropped the chain on the edge of the cylinder, after which the beads fell down on the table next to the cylinder.

For the Square-Dowel, the model picked up the plastic square and placed it onto the dowel but failed to place the hole of the square over the cylinder on the dowel, so as to fit the objects together.

In all demonstrations, the model's eye gaze was always directed at the moving object until it approached the goal region on the stationary object. At that moment, the model's eye gaze was directed mainly at the goal region until the objects were placed back towards their initial location. To control for effects of laterality, the movement direction for each object was counterbalanced between infants. That is, half of the infants viewed a left-to-right movement on the second and fourth trial and a right-to-left movement on the first, third and fifth trial, whereas for the other half this was reversed. Immediately after each set of failed attempts, the screen went blank, and the experimenter placed the objects pairs on the table in front of the infant while uttering 'Look'. A small webcam (Sonix SN9C201, Taiwan) next to the screen recorded infants' behavior after each demonstration. After the five trials were finished, parents and infants were thanked for their cooperation and the parents were debriefed.

Behavioral scoring and data reduction

A total of 15 infants were excluded from the analysis due to experimenter error (1) or equipment failure (4), inattentiveness of the infant[1], or refusal to touch one or more object pairs within the first 20 seconds (7) or interference by the parent (3). As a primary measure, we scored whether the remaining 46 infants with valid data performed the target act (yes/no) within a 20-second response period, starting when they first touched the object (similar to the studies reported by Bellagamba & Tomasello, 1999; Huang & Charman, 2005; Meltzoff, 1995). For the Dumbbell, the goals consisted of pulling one of the squares apart from the rest of the object. For the Box-Stick, the goal was to insert the stick into the box, which then would activate the buzzer inside. We noticed that infants experienced difficulties activating the buzzer, because the hole was too deep and/or the button inside to rigid. Therefore, we counted the behavior as correct whenever an infant placed the stick partly into the hole (i.e., when all four corners of the end of the stick were inside the hole). For the Loop-Prong, the goal was to drape the loop onto the prong. For the Beads, the goal was to place the chain of beads in the cylinder. Finally, for the Square-Dowel, the goal was to place the square onto the dowel. Next, we made proportion scores of the total amount of target acts completed per infant. Whether infants successfully produced the actions afforded by the object was scored live by an experimenter. In addition, one-third of the trials were scored after the data were collected by someone who was unaware of the goals of the study. There were no discrepancies between the two coders (Cohen's $\kappa = 1$), meaning that successes and failures were clear. Finally, we measured the time in seconds infants needed to complete the target acts with the object pairs whenever they were successful.

Eye-tracking measures; spatial regions of interest analysis

In order to analyze what infants looked at during the failed attempt demonstrations with the different object pairs, we identified several regions of interest (ROI) in terms of x and y coordinates on the screen (see Figure 3.2). The ROI used were 1) the eye region of the model, 2) the complete static object in the scenery, 3) the specific region of the static object that can be combined with the dynamic object (goal region), 4) the dynamic object that can be combined with the static object and finally, 5) all else. We tracked the movement of the dynamic object frame-by-frame across the demonstrations with different object pairs, using Adobe After Effects (Adobe After Effects CC, Adobe, 2015). Using MATLAB (MATLAB, Version 9), we registered to which ROI infants were looking on each timeframe at 60 Hz. Whenever there was an overlap in looking at the dynamic object and another ROI, only looking at the dynamic object was registered. The duration infants looked at each ROI

[1] Some infants had a preference to explore the room or interact with the parent instead of attending to the experimental stimuli.

during each failed attempt demonstration was calculated by adding up all timeframes in which infants looked at the pixel coordinates that corresponded with the ROI per object pair. Next, we averaged the looking time to each ROI across the object pairs for each infant that was used to calculate the percentages of looking time towards each ROI for infants in each condition. Only the object pairs 'Box-Stick', 'Loop-Prong' and 'Beads' were included in the eye movement analysis because the demonstrations for these three object pairs extended across a large portion of the screen, eliciting clearer eye movements in the infants.

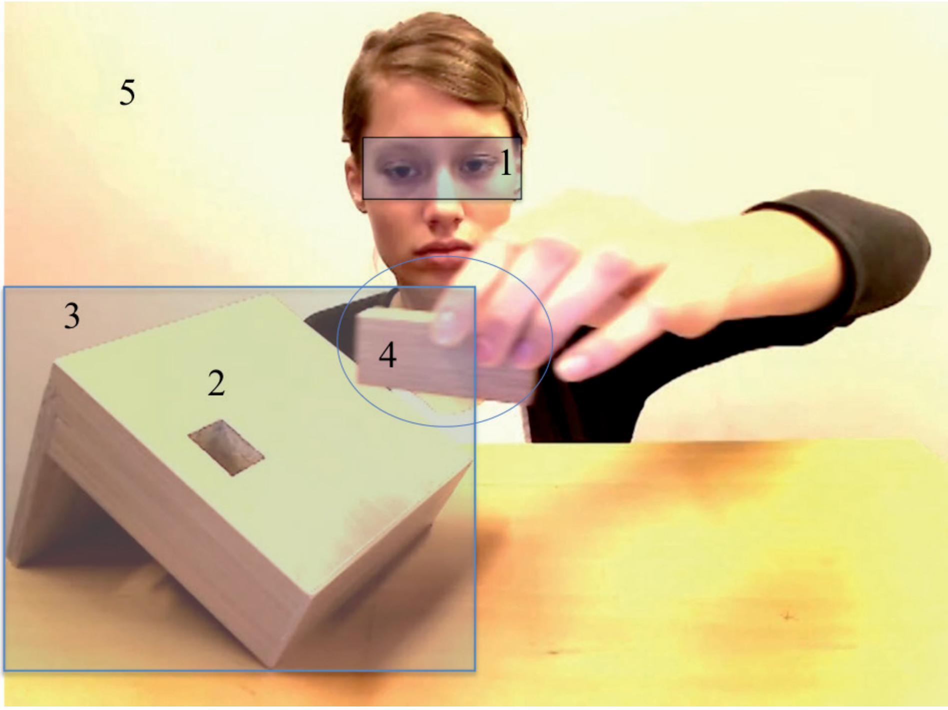

Figure 3.2. The regions of interest consisting of 1) the eye region, 2) the goal region of the static object, 3) the whole static object, 4) the dynamic object and, 5) all else.

Eye-tracking measures: spatiotemporal regions of interest analysis

Individual time series of eye movements were created by registering to which ROI each infant was looking at for each timeframe. Then, we collapsed these individual time series in percentages across conditions and according to whether infants realized the target affordance of not. These time series allowed us to perform a descriptive analysis on the spatiotemporal order in which infants looked at the different ROIs. Further, for each ROI at each timeframe for all infants, we calculated the 95% confidence intervals to assess whether infants in each condition would look at a certain ROI at a certain time significantly more than would be expected from the total amount of infants included in this study. To this end, the percentages of infants looking at each ROI per time frame were first transformed into proportions within a Z-distribution. Next, difference score times series were created by subtracting the proportion of looking time at each ROI at each timeframe of unsuccessful infants from successful infants. The resulting difference score indicates in which time frames more successful infants looked at a particular ROI than unsuccessful infants (indicated by a positive proportion) and vice versa (indicated by a negative proportion). Then, a confidence interval of 95% was calculated for the percentage of infants looking at each ROI per time frame. This was subsequently used to compare the looking proportions to ROIs at each time frame for successful and unsuccessful infants across the three experimental conditions.

Finally, in order to determine whether condition type influenced these difference scores between correct and incorrect, a permutation test was applied to the following ROIs: 1) the eye region of the model, 2) the specific region of the static object that can be combined with the dynamic object (goal region), 3) the dynamic object that can be combined with the static object. The temporal order of values of the difference scores time series was resampled 999 times after which the rank of the observed difference score among the 999 resampled difference scores for each time point (i.e., ranging from 1 to 1000) was evaluated. Specifically, a p-value was calculated by dividing the number of difference scores that were equal to the observed difference or more extreme by the number of values in the distribution. The alpha level was adjusted from .05 by a factor 3 for multiple comparisons, taking into account that the looking at one of the three analyzed ROIs affects the timing and duration of looking to another ROI. Because the observed time series are autocorrelated, the method of random block size resampling (or, stationary bootstrap, cf. Politis & Romano, 1994) was applied. First, a time series is covered with blocks of different sizes that are randomly drawn from a geometric distribution (the mean expected block size was 4, this was based on an inspection of the partial autocorrelation functions of the observed time series, which had significant correlations up to lags of 3, 4 and 5). Second, the surrogate time series for the permutation tests is generated by randomizing the order of the blocks, preserving the temporal order of values within each block. Finally, the ranks of observed and surrogate differences are calculated, yielding a p-value for each timepoint (see, e.g., Vink, Hasselman, Cillessen, Wijnants, & Bosman, 2018).

Results

Behavioral measures

Even though they had only seen failed attempts by the model, on average, infants in the EC (Eye Contact) condition correctly performed 64% of the actions afforded by the objects.

Infants in the ES (Eye Salience) and NEC (No Eye Contact) condition correctly performed 65% and 46% of the target acts, respectively. This overall success rate of performing the target acts across the conditions is comparable to Huang and Charman (2005, experiment 3, object movement failed attempt) and Huang, Heyes and Charman (2002, failed attempt condition). A simple regression analysis revealed that the proportion of target acts produced was not influenced by age, sex, or effects of laterality (all predictors n.s.).

To test whether the proportion of target acts completed differed between conditions, we fitted mixed effects logistic model for binary outcomes (incorrect = 0, correct = 1) using the R package lme4 (Bates et al., 2015; R Core Team, 2016). Condition was entered as a fixed factor (taking NEC as the reference category) in a model with random intercepts for target objects and participants. The estimated odds of correctly performing the target action in the NEC condition did not deviate significantly from 50% ($\beta_{intercept} = -0.15$, $SE = 0.22$, Wald Z = -0.67, $p = .5$, OR = 0.86, 95% CI [0.54, 1.29]). However, relative to the NEC condition, the odds of producing a correct action increased significantly for the EC condition ($\beta = 0.73$, $SE = 0.33$, Wald $Z = 2.21$, $p = .02$, OR = 2.07, 95% CI [1.08, 4.29]) as well as the ES condition ($\beta = 0.178$, $SE = 0.33$, Wald $Z = 2.37$, $p = .02$, OR = 2.19, 95% CI [1.22, 4.40]). The odds ratios did not differ across the ES and EC conditions, as can be seen in Figure 3.3, which displays the estimated OR with 95% bootstrapped confidence intervals for each condition.

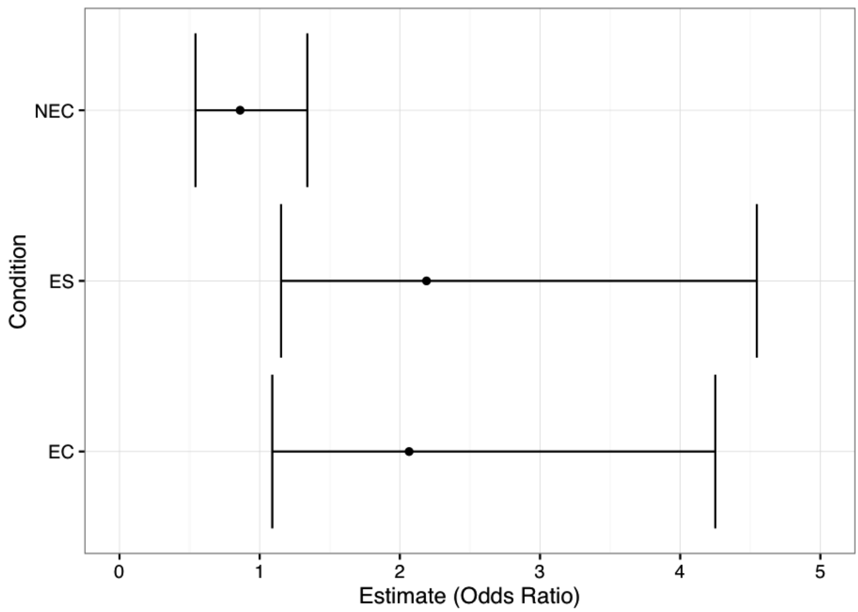

Figure 3.3. Odds Ratios of effect estimates of the mixed effects model for each of the conditions EC, ES and NEC with 95% bootstrap Confidence Intervals (1000 bootstrap simulations).

In addition, a mixed effect model with subjects and objects as random factors showed that infants in the EC condition and ES condition combined (M = 6.6 seconds, SD = 3.4 seconds) needed significantly less time to complete the target acts than infants in the NEC condition (M = 9.6 seconds, SD = 4.8 seconds) whenever infants made a successful attempt (β_{EC+ES} = -2.41; SE = 1.13; Satterthwaite's method: $t(42.6) = -2.14$, $p < .039$).

Eye-tracking measures: spatial regions of interest analysis

As can be seen in Figure 3.4, there were no differences across conditions for the total amount of looking time to the different ROIs (all differences n.s.). This reveals that making the eyes

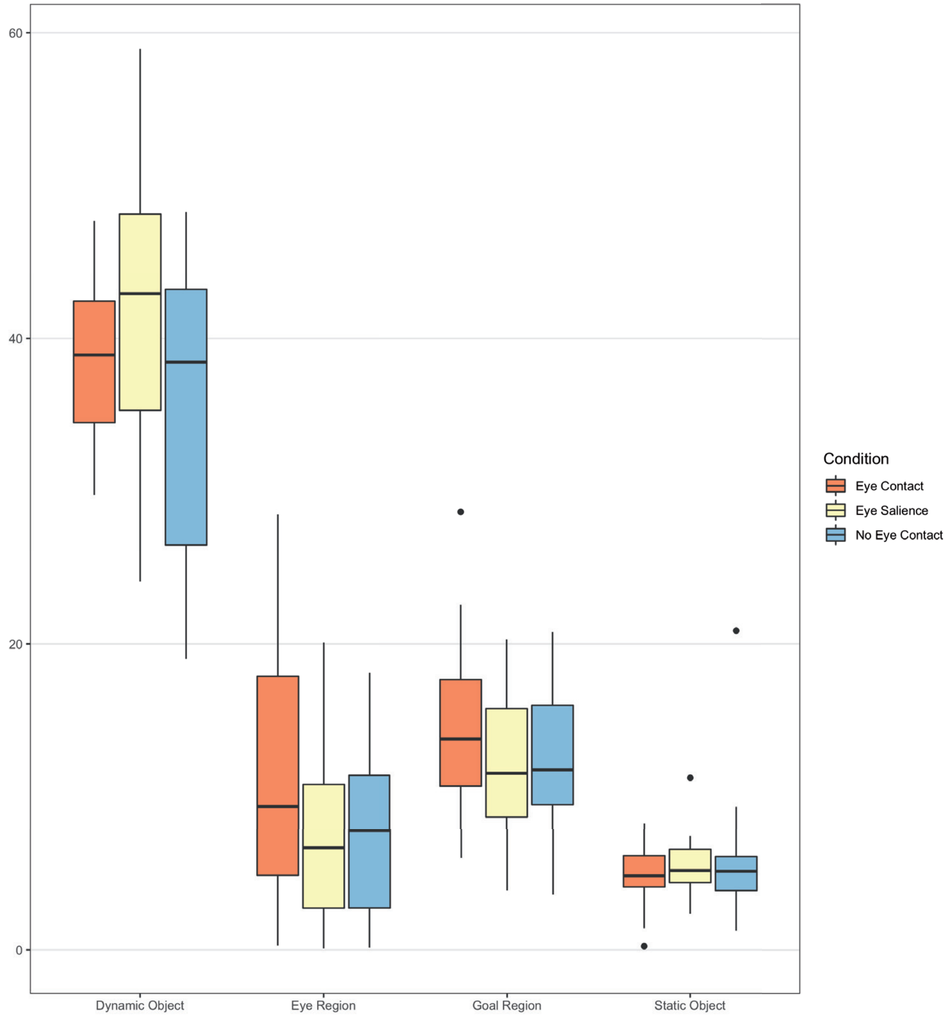

Figure 3.4. Average percentage of time spent looking at a Region of Interest relative to the total amount of looking time of infants across conditions.

of the model salient (naturally or artificially) or covering the whole face area prior to the failed attempt demonstration did not affect the duration of time infants spend looking at the different ROIs thereafter. In addition, we also did not find any difference in looking duration towards the ROIs during the failed attempt demonstrations between successful and unsuccessful attempts (all differences n.s.).

Eye-tracking measures: spatiotemporal regions of interest analysis

Figure 3.5 depicts the percentage of infants looking at each ROI[2] per time frame during the failed attempt demonstrations for infants who correctly performed the target affordance (Figure 3.5 upper panel) and for those who did not (Figure 3.5 lower panel). Most infants tended to look at the eye region during the manipulation phase (where the conditions differ), after which they followed the object movements as the model conveys one object to the other three times in succession. Remarkably, the eye movement patterns are quite similar for both groups, irrespective of whether the infant is successful or unsuccessful in realizing the target affordance right after the demonstration. This reveals that although the conditions yielded significant differences in terms of affordance realization, the effects of these manipulations on infants' looking behavior when watching the failed attempt demonstrations must be quite subtle.

In Figure 3.6, the time courses of successful and unsuccessful attempts as displayed in Figure 3.5 were z-transformed and turned into a difference series in which positive z-scores indicate relatively more infants who produced a successful attempt were looking at the ROI and negative z-scores indicate relatively more infants who produced an unsuccessful attempt were looking at the ROI. The z-scores were calculated following the procedure one would follow conducting a z-test of observed versus expected proportions. There were five mutually exclusive ROIs, so we considered the expected proportion of infants looking at a specific ROI n, at a specific point in time t, denoted as ROI(n,t), to be 1/5 (p_0), a priori, and consequently, not looking at ROI(n,t) to be 1-p_0. The observed proportion (p_1) was simply the number of infants observing ROI(n,t) divided by the total number of infants observing an ROI at t (N_t). From this information a proportion Z-score can be calculated whose magnitude reflects the deviation of the observed proportion from the expected proportion as $Z(n,t) = (p_1 - p_0)/\mathrm{sqrt}(p_0 * (1-p_0)/N_t)$. The time courses of successful and unsuccessful attempts were standardized separately in order to account for the differences in N_t between infants producing successful and unsuccessful attempts when rescaling the data. By design, the

2 Note that for clarity of presentation, the ROIs 'Static Object' and 'Other' were not included in Figure 3.5, Figure 3.6, or in the discussion of the results, because the groups did not display any relevant difference in timing or amount of looking for this ROI.

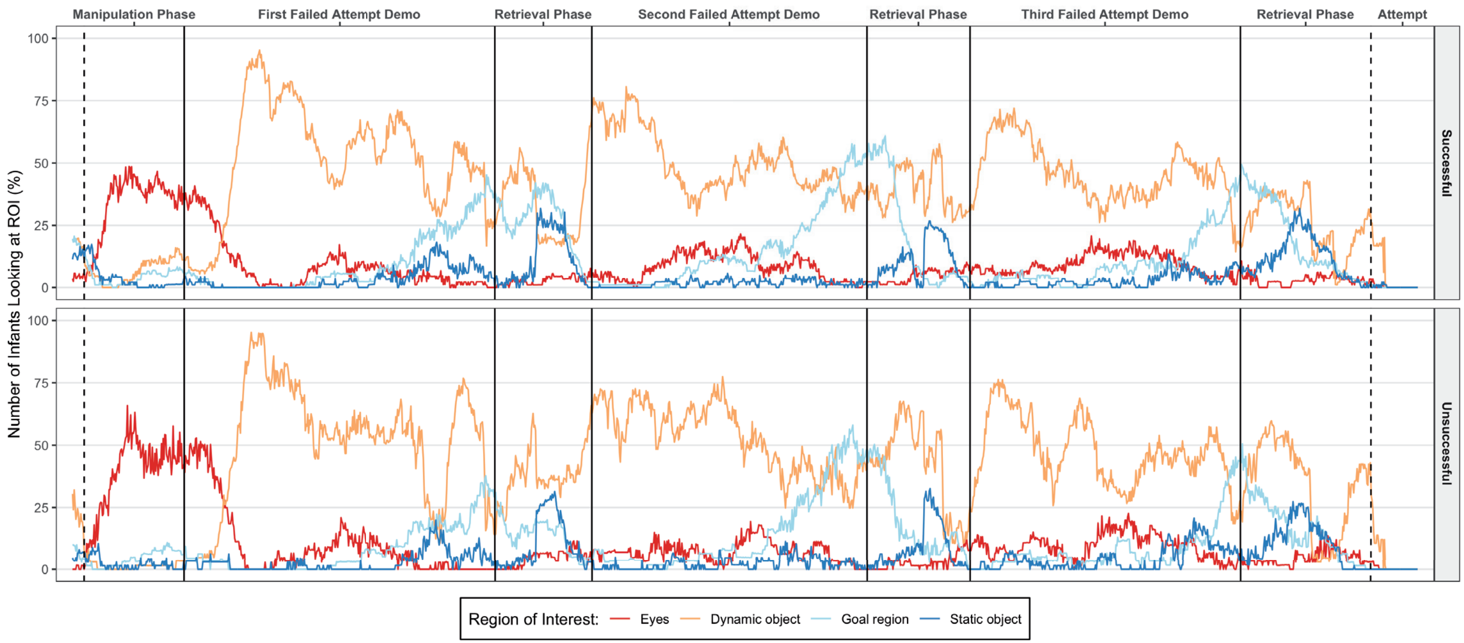

Figure 3.5. Time series displaying the percentage of infants looking at an ROI (different lines) at a given time, averaged across conditions. The upper panel shows infants successful in affordance completion, the lower panel the infants that were unsuccessful.

difference between these Z-scores (correct – incorrect, $Z_c - Z_{inc}$) at time t, reflects whether an ROI(n,t) was observed more often than expected, by infants who made a correct attempt (positive Z-values) or an incorrect attempt (negative Z-values) at realizing the affordance.

In order to provide some notion of whether these differences were substantial, we focus only on values outside the range of $-1.96 < Z < 1.96$, or,a difference of about two standard deviations in both directions. In Figure 3.6 all values falling within this range are masked by a grey panel on top of which line marks are drawn whose color corresponds to the ROIs of which $Z_c > Z_{inc}$ (positive Z) or $Z_c > Z_{inc}$ (negative Z). In effect, the two-colored bands in the center of each graph provide a description of the average time course of ROIs that most likely produced successful and unsuccessful re-enactments of the model's failed attempts in each condition.

Figure 3.6 reveals various differences between the conditions and between the successful and unsuccessful groups. One noticeable difference is that successful infants were looking more towards the dynamic object as it moved towards the goal region than unsuccessful infants during the failed attempt demonstrations. This difference was significant for 798 timestamps (≈ 13 seconds) on average per failed attempt demonstration. In addition, successful infants tended to look more at the goal region when the model conveyed the dynamic object in close proximity to it. Remarkably, successful infants also tended to look more at the goal region significantly more during the retrieval phase. On the whole, successful infants looked at the goal region significantly more than unsuccessful infants for 270 timestamps (≈ 4.5 seconds) on average per failed attempt demonstration during the retrieval phase. Instead, unsuccessful infants tended to look more at the dynamic object for 210 timestamps (≈ 3.5 seconds) when the model retrieved the object towards the starting position of the failed attempt demonstration. The aforementioned differences, however, seem only to be the case across successful and unsuccessful infants within the EC and ES conditions. On the whole, differences in looking behavior (i.e., differences in proportions of looking at certain ROI at certain timeframes) between successful and unsuccessful infants are smaller for infants in the NEC condition, compared to the EC condition ($P_{EC} = 0.58$, $P_{NEC} = 0.24$; $Z_{EC\text{-}NEC} = 20.96$, $p < .001$) and ES condition ($P_{ES}= 0.57$, $P_{NEC} = 0.24$; $Z_{ES\text{-}NEC} = 20.82$, $p < .001$). In addition, the differences in looking behavior between successful and unsuccessful infants across the EC and ES conditions were comparable ($P_{EC} = 0.58$, $P_{ES} = 0.57$; $Z_{EC\text{-}ES} = 0.60$, $p = .33$).

To assess the influence of condition type on the difference scores between correct and incorrect within each condition, a permutation test was used in which we focussed on difference scores in looking at the 1) eyes of the model, 2) the dynamic object and 3) the goal region where the dynamic object and static object can be combined. Figure 3.7 displays

3

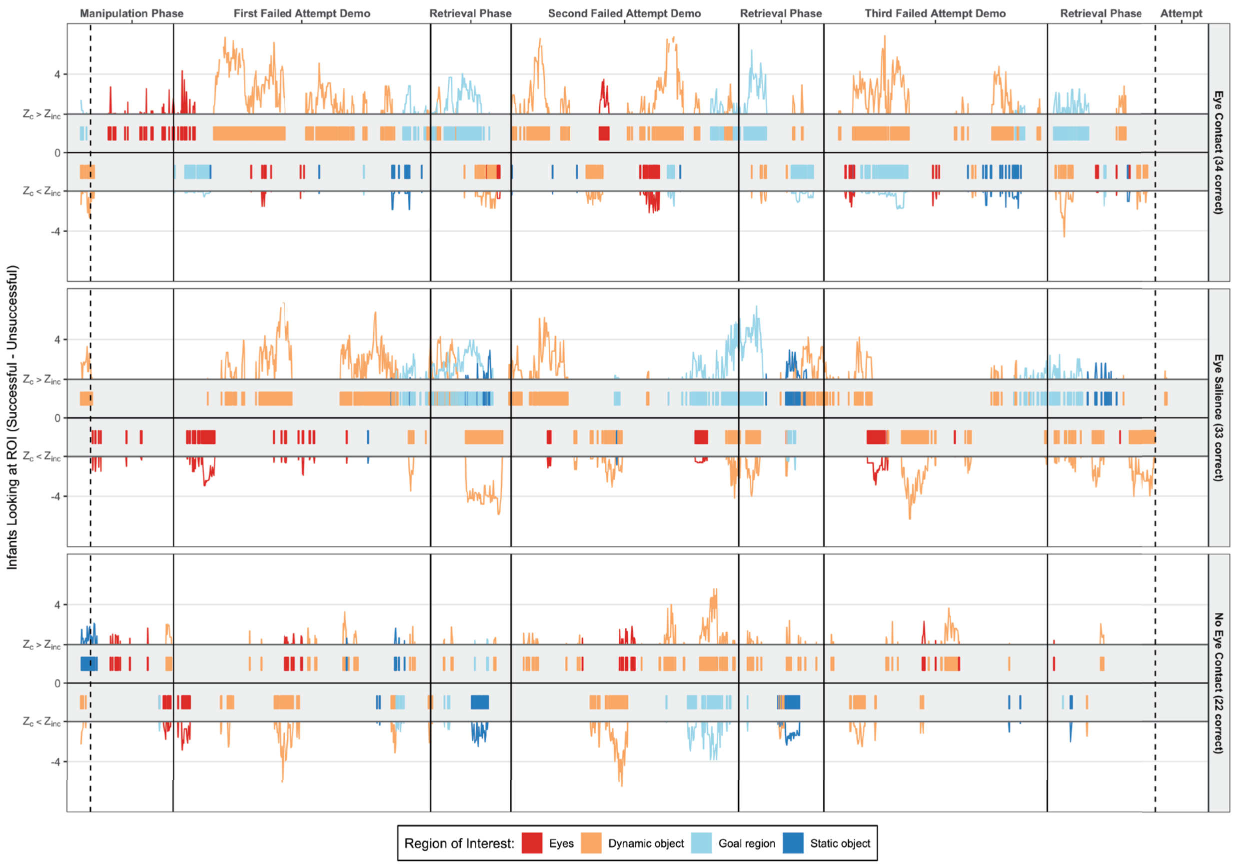

Figure 3.6. Difference score times series of successful infants (top of each panel) versus unsuccessful infants (bottom of each panel) across all conditions. The y-axis represents the proportion Z scores. The EC condition is in the top panel, ES condition in the middle panel, and NEC in the bottom panel. The color codes in the center of the graphs indicate whether significantly more successful or unsuccessful looked at a specific ROI (see text for details).

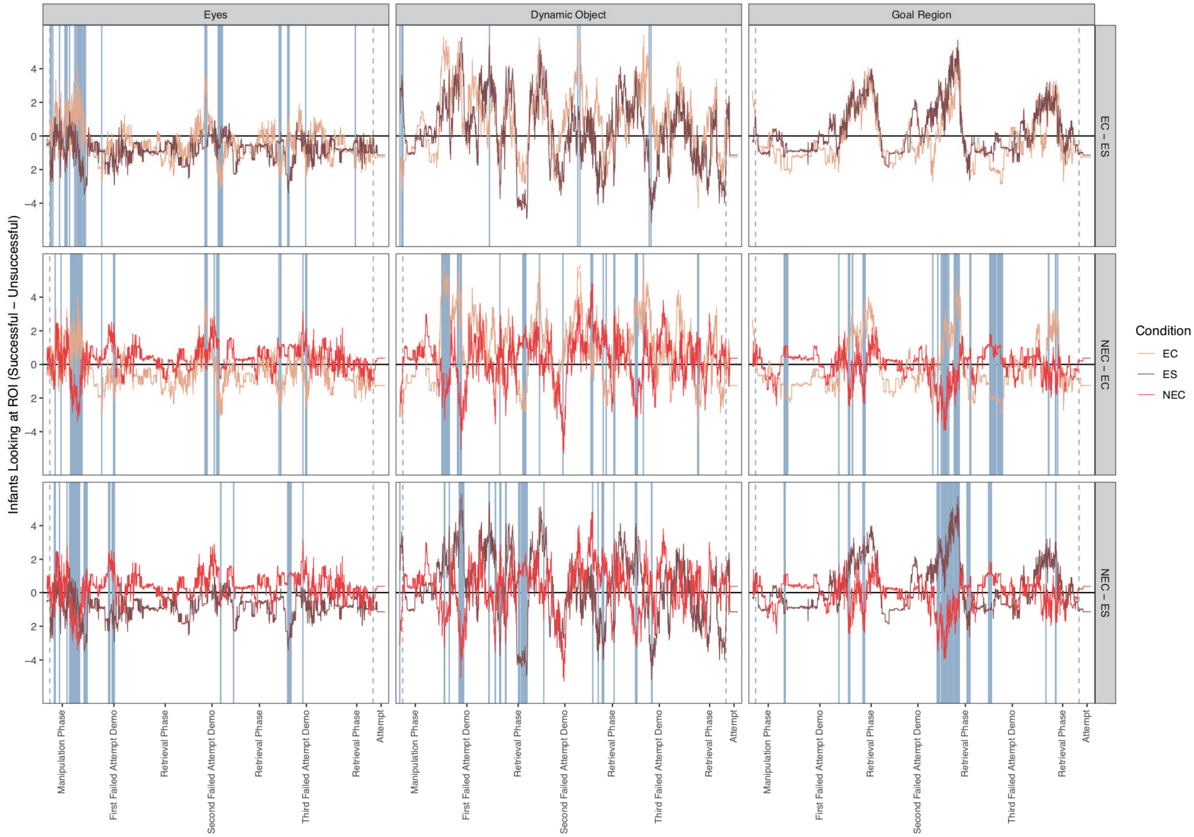

Figure 3.7. Presented are comparisons across conditions (EC and ES in the top panel, NEC and ES in the middle panel, and NEC and EC in the lower panel) of the time series of difference scores between successful and unsuccessful infants per region of interest (Eyes in the left panel, Dynamic Object in the middle panel and the Goal Region in the right panel). The y-axis represents proportions of Z scores. The vertical colored bands indicate significant differences across the conditions according to a permutation test.

3

comparisons across conditions of the time series of difference scores between successful and unsuccessful infants per region of interest and per condition. Significant differences are marked by the light blue vertical bands. Looking at the eye region in the EC and ES conditions prior to the failed attempt demonstrations is related to successfully find the target affordance after whereas this is not the case for infants in the NEC condition (note that the whole face of the woman was not visible for infants in the NEC condition during this phase). For the ROI related to the dynamic object, there are hardly any significant differences between the EC and ES conditions. However, for the ROI related to the dynamic object, both the EC and ES conditions differ from the NEC condition significantly during the timeframes in which the model looks at the dynamic object for the first time. Specifically, more successful infants in the EC and ES conditions tend to look at the dynamic object right after the model shifts her gaze towards it whereas this difference between successful and unsuccessful infants is absent across infants within the NEC condition. The reverse pattern can be seen during the phase in which the model retrieves the object to the initial location in order to prepare for the second demonstration: Not only are successful infants likely to look at the goal region during this phase (as noted before), but successful infants in the EC and ES conditions are likely to do so more than successful infants in the NEC condition.

Discussion

In the current study, we investigated whether attracting the attention of infants towards the eyes of a model before she performed a failed attempt to combine two objects would promote affordance realization by infants thereafter. Previous studies have shown that attracting the attention of infants towards the eyes of another promotes gaze following by following head turns (e.g., Senju & Csibra, 2008) and eye turns (de Bordes et al., 2013). We reasoned that the gaze of the model can inform the observing infants about the opportunities for action the model is trying to realize by manipulating the objects, and this likely enables the infants to learn which actions the objects might afford (i.e., affordance learning), even though the model fails at her attempts. Following the gaze of another person could facilitate learning about a specific affordance by means of directing the attention of infants towards relevant actions on (parts of) objects. As the actions performed by the model all belong to the action-repertoire of the observing infants, we predicted that infants had a bigger chance to realize the target affordance after their attention was drawn towards the model's eye region (socially or artificially) prior to the failed attempt demonstration than if their attention was directed towards the head region in a non-specific way.

Eye salience manipulations

We found that when the model's eyes were made salient prior to each demonstration, either as part of an ostensive signal (eye contact) or artificially (blinking and moving dots covering the eyes), the success rates of infants in realizing the target actions afforded by the objects were higher (65%) than when the entire head region of the model was covered (46%). We did not find a difference in affordance realization between the eye contact and eye salience conditions, which replicates earlier findings indicating that drawing the attention of infants to the eye region prompts infants to follow gaze, irrespective of how their attention was drawn towards the eyes (i.e., ostensively or non-ostensively; see de Bordes et al., 2013; also see Gredebäck et al., 2018; Szufnarowska et al., 2014 for similar results). The suggestion that gaze following plays a facilitatory role might also explain why Belagamba and Tomasello (1999) found that most 18-month-old infants but not 12-month-old infants were able to carry out the correct actions after seeing a failed attempt demonstration. Proficiency in following eye gaze seems to reach a functional level only from 18 months of age (Corkum & Moore, 1995), allowing the 18-month-olds but not the 12-month-olds to benefit from observing gaze behavior in their study.

These results do not imply that making the eyes salient in the EC and ES conditions is either a necessary or sufficient condition for infants to successfully realize the target affordance. The average success rate of the infants in the NEC condition was still 46% (also see Shneidman et al., 2014 for comparable results with 18-month-olds). Also, in the study by Huang and Charman (2005; Experiment 3), 56% of the infants were still able to successfully re-enact the failed attempt when it was executed without a model being visible, indicating that the object movements themselves, naturally draw the attention of infants towards crucial parts of the object and actions onto the objects that can lead infants to realize the target affordance.

Huang and Charman (2005) used the concept of stimulus enhancement in order to explain why infants in their study were able to realize the target affordance after seeing the failed attempt demonstration or the 'ghost' failed attempt demonstration. Stimulus enhancement, such as object movements, directs the attention of the observer to an object, increasing the chance that observers manually explore that object and discover its affordances thereafter (see Fritz et al., 2000; Galef, 2013; Heyes et al., 2000; Hopper, 2010; Hoppitt & Laland, 2013). However, it might be that when object movements (and gaze alike) directs the attention of the observer towards specific actions performed onto that object, affordance learning can occur based mainly on these observations alone instead of manual exploration after viewing the demonstration. In this sense, object movements and the gaze of the model are not merely drawing the attention towards the objects that increases the chance that infants interact with the objects. Rather, the object movements and the gaze of the model might guide (or educate) the perceptual system of infants by which they attune to critical aspects of the failed attempt through which they can learn a certain affordance. In the current study, this point is illustrated by the fact that infants in the EC and ES conditions needed less exploration time compared to infants in the NEC condition in producing the target act.

Spatiotemporal structure of eye movements

The analysis of the spatiotemporal structure of eye movements of infants allowed for a closer look at how the model's gaze directed and guided the attention of infants, leading them to attune to critical aspects of the failed attempt demonstrations through which they could learn affordances. Although most infants followed the object movements performed by the model, infants that were successful at realizing the target affordances in the Eye Contact and Eye Salience conditions followed the object movements that matched the gaze direction of the model during the failed attempt demonstrations in comparison to unsuccessful infants in these conditions. Specifically, successful infants in the EC and ES conditions looked at the object movements more often while the model conveyed one object towards the other before she 'failed' to combine the two than unsuccessful infants in these conditions. In addition, the successful infants in the EC and ES conditions also looked significantly more at the goal region just before and after she 'failed' in comparison to unsuccessful infants in these conditions and successful infants in the NEC condition. In our view, the looking patterns of the successful infants in the EC and ES conditions closely resemble, and thus specify, the actions necessary to relate the dynamic object with the stationary object by which the infants were able to realize the target affordance shortly after viewing the demonstration. In contrast, unsuccessful infants across conditions looked at the object movement in a less-specific manner. For instance, they looked more often to the object movements after the model failed to realize the target affordance and retrieved the objects towards the initial location, which is deemed as less informative for realizing the target affordance. So, although both groups looked at the same regions of interest for an equal amount of time, the difference in the temporal order in which infants looked at the object movements as a function of condition seems to be crucial for producing a success or not.

In order to directly assess the effect of condition, we compared the sizes of difference scores between infants that were successful and infants that were unsuccessful in affordance realization across conditions of looking at the eyes, dynamic object and goal region across the time series. When looking at the differences between successful and unsuccessful infants across conditions in general, it is noticeable that the looking pattern differs between these groups in the EC and ES conditions but not in the NEC condition. Specifically, we found significantly larger difference scores in the ES and EC conditions of looking at the eye region of the model prior the failed attempt demonstration and looking at the dynamic object right after the model shifts her gaze towards it for the first time in comparison to the NEC condition. This means that in EC and ES conditions, success was more strongly related to looking at the eye region of the model prior the failed attempt demonstration and looking at the dynamic object right after the model shifts her gaze towards it for the first time than the success of infants within the NEC condition. In the NEC condition, the eyes were made invisible at the start of the demonstration so initial gaze following and subsequent matching of gaze could not have led to subsequent differences in looking behavior between successful and unsuccessful infants. This result combined with the previously noted result that more infants in the EC and ES conditions were successful than infants in the NEC condition suggests that the manipulation, making the eyes of the model salient or not at the start of the failed attempt demonstrations, worked as intended: whenever the eye region was made salient, infants had higher success rates in finding the target affordance and of this successful group, more of them looked at the eye region and matched the subsequent eye gaze directions of the model in comparison to successful infants in the NEC condition.

Limitations and future directions

Huang and Charman (2005) have shown that infants produce the target acts, independent of whether they viewed the demonstrations live or not. Therefore, we chose to present the failed attempt demonstrations on a video screen with an inbuilt eye-tracking system instead of live-failed attempt demonstrations in order to avoid possible and unintended differences in the demonstrations across infants. In addition, this allowed us to measure the eye movements of infants and analyze differences between them based on subtle variations (i.e., the conditions contrasts). However, this study setup refrained us to measure and analyze the dynamic interaction and coordination of gaze between the infant and the model that occurs naturally (see Yu & Smith, 2013, 2017 for examples). This could further elucidate how infants and their caregiver adapt and attune their behavior dynamically in ways that allow infants to learn effectively.

Next, we choose to analyze the behavior of infants on a group level, assuming that infants across conditions react in similar ways. However, there are probably multiple ways in which infants can learn affordances based on demonstrations as used in the current study. Whereas some infants benefit a lot from using the gaze of a caretakers in learning affordances, others might depend more on the object manipulations that lead them to manually explore the objects from which affordances can be learned or perhaps a combination of two. This multicausality in learning and development is easily overlooked when looking for main effects of study manipulations but should be a topic of future investigation in order to better understand the complex interplay of contextual and personal factors that causes the observed variability of learning success across infants (see Thelen et al., 2001; Yu & Smith, 2017).

Finally, it should be noted that the retrieval phase is usually ignored in paradigms like the Re-enactment procedure (e.g., Bellagamba & Tomasello, 1999; Huang & Charman, 2005; Meltzoff, 1995;). However, the analysis in the current study reveals that it actually might be a highly relevant part of the demonstration phase as looking at the dynamic object being moved away from the other object might specify anything but combining the two objects together in a specific way, possibly refraining infants to learn the target affordance. Therefore, in future studies, attention should be devoted to what infants can observe before, between and after the experimental of learning phase.

The idea that social learning requires a form of social bias of the infant (e.g., intention detection) has thus far been tackled in studies showing that infants can realize object affordances with objects at a similar rate of success after seeing either a social or a non-social demonstration with those objects (e.g., Horne et al., 2009; Huang et al., 2002; Thompson & Russell, 2004; also see Heyes (2012) for a similar view on animal learning). As an alternative to these social biases and following theories of direct perception and ecological psychology (Gibson, 1979; Gibson & Pick, 2000; Meagher & Marsh, 2014; Thelen et al., 2001), we argue that social information like gaze direction is picked up and used similarly as other information in our physical environment, that is, as specifying affordances by attuning the perceptual system (Gibson, 1979; Gibson & Pick, 2000; Haith, 1998; Hellendoorn, 2014; Ingold, 2001).

Acknowledgments

We would like to thank Vivian Pasman for helping with the data-collection, professor Paul van Geert for supporting this study, and Dr. Ludger van Dijk and the anonymous reviewers on providing helpful suggestions on earlier drafts of this paper.

Chapter 4

Children's perception of facial expressions

P. F. de Bordes
F. Hasselman
R. F. A. Cox

Based on:
De Bordes, P. F., Hasselman, F., & Cox, R. F. A (in press). Children's perception of facial expressions. *Developmental Psychology*.

Abstract

This study investigated the developing ability of children to identify emotional facial expressions in terms of the contexts in which they generally occur. We presented Dutch 6- to 9-year- old primary school children (*N* = 164, 98 girls) prototypical contexts for different emotion categories and asked them whether different kinds of facial expressions belonged to those contexts or not, using a two-alternative forced-choice task. Correct and incorrect responses were quantified into a single index using Signal Detection Theory, representing children's sensitivity to perceive each facial expression as categorically different from each of the others in terms of their prototypical contexts. Results show age-related improvements in identifying facial expressions as belonging to their prototypical contexts. In addition, we found that older children not only made less misidentifications, but also misidentified less kinds of facial expressions to the prototypical contexts. Furthermore, the kinds of misidentifications children made suggests that they do not identify facial expressions based on their conceptual emotional valence. Results were discussed from a perceptual learning account.

Keywords: Facial expressions, Signal Detection Theory, Shannon Diversity Index, Social Development

Introduction

Humans are affected by facial expressions of others and tend to adjust their behaviour accordingly (Gao & Maurer, 2009; Leppänen & Nelson, 2009; Zebrowitz, 2011). The ability to perceive distinct facial expressions is crucial for social adaption and social exchange (Bal et al., 2010; Gao & Maurer, 2010; Mostow, Izard, Fine, & Trentacosta, 2002). This ability starts to develop from infancy (Cong et al., 2019; Lee, Cheal, & Rutherford, 2015) and continues throughout childhood and adolescence (Herba & Phillips, 2004; Montirosso, Peverelli, Frigerio, Crespi, & Borgatti, 2010; van Beek & Dubas, 2008; Widen & Russell, 2003). Although this development has been well documented, research outcomes and suggested developmental pathways seem to vary according to how the ability is operationalized and theoretically conceived (Calvo & Marrero, 2009; Durand, Gallay, Seigneuric, Robichon, & Baudouin, 2007; Rodger, Lao, & Caldara, 2018; Rodger, Vizioli, Ouyang, & Caldara, 2015; Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000; Widen & Russell, 2008;). In the current paper, we will argue that the perception of facial expressions is contextualized. In addition, we present a novel method of analysing the development of children to identify facial expressions as belonging to the prototypical contexts in which they are normally perceived.

Measuring categorical perception of facial expressions

When it comes to investigating the development of children in the ability perceive facial expressions, a wide range of research methods have been used such as labelling tasks, discrimination tasks and recognition tasks (e.g., Herba et al., 2006; Rodger, Lao, & Caldara, 2018; Thomas et al., 2007; Vicari et al., 2000; Widen, 2013). Most of these methods aim to measure the developing ability of children to perceive each facial expression in isolation (e.g., Durand et al., 2007; Rodger et al., 2015), in reference to only neutral expressions (e.g., Gao & Maurer, 2009, 2010) or in reference to just a few other categories of expressions (e.g., Thomas et al., 2007). This might offer a limited view on their ability to perceive distinct facial expressions for two reasons. First, the number and kinds of categories of facial expressions available during testing likely affects how children categorize the perceived facial expressions (Bimler & Kirkland, 2001; Russell & Fehr, 1987). Second and more importantly, each of the discrete categories of facial expressions are unlikely to be learned in isolation. This is because being able to perceive a particular facial expression as a discrete category entails that it is perceived as different from other categories of facial expressions. For instance, being able to perceive facial expressions of anger entails the ability to perceive a happy facial expression as 'not angry'. To overcome these limitations, it seems fruitful to use a forced-choice task in which children need to identify different facial expressions as belonging to an emotion category and to analyse both the correct and the incorrect responses. This can reveal to what extent children perceive facial expressions as categorically discrete and mutually exclusive. Based on such an analysis, Widen and Russell (2003, 2008) found that although five-year-old children are not yet proficient in verbally labelling expressions of anger accurately, they rarely mislabelled expressions of anger as happy or sad. This shows that although young children might not have the ability to label angry expressions properly, they are able to exclude angry expressions from other discrete categories, which is at least equally important in terms of attaching meaning to the perceived angry expression. In addition, they found that whereas four-year-old children used the label 'happy' to include happy, surprised and fearful expressions, five-to-six-year olds would be less inclined to mislabel surprised and fearful expressions as 'happy' and were able to label those expressions appropriately (Widen & Russell, 2003, 2008). In general, Widen and Russell (2003, 2008) have found that the categories of facial expressions emerge gradually, at least in children's vocabulary, from one or a few broad categories around the age of two years to several ones that can cover each of the universal expressions discretely around the age of six-to-seven years (i.e., broad-to-*differentiated pattern*, see Widen (2013) for an overview).

However, the development in label use might not be an optimal indication of children's developing perceptual sensitivity for the distinctions between facial expressions as knowledge of the labels does not mean that children know how to apply them correctly (Vicari et al., 2000; Widen, 2013). In addition, even though children from six to seven years seem to be able to label all of the facial expressions (Widen & Russell, 2008), studies have shown considerable improvement of children after the age of six years in their ability to label, discriminate, categorize and recognize them (e.g., Herba, Landau, Russell, Ecker, & Phillips, 2006; Rodger, Lao, & Caldara, 2018; Thomas, De Bellis, Graham, & LaBar, 2007; Vicari et al., 2000; Widen, 2013). Therefore, a study design that does not require verbal responses and measures both correct and incorrect identifications of facial expressions as belonging to emotion categories might be more suitable for investigating the differentiation process by which children become increasingly sensitive for the perceptual distinctions between facial expressions. This could be established by making a certain emotion category salient to children after which they have to identify facial expressions as belonging to that category or not. If children learn to perceive each of the facial expressions as perceptually discrete and mutually exclusive from each of the others, this should be reflected in an increase in correctly perceiving each of the universal expressions as belonging to a discrete emotion category and a decrease of incorrectly perceiving expressions as belonging to that same discrete emotion category. As studies have shown that children are better in non-verbal emotion recognition tasks than verbal labelling tasks (Klinnert, Campos, Sorce, Emde,

& Svejda, 1983; Vicari et al., 2000; Widen & Russell, 2008) and seven-year-olds children are able to label each of the universal expressions above chance level (Durand et al., 2007; Widen, 2013), it is likely that children from six years of age are to be able to identify them non-verbally above chance level as well.

Context dependent categorical perception of facial expressions

Ekman and Friesen (1971) proposed that there are six categories of facial expressions (i.e., 'anger', 'disgust', 'fear', 'happy', 'sad', and 'surprise') that are perceived in a similar way by humans across different cultures (also see Darwin, 1872; Ekman, 2003). This seems to be the reason why the abundance of the aforementioned studies on how children develop in perceiving facial expressions include a selection of these categories. However, recent studies have shown that the sociocultural upbringing of humans largely affects how humans view these categories of facial expressions, suggesting that these six categories are not perceived similarly across different cultures (e.g., Jack, Garrod, Yu, Caldara, & Schyns, 2012). Socio-cultural differences in categorical perception is a phenomenon that is not restricted to facial expressions as it is also evidenced in the perception of categories of colour and phonemes (e.g., Barrett, 2006; Goldstone & Hendrickson, 2010; Özgen & Davies, 2002; Ozturk, Shayan, Liszkowski, & Majid, 2013). These cultural differences likely have a social constructivist origin, meaning that the perception of categories of facial expressions and their meaning emerges from context-dependent functional relations between the perceived expressions and the possible actions that are adaptive in response to them (cf. (social) affordances, Gibson, 1979; Walker-Andrews, 1997; Widen, 2013; Withagen & Michaels, 2007, also see Träuble & Pauen, 2007).

With the term 'context', we refer to perceptual information that, together with the perceived facial expression, can be meaningfully connected to potential actions of the observer. Social referencing provides an illustrative example of the ability of (preverbal) infants to use another person's emotional expression to guide their own behaviour in various contexts (Klinnert et al., 1983; Ruba & Repacholi, 2019; Walker-Andrews, 1997). Context includes learned social practices and norms in relation to the perceived facial expression, the specific (social) environment in which the facial expression is perceived, but also the emotional state and personal history of the observer. The idea that specific experiences shape the way facial expressions are perceived is in line with the finding that children who have experienced an abnormal amount of hostility and anger in their life tend to respond differently to angry expressions in respect to other expressions, and in comparison with their peers (Ardizzi et al., 2017; Pollak, Messner, Kistler, & Cohn, 2009; Pollak & Tolley-Schell, 2003). Hence, perception of facial expressions is shaped by specific exposure and experience with facial expressions over time and across specific (cultural) contexts (see Goldstone, 1998; Goldstone & Hendrickson, 2010; Kelly et al., 2011; Leppänen & Nelson, 2009; Pollak & Kistler, 2002; Pollak et al., 2009).

The extent to which a context can shape the facial expressions we perceive is also evidenced by the fact that humans are likely to perceive expressions (e.g., happiness) that are congruent with the context (e.g., a birthday party) in which they are perceived (Carroll & Russell, 1996; also see Ruba, Meltzoff, & Repacholi, 2019). As the meaning (cf. affordances) of facial expressions is intrinsically context-dependent (Barrett, 2006; Carroll & Russell, 1996; also see Marchi & Newen, 2015), it might explain why children tend to perform much better in facial expression recognition tasks and labelling tasks when the task is preceded with a context in terms of a story, behavioural consequences, or pictures depicting a scene that fits to the perceived facial expressions (e.g. Vicari et al., 2000; Widen & Russell, 2004, 2013). For these reasons, it seemed appropriate to investigate the developing ability of children to identify facial expressions according to the prototypical context in which they are normally conceived in their culture. In doing so, we expected that young and relatively unexperienced children are likely to misidentify all kinds of facial expressions as belonging to the prototypical contexts that each display specific emotion categories. However, we expected that across age and with experience, children would display a decrease of diversity of expressions that are incorrectly identified as belonging to the prototypical contexts, similar to the broad-to-differentiated pattern (Widen & Russell, 2003, 2008) discussed earlier. Hence, the aim of the current study was to track age-related differences in the manner different facial expressions are perceived in different prototypical contexts (and not the effect of prototypical context itself).

Flexible boundaries between categories of facial expressions

Widen and Russell (2003, 2008) found that incorrect labelling by children up to seven years of age seems to be systematic rather than random and that children are likely to use the same label to describe different expressions with a similar conceptual valence (either pleasant or unpleasant) or level of arousal. However, this result might also be partly due to the fact that children are biased to use some labels more than others in describing the different facial expressions, favouring labels that are more commonly used than others (Widen, 2013). Moreover, when a specific facial expression such as fear is incorrectly identified as belonging to a prototypical context in which anger is normally expressed, this does not necessarily have to imply that expressions of anger are also incorrectly identified as belonging to a prototypical context in which fear is normally expressed. For example, given a certain context such as a bullying situation at school, expressions of disgust might be perceived as expressions of anger by children as both expressions could be seen as suitable to the situation (Danovitch &

Bloom, 2009). However, that does not imply that they would perceive expressions of disgust as expressions of anger in another context such as a display of infested food where expressions of anger are generally less suitable. Therefore, when facial expressions need to be categorized as belonging to context that makes a particular emotion category salient, we do not expect that incorrect categorizations are always necessarily bidirectional. Such a result would stress that the differences between the facial expressions that represent different emotional states do not lie on internalized orthogonal continua such as pleasant-unpleasant or high arousal-low arousal, on which strict category boundaries can be identified (see Bimler & Kirkland, 2001; Young et al., 1997; Widen & Russell, 2008). Rather, the appearance of category boundaries might be much more flexible and contextualized, as has been shown to be the case for the speech sound categories (Case, Tuller, Ding, & Kelso, 1995; Hasselman, 2015).

The current study

In the current study, a two-alternative forced-choice task was used in which six- to nine-year-old children growing up in a West-European culture (i.e., the Netherlands) were instructed to identify seven different facial expressions ('anger', 'disgust', 'fear', 'happy', 'sad', 'surprise', and 'neutral') presented in succession as belonging to a target emotion category or not in six different conditions. Each condition was preceded with a presentation of a prototypical context which was either 'anger', 'disgust', 'fear', 'happy', 'sad', or 'surprise'. To operationalize their ability to identify facial expressions into emotion categories displayed by prototypical contexts (called 'perceptual sensitivity' henceforth), the results were analysed using Signal Detection Theory (SDT). SDT takes both correct responses (Hit ratio) and incorrect responses (False Alarm ratio) into account in constructing a single index of perceptual sensitivity for all categories together and for each category separately (see Stanislaw & Todorov, 1999; Thomas et al., 2007). There are several advantages of applying SDT to the measures of the proposed forced-choice task. First, it capitalizes on both correct and incorrect responses statistically, similar to the qualitative analysis on correct and incorrect labelling of facial expressions by children as Widen and Russell did (see Widen, 2013 for an overview). Second, it teases apart the response bias respondents can potentially have from the perceptual sensitivity (Durand et al., 2007; Lynn & Barrett, 2014). Response bias refers to the potential general tendency to identify facial expressions as either belonging or not belonging to the prototypical context, independent from the type of prototypical context and/or the facial expressions shown within a condition. This seems especially relevant as Widen and Russell (2003, 2008) have shown that children are biased in using some labels more frequently than others in describing facial expressions, which could mean that children in the current study might be more inclined to identify facial expressions as belonging to a certain category than others. Third, this operationalization of perceptual sensitivity has considerable conceptual overlap with the notion of 'differentiation' as put forward by Gibson and Gibson (1955), as the sensitivity for perceiving distinctions leading to differential responses plays a central role in both. Finally, Shannon's Diversity Index (see Begon, Harper, & Townsend, 1996; de Jonge-Hoekstra, van der Steen, & Cox, 2020) was used to index the diversity of incorrect responses (i.e., kinds of false alarms) across age groups.

Methods

Participants

The final sample consisted of 164 primary school children (98 girls) between ages five and nine (M_{age} = 7 years; 7 months, *SD* = 1 year and 2 months) from four different school grades from five different schools in the Netherlands (see Table 4.1 for an overview).

Table 4.1. Sample characteristics in terms of age in years (mean and standard deviation), distribution of children across gender, and distribution of children across school grades and their accompanying age groups

					Age					
		N			Boys		Girls		Total	
Age group	School grade	Boys	Girls	Total	*M*	*SD*	*M*	*SD*	*M*	*SD*
Six-year-olds	2	16	28	45	6.30	0.50	6.14	0.32	6.20	0.40
Seven-year-olds	3	16	17	33	6.94	0.32	6.91	0.36	6.92	0.33
Eight-year-olds	4	13	24	37	8.11	0.38	8.10	0.35	8.11	0.36
Nine-year-olds	5	19	29	49	8.86	0.29	9.01	0.35	8.95	0.33

Three participants (two that were in the second grade and one that was in the fifth grade) were excluded from the analysis because of computer failure. The gender of two participants (one in second grade and one in the fifth grade) was registered incorrectly and is therefore missing in the analysis. In constructing age groups, we aimed to minimize intragroup variation and maximize intergroup variation in terms of the (social) contexts in which they normally view facial expressions. We reasoned that children within a school grade are more likely to be similar in the types of contexts in which they perceive facial expressions and display more commonalities in their ability to interact within these contexts than children of the same chronological age per se. Therefore, the age-groups of children were comprised of children attending the same school grade and hence have some overlap, although very little, in terms of chronological age. Prior to participation, informed consent was acquired from one parent of each child. This study was part of a research program titled "the perceptual basis for emotion recognition and perspective taking" and was approved by the ethical committee of the Faculty of Social Sciences, Utrecht University, under protocol number FETC18-037.

Materials

The Facial Expression Perception task (FEP task) consisted of 72 trials. In each trial, a photograph was presented with a face displaying one of the six facial expressions (i.e., anger, disgust, fear, happy, sad, or surprise) or the neutral expression. The person displaying the facial expression could be either male or female, and was either a child, middle-aged adult, or an older adult. The photographs were shown in grayscale with the background and hair around the face removed to reduce distraction and to ensure the saliency of the facial expressions[1]. Trials were divided over six conditions. Each condition contained six trials displaying one of the expressions that fit the emotion category of that condition and six trials displaying the remaining expressions and the neutral expression as distractor emotions. Within each condition, both the target emotion and distractor emotions were each expressed three times by a male face and three times by a female face. For each gender, two were children, two were middle-aged people, and two were older people. Trials were shown on a 15" display with a resolution of 1280 × 800 pixels and E-prime 2 was used for stimuli presentation and data collection (Schneider, Eschman, & Zuccolotto, 2012). Participants responded by using the computer keys 'L' (marked blue) for whenever they perceived the target emotion and 'D' (marked red) for whenever they did not perceive the target emotion.

4

Procedure

Warming up and training

The task took place in a quiet room in the school the children were attending. After being seated in the room, the experimenter would typically chat with the child and asked whether

[1] Pictures were retrieved from the internet using Google Search, using the search words "expression" in combination with each of the tested expressions (e.g., "sad expression") and filtered on "free to use, share of modify". A pilot study was performed to test the validity of the stimulus material. Students of the faculty of Social Sciences of Utrecht University (*N* = 27, twenty were female, M_{age} = 20 years and 4 months, *SD* = 1 year and 6 months) participated on voluntary basis with informed consent. During the end of a lecture, students completed an online expression labelling task via their smartphone in silence. 72 pictures of expressions were presented in succession and in random order. Participants labelled each expression by pressing one out of seven options: the six used expressions and the neutral expression. Nine expressions were labeled only 50% or less correct and were therefore replaced. two independent raters labelled all replacements correctly. With the remaining pictures, students labelled 83.86% correct on average which is typical performance level for adults (cf. Lawrence, Campbell, & Skuse, 2015), and, in our view, indicated that the material is valid for experimental use.

they would like to play a small game about emotions. Next, the experimenter explained that she would sit next to the child during the task and that the child could always take a break, stop the game or ask questions if needed, and finally, that there are no wrong answers.

Prior to the task, children did a short training session to get accustomed to pressing the computer keys and specifically to press the blue key to confirm the presence of a target and the red key otherwise. To this end, they were instructed to push the blue key whenever a blue screen was presented and otherwise press red. On four different trials, participants were shown successively a blue screen and a red screen twice and could only proceed to the next trial if they pressed the correct button. Although a few children made an error on the first training trial, all children seemed to understand the task after and responded correctly on the last three trials. After the training session, the experimenter asked if the child had any further questions or required a bathroom break, and if not, she would proceed to the first condition of the FEP task.

Testing phase

Each condition started with a presentation of a prototypical context for a particular emotion whereafter children engaged in a forced-choice procedure in which they identified facial expressions as either belonging to that prototypical context or not. At the start of each condition of the FEP task, emotions were briefly discussed according to a script by describing situations with visual aids in which each of the six emotions would be appropriate. With each emotion, an example was given of a situation, typically leading to experiencing that particular emotion. This created the necessary context for the subsequent task in which children had to identify facial expressions as belonging to that context or not.

For happiness, pictures of an ice-cream, a playground and jumping people on the beach were shown and the child was told that most people are happy when they get to eat ice-cream.

For angriness expressions, a picture of two children was shown in which one child had a considerable bigger piece of cake than the other, a picture depicting a mother disciplining her child, a thief running away with a large sum of cash and an almost intact ice-cream in upside down position on the street. The child was told that a child could get angry if someone else would intentionally break his or her new toy.

For sadness, pictures were shown of a child in front of a window and looking at the rainy and cold weather outside, a broken smartphone, a funeral scenario and devastating looking football players after losing a match. The child was told that people tend to get very sad when their pet passes away.

For surprise, pictures were shown of someone wearing a blindfold and about to receive a cake, a nicely packed present, someone waiting around a corner to surprise another person that is approaching that corner, and a car that is parked on a street that is curved upside down. The child was told that people would be surprised if it would snow in the middle of the summer.

For fear, pictures of a thunderstorm, a barking dog showing its teeth, a shady (creepy) alley and a person wearing a cape with everything black except for two white dots for eyes were shown. The child was told that people tend to get scared in the dark or whenever they see monsters.

For disgust, pictures of dog poo on the street, a dead and half eaten insect, a garbage can with the stench visualized as smoke/fog and a plate of something that could once have been food were presented. The child was told that some people tend to get disgusted and say 'yucky' ('bah' in Dutch) whenever they have to eat something they don't like to eat, such as 'brussels sprouts' (for Dutch children, Brussels sprouts are notorious for its bitter taste)[2].

After each emotion was briefly described, the experimenter asked children to describe a situation that would result in experiencing the discussed emotion. Answers were mildly praised. Before the FEP task started, the experimenter explained that a series of faces would be shown in succession, each expressing an emotion, and that they had to press the blue button whenever the face was expressing the same emotion as they just discussed, and otherwise the red button. After this instruction in the first condition, the experimenter would ask which button had to be pressed upon seeing the target emotion. Children would typically respond by saying 'the blue button' and if not (which happened only in a few cases), the experimenter would repeat the instruction where after they typically gave the correct answer, revealing that they understood the task instructions.

During the presentation of the trials in each condition, children were wearing headphones in order to reduce possible distracting noises from the surrounding area and to reduce the tendency to talk during the task. The order of presentation of the conditions and pictures within the conditions was randomized. After children had completed the six conditions, they were asked if they liked to play the game and were given a sticker to thank them for participating, after which they returned to their classroom.

[2] Three independent raters categorized the prototypical contexts, consisting of the collection of pictures, in terms of one out of six emotion categories used in the experiment. All raters categorized the prototypical contexts as belonging to their emotion categories as intended (100% agreement), revealing that the found effect was unlikely to be due to ambiguity in terms of how the prototypical contexts were constructed.

Measures and coding data

Perceptual sensitivity

On each trial, reaction time (in milliseconds) and performance (correct or incorrect) were registered. Performance of participants on each trial was recoded into either a '*hit*' whenever participants correctly identified the perceived facial expression as belonging to the emotion category, '*miss*' whenever they failed to identify the perceived facial expression as belonging to the emotion category, '*false alarm*' when they incorrectly identified the perceiving facial expression as belonging to the emotion category, or '*correct rejection*' when they correctly did not identify the perceived facial expression as belonging to the emotion category. Next, perceptual sensitivity (in SDT, known as '*d-prime*') was calculated for each participant for the whole task and per condition, using the following formula: $d' = z^{-1}(HIT) - z^{-1}(FA)$, where *FA* is the proportion of false alarms on the trials that contained a distractor, *HIT* is the proportion of hits on the trials that contained a target and *z* is the inverse of the normal distribution function (MacMillan & Creelman, 2004; see Stanislaw & Todorov (1999) for an extended overview). A simple way to interpret the perceptual sensitivity index is that a score above zero means that the proportion of hits is larger than the proportion of false alarms, meaning that the participant is able to distinguish the target emotion(s) from the distractors above chance level (MacMillan & Creelman, 2004). To avoid divisions by zero, the proportion of false alarms and proportion of hits were adjusted whenever they were either one or zero by replacing the zeros with $X_{adj} = \frac{1}{2N_{trials}}$, and the ones by $X_{adj} = 1 - \frac{1}{2N_{trials}}$ before *d'* was calculated, as suggested by MacMillan and Creelman (2004).

Criterion bias index

The criterion bias index signifies the extent to which participants are biased to either confirm or disconfirm perceiving the target, regardless of the expression shown, and was used as a measure of the reliability of the data (see Stanislaw & Todorov, 1999). The criterion bias index (*c*) was calculated for each participant. In order to calculate *c*, the following formula was used: $c = \frac{-(z^{-1}(HIT) + z^{-1}(FA))}{2}$. Negative scores indicate a tendency to confirm the presence of a target, positive scores indicate a tendency to disconfirm the presence of a target, and scores around zero indicate that the participant was not biased in any direction (Stanislaw & Todorov, 1999). In order to calculate c whenever the proportion of false alarms and proportion of hits were either one or zero, these proportions were adjusted using the same formulas as presented in the section above.

Shannon's diversity index of false alarms

Shannon's Diversity Index (*H*) was used to index the diversity of distractor facial expressions that were incorrectly identified (i.e., false alarms) as belonging to the prototypical contexts of each target emotion for participants of each age group. This index can reveal to what extent the type of false alarms in each condition made by participants of an age group is either random or non-random. If it is random, it means that participants of an age group made many different kinds of false alarms in a condition. For instance, five-year-olds might mistakenly identify expressions of anger, fear, sadness and surprise as belonging to a prototypical context for disgust. Conversely, if it's non-random, it means that participants of an age group and for a target emotion made one or a few specific kinds of false alarms. For example, nine-year-olds might mistakenly identify an expression of anger as belonging to a prototypical context for disgust but do not make other mistakes in identifying expressions as belonging to the prototypical context of disgust. We calculated *H* for each age group and for each condition using the following formula: $H = -\sum_{i=1}^{s} p_i \, ln \, p_i$, where *s* is the number of different kinds of possible false alarms that were made in each condition (which can be between one and six as there are six possible kinds of false alarms), and p_i is the proportion of observed false alarms of a particular kind relative to the total number of false alarms across all kinds of false alarms that have been made by children of that age group on a particular condition (see Begon et al., 1996). As p_i is a proportional measure, *H* is invariant with respect to the absolute number of false alarms that a particular age group made in a particular condition. In other words, *H* reflects the distribution of false alarms independently from the absolute number of false alarms made by an age group or in a condition. When *H* is high, it means that the proportions of false alarms are uniformly and randomly distributed across the different distractor emotions, indicating that participants of that age group identified a high diversity of facial expressions as belonging to the context that described the target emotion. Conversely, a low *H* indicates that the proportions of false alarms are distributed unevenly across the different distractor emotions, indicating that participants of that age group identified a low diversity of facial expressions to the context that described the target emotion. The average and standard deviation of *H* across the six conditions were calculated for each of the four age groups. Finally, the calculated values of *H* were transformed into Shannon's equitability (E_H) to fit into a scale that runs from zero (no diversity and no randomness) to one (maximum diversity, complete randomness) in order to easily interpret the differences in values across age groups and emotion categories. This transformation was performed, using the following formula: $E_H = H/(\frac{H}{ln\,S})$.

Data preparation & analysis

Analysis of reaction time

Reaction times (in ms) were averaged across trials and across conditions for each participant. In order to test whether average reaction times differed across age group, a one-way ANOVA was performed using the average reaction time per participant. All assumptions for performing the ANOVA were checked. Two participants were removed from the analysis, as their average reaction time deviated 3 *SD* from the average reaction time in their age group. According to Levene's test, variances in reaction times between children from different age groups were not equal ($F(3, 158) = 7.89$, $p < .001$. Data transformations as suggested by Howell (2007) and Tabachnick and Fidell (2007) did not seem to change this outcome. Therefore, the alpha was set to .01 in order to reduce the chance on a type-1 error (Moder, 2010) and the non-parametric Games-Howell post-hoc test was used to analyse which age groups differed from each other in terms of average reaction times (Ruxton & Beauchamp, 2008). Effect sizes for the significant group differences in reaction time were calculated in terms of Cohen's *d*.

Analysis of SDT measures and Shannon's Diversity Index of false alarms

In order to test if the children in our sample were able to relate each of the facial expressions to its contextual description above chance level, we tested whether the perceptual sensitivity (*d'*) was above zero for each of the four age groups and for each of the six conditions, using 24 separate *t*-tests. In order to account for an inflated type 1 error, a Bonferroni correction was applied, meaning that $\alpha_{critical}$ of 0.05 was divided by 24, resulting in $\alpha_{altered}$ of .002. All assumptions for performing these tests were met. In order to test whether average perceptual sensitivity (*d'*) for the different expressions combined on the FEP task differed across participants from the four age groups, a one-way ANOVA with a Bonferroni post-hoc tests were performed. All assumptions for performing the ANOVA were checked and met and no outliers were removed as all obtained sensitivity scores did not deviate more than 3 *SD* from the average. Effect sizes for the significant group differences (as indicated by the Bonferroni comparison test) in average *d'* on the FEP task were calculated in terms of Cohen's *d*. To test whether children across different age groups differed on d' on each of the six emotion categories while taking the within subjects variance across the conditions into account, a MANOVA was performed with Bonferroni post-hoc comparison tests, using age group and gender as the independent variables and the *d'* for each emotion category as the dependent variables. Pillai's' Trace was used as an indication of whether age groups generally differed from each other whereas Bonferroni post-hoc comparison tests were used to indicate age group differences for each emotion category. Effect sizes for the significant group differences as indicated by the Bonferroni test in *d'* for each emotion category were calculated in terms of Cohen's *d*.

Next, η^2 of the interaction of emotion category and age group of all the motion categories were calculated in order to compare the sizes of age difference in *d'* between the emotion categories. The η^2 of the intercept of all the motion categories were calculated in order to compare the *d'* of participants for each emotion category independent from age group. Three univariate outliers were removed as the obtained *d'* deviated more than 3 *SD* from the average *d'* in that age group. In addition, one multivariate outlier was removed, based on a deviation from the expected the Mahalanobis distance ($p < .001$ χ^2-distrubution). All assumptions for performing the MANOVA were checked and met.

In order to test whether the criterion bias index *c* for the FEP task differed across participants from different age groups, a one-way ANOVA was performed using the *c* for the FEP task per participant. All assumptions for performing the ANOVA were checked and met. An additional MANOVA with post-hoc Bonferroni comparison tests test was performed in order to check whether children across age groups differed in *c* for each emotion category separately. Pillai's Trace was used as an indication of whether age groups generally differed from each other whereas Bonferroni post-hoc comparison tests were used to indicate age group differences for each emotion category. Effect sizes for the significant group differences as indicated by the Bonferroni comparison test in *d'* for each emotion category were calculated in terms of Cohen's *d*. Four univariate outliers were removed as the obtained *c* deviated more than 3 *SD* from the average c in that age group. All assumptions for performing the MANOVA were checked and met.

To test whether children from different age groups differed on *H*, a one-way ANOVA with Tukey's HSD post-hoc tests was used. Effect sizes for the significant group differences (as indicated by the Tukey's HSD tests) in *H* were calculated in terms of Cohen's *d*. All assumptions for performing these analyses were checked and met. Finally, a difference score in *H* was calculated between the nine-year-olds and six-year-olds in order to observe age trends in this measure.

Results

Reaction time

Average reaction time on the task seemed to differ across children from different age groups, $F(3, 158) = 18.33$, $p < .001$, $\eta^2 = .26$. Non-parametrical post-hoc Games-Howell tests indicated that six-year-olds ($M = 3438$ ms, $SD = 1197$ ms, $N = 45$) were significantly slower than 7-year-olds (M = 2552, SD = 612, N = 33, p < .001, Cohen's d = 1.04), 8-year-olds olds (M = 2344, SD = 653, N = 37, p < .001, Cohen's d = 1.13), and nine-year-olds olds (M = 2282, SD = 640, N = 49, p < .001, Cohen's d = 1.20). No other age group differences were found (all p > .05).

Perceptual sensitivity and response bias measures

The distribution of d' scores for each emotion category of all children differed significantly from zero (all t-tests, p < .001), indicating that on average, children had a higher Hit ratio than a False alarm ratio, meaning their performance on the FEP task was above chance level. Children differed in terms of overall perceptual sensitivity (d') across age on the FEP task, $F(3, 160) = 16.22$, p < .001, $\eta^2 = .23$. Post-hoc Bonferroni comparison tests revealed that in terms of perceptual sensitivity scores, six-year-olds (M = 1.44, SD = .68) scored lower than 8-year-olds (M = 2.13, SD = .65, p < .001, Cohen's d = 1.04) and 9-year-olds (M = 2.25, SD = .50, p < .001, Cohen's d = 1.36), and seven-year-olds (M = 1.77 , SD = .62) scored lower than 9-year-olds (p = .003, Cohen's d = .85). No other age group differences were found (all p > .05). These group differences suggest a positive relation between age and overall perceptual sensitivity (see Figure 4.1). A MANOVA yielded a similar result in showing that children across age generally differed in perceptual sensitivity for the individual emotion

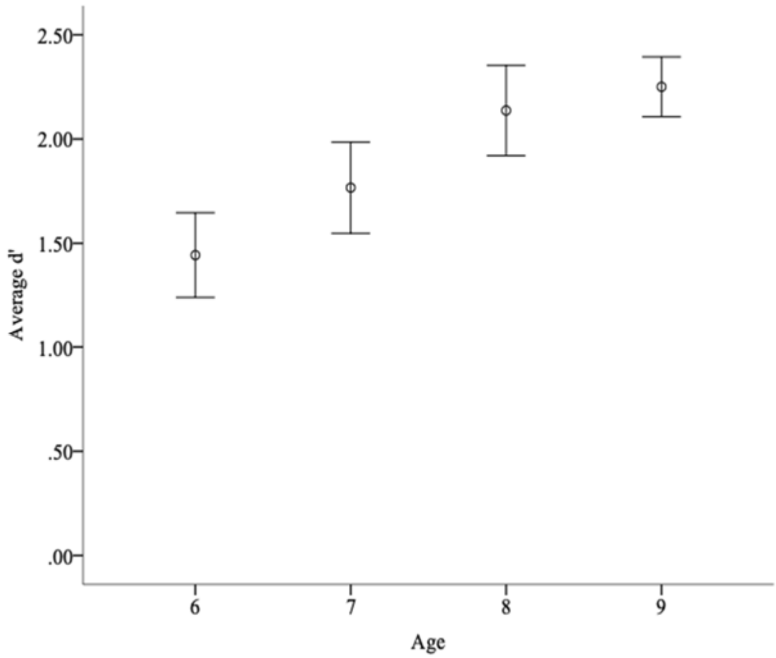

Figure 4.1. Average perceptual sensitivity (d') with standard error bars for the six emotion categories combined per age group.

categories on the FEP task, Pillai's' Trace = .28, $F(18, 459) = 2.59$, $p < .001$, $\eta^2 = .092$. Post-Hoc Bonferroni comparison tests results suggest a general positive relation between age and d' for each emotion category (see Figure 4.2). In addition, children across the tested age groups differed in their perceptual sensitivity for the emotion categories in the following order of size of difference (from large to small): 'surprise' ($\eta^2 = .11$), 'sad' ($\eta^2 = .09$), and 'disgust' ($\eta^2 = .09$), 'fear' ($\eta^2 = .08$), 'happy' ($\eta^2 = .08$), and 'angry' ($\eta^2 = .07$). Based on the partial η^2 of the intercepts of the individual emotion categories that control for age, perceptual sensitivity for the emotion categories had the following order (from high to low): 'happy' ($\eta^2 = .94$), 'angry' ($\eta^2 = .84$), 'sad' ($\eta^2 = .77$), 'surprise' ($\eta^2 = .76$), 'disgust' ($\eta^2 = .76$), 'fear' ($\eta^2 = .68$). Average response bias index did not differ across age, $F(3, 160) = 1.94$, $p = .126$ and the average response bias index for the entire test population was close to zero (M = -.004, SD = .33). However, a MANOVA revealed a difference in c between children across age for the individual emotion categories on the FEP task, Pillai's' Trace = .15, $F(18, 459) = 1.35$, $p = .050$, $\eta^2 = .050$. Post-hoc Bonferroni comparison tests revealed only one age group difference, namely that six-years-olds (M = .26, SD = .44) less often confirmed the presence of the emotion category than nine-year-olds (M = .04, SD = .28, p = .028, Cohen's d = .59).

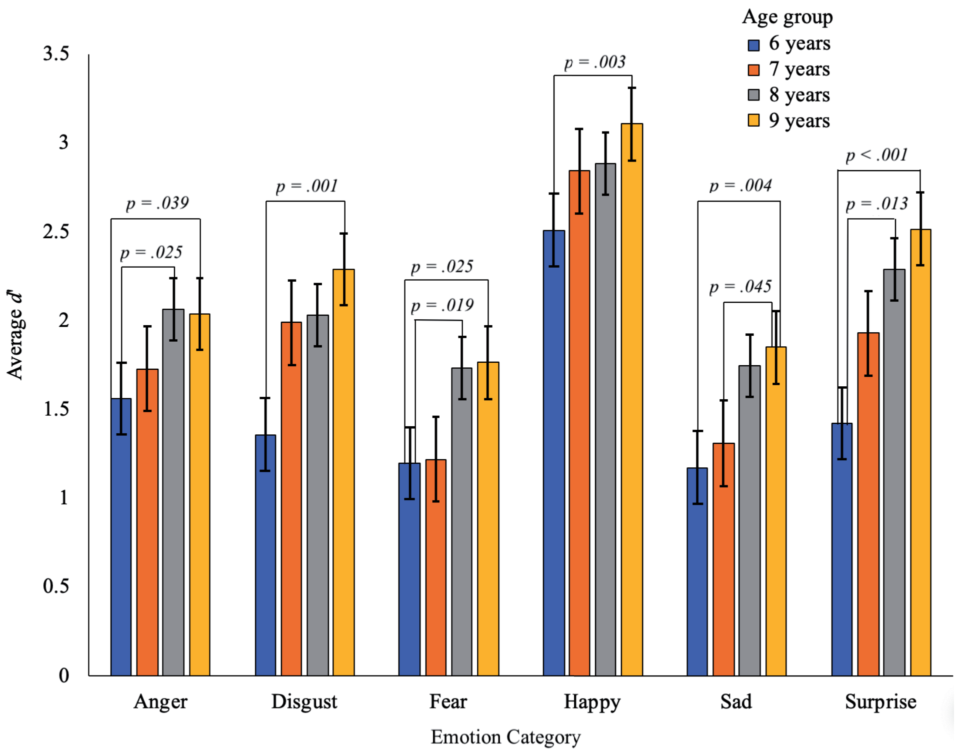

Figure 4.2. Average perceptual sensitivity (d') with standard error bars for each of the six emotion categories per age group.

4

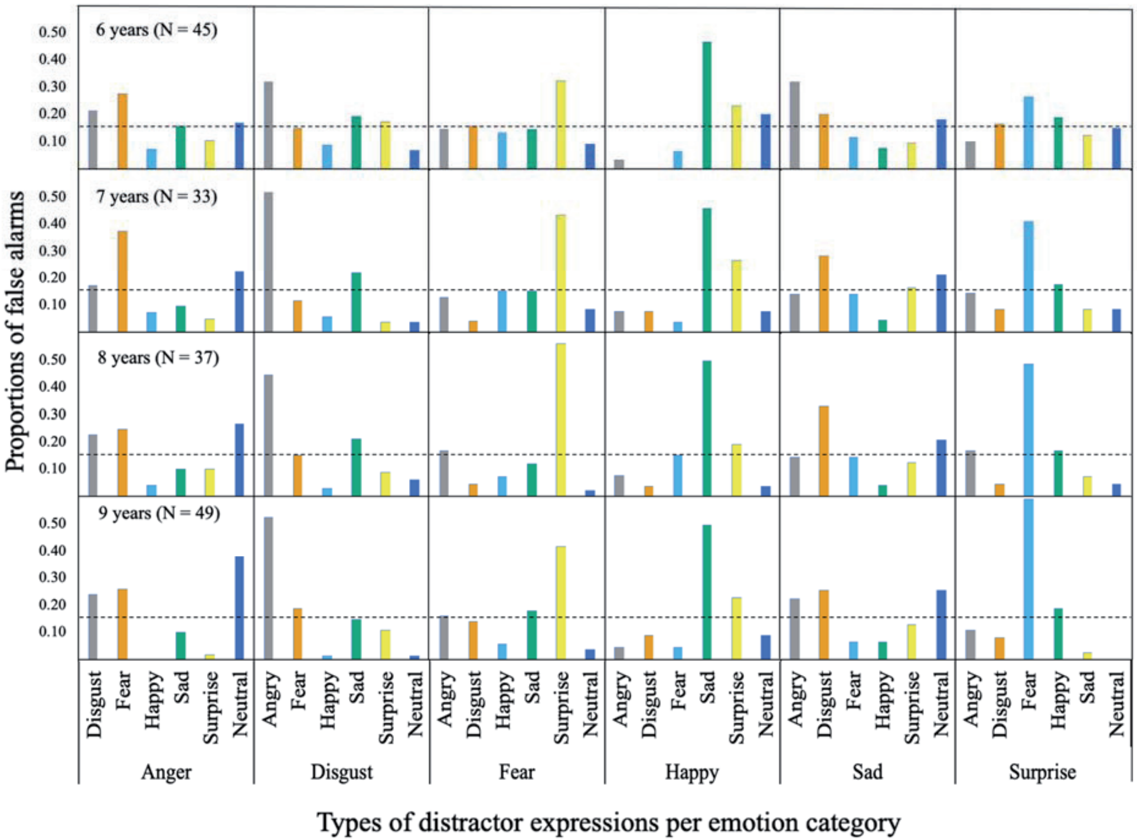

Figure 4.3. Proportions of false alarms distributed across distractor expressions per emotion category for each age group. The dotted vertical reference line (on 16.66) indicates the proportions of false alarms per emotion category that would be expected in a uniformal distribution. False alarm proportions deviating from this line indicate a low Shannon's Diversity Index for the emotion category to which the false alarm proportions belong. *Note*. The average number of false alarms per emotion category decreases as the age of children is increasing.

Shannon's Diversity Index of false alarm distractors

Children across age groups differed in average *H*, $F(3, 160) = 19.76$, $p < .0001$, $\eta^2 = .27$. Post-Hoc Tukey's HSD test results suggest a general negative relation between age and *H* as they indicated that six-year-olds ($M = 1.64$, $SD = .16$) scored higher than seven-year-olds ($M = 1.53$, $SD = .12$, $p = .01$, Cohen's $d = .78$), eight-year-olds olds ($M = 1.49$, $SD = .13$, $p < .0001$, Cohen's $d = 1.03$), and nine-year-olds olds ($M = 1.40$, $SD = .17$, $p < .0001$, Cohen's $d = 1.45$), seven-year-olds scored higher than nine-year-olds ($p = .001$, Cohen's $d = .88$) but not higher than eight-year-olds ($p = .69$), and eight-year-olds scored higher than nine-year-olds ($p = .04$, Cohen's $d = .59$). The standardized scores of *H* (E_H) are presented in Table 4.2, and shows that young children have larger values of E_H than older children for most expressions, meaning that the distribution of false alarms is more uniform for younger children than older children for these expressions. This is further illustrated in Figure 4.3 which shows that distributions of proportions of false alarms for most emotion categories seem to be rather uniformly distributed across the different distractor expressions for six-year-olds, compared to older children, especially compared to nine-year-olds. The largest decline in E_H across age was for the expression 'Surprise', followed by 'Fear' and 'Disgust', suggesting that children develop in distinguishing these particular expressions from others across the tested age range (see Table 4.2). The distribution of false alarms does not seem to change for the emotion categories 'Sad' and 'Happy' across ages, considering the small difference scores of E_H between six-year-olds and nine-year-olds (see Table 4.2).

4

Table 4.2. Shannon's Equitability Index (E_H) of the distribution of false alarm rates across distractor expressions for each emotion category per age group and difference scores of E_H between six-year-olds and nine-year-olds per emotion category

Age group	Anger (E_H)	Disgust (E_H)	Fear (E_H)	Happy (E_H)	Sad (E_H)	Surprise (E_H)
6	.95	.93	.95	.81	.94	.97
7	.88	.76	.86	.80	.94	.89
8	.92	.82	.73	.78	.92	.80
9	.85	.72	.86	.78	.92	.72
Difference score	.10	.21	.09	.03	.02	.25

Discussion

In the current study, we investigated the difference in ability across six- to nine-year-olds to identify emotional facial expressions as belonging to prototypical contexts for each emotion. We found that the perceptual sensitivity of children of all age groups was above chance level for all emotions and that older children are more perceptually sensitive to perceive each of the six expressions as distinctly belonging to specific emotion categories than younger children (cf. Herba et al., 2006; Rodger et al., 2018; Thomas et al., 2007). Specifically, with increasing age and experience, children become better in correctly identifying facial expressions as belonging to their prototypical contexts and made fewer mistakes in terms of incorrectly identifying facial expressions as belonging to the prototypical contexts. The same pattern was observed when looking at the age-group differences in terms of the perceptual sensitivity for each of the used expressions separately, although these differences were more pronounced for some emotion categories compared to others. Specifically, perceptual sensitivity for the facial expressions 'surprise' 'sad', and 'disgust' seemed to develop stronger across the tested age groups than the perceptual sensitivity for the facial expressions 'fear', 'happy', and 'angry'. A similar age-related pattern has been observed in other studies that have shown that children within this age range improve in identifying these facial expressions (cf. Durand et al., 2007; Gao & Maurer, 2010; Rodger et al., 2015). In addition, the perceptual sensitivity for identifying facial expressions as belonging to their prototypical contexts across all age groups had the following order (from high to low): 'happy', 'angry', and 'sad', 'surprise', 'disgust', 'fear' (cf. Durand et al., 2007; Herba et al., 2006; Rodger, Lao, & Caldara, 2018; Segal, Reyes, Gobin, & Moulson, 2019; Thomas et al., 2007; Vicari et al., 2000; Widen & Russell, 2013). The fact that the perceptual sensitivity for happy facial expressions was higher than for any other expression is in line with studies that have shown that humans of all ages tend to be very accurate in classifying, discriminating and identifying happy faces (Gao & Maurer, 2009, 2010; Herba et al., 2006; Rodger et al., 2018; Thomas et al., 2007; Vicari et al., 2000; Widen, 2013). Perhaps, this finding could be explained by the idea that the tested children have a disproportionate amount of experience with perceiving this particular expression as they are generally approached in a positive manner in the Netherlands (see UNICEF, 2007). In addition, perhaps the pleasant atmosphere experimenters generally try to maintain during testing might unintentionally cause children to have an increased perceptual sensitivity for happy expressions like in the current study. More generally, the found age differences suggest that the children up to nine years develop in their ability to perceive facial expressions as categorically distinct. However, this ability seems to develop further during adolescence and adulthood (see e.g., De Sonneville et al., 2002; Kessels, Montagne, Hendriks, Perrett, & de Haan, 2013; Lawrence, Campbell, & Skuse, 2015). Finally, we calculated the response bias index in order to control for the possibility that children were biased to systematically confirm or disconfirm that facial expressions belonged to a certain emotion category. This bias could occur as an effect of fatigue, lack of concentration or not understanding a certain emotion category or the instructions of the FEP task. The average response bias index was close to zero for all children and did not differ across ages[3], indicating that children understood the task specifics and the emotion categories as displayed by the prototypical contexts.

Next, we found a negative relation between age group and the diversity of false alarms, meaning that older children identify fewer kinds of facial expressions as belonging to an emotion category than younger children (cf. Widen and Russell, 2008). For example, six-year-olds incorrectly identified a high diversity of expressions as belonging to the prototypical context of surprise such as happy, anger, sadness, fear, neutral, and disgust. However, nine-year-olds would merely incorrectly identify expressions of fear as belonging to surprise but hardly any other expressions. The age-related decline of diversity of false alarms for each target facial expression is most pronounced for expressions of anger, disgust and surprise, which is in accordance with other study results using various methods in showing that children between six and nine years of age develop vastly in terms of perceiving these facial expressions (cf. Gao & Maurer, 2009, 2010; Herba et al., 2006; Rodger et al., 2018; Thomas et al., 2007). Together, these results show that with age and experience, children seem to improve in differentiating facial expressions, allowing them to better identify them as belonging to their respective prototypical contexts.

Finally, the kinds of misperceptions for each target emotion were qualitatively analysed in order to asses if incorrect identifications across target emotions were made bidirectionally. We found that when children incorrectly identified a certain distractor facial expression such as anger for a target emotion such as disgust, this did not imply that the opposite was also true (cf. Widen & Russell, 2013). That is, if an anger was the target emotion, children would not necessarily incorrectly identify disgust as belonging to that emotion but would more often incorrectly identify neutral or fear as anger. Indeed, it is conceivable that a prototypical context for the emotion disgust (e.g., someone placing a plate of rotting food on your table) could evoke feelings or displays of anger towards the cause of that context. Alternatively, a prototypical context of anger (e.g., someone breaking your brand-new toy) might be less likely to evoke feelings of disgust (cf. Danovitch & Bloom, 2009). Sad expressions were often incorrectly identified as happy expressions while angry expressions and expressions of disgust were most often incorrectly identified as sad expressions. The fact that sad expressions were often incorrectly identified as belonging to the happy category seems peculiar and is not in line with the findings of Widen and Russell (2008) that have performed a similar analysis on mislabelling facial expressions by children up to five years of age. Therefore, we can merely speculate in accounting on this finding. For example, perhaps children found the sad face to be suitable for the prototypical context of happy that included

[3] Apart from a small and in our view, neglectable difference between six-year-olds and nine-year-olds in the happy emotion category with six-year-olds being slightly more conservative than nine-year-olds.

the statement that eating ice-cream makes people happy as they did not have an ice-cream at the time of testing, making them feel sad about it and leading them to respond accordingly. Only the combination of expressions of fear and surprise seemed to operate bidirectionally in that expressions of fear were often incorrectly identified as expressions of surprise and vice versa (cf. Widen & Russell, 2008). These findings contradict the idea that especially expressions that are dissimilar in emotional valence (i.e., happy is pleasant, sad is not) are not likely to be confused with each other. Instead, these results imply that emotional states cannot be placed on simple continua that separate the full domain of human emotions into distinct non-overlapping categories along a dimension such as either pleasant or unpleasant.

Study limitations & future research

In the current study, children had a West-European background which could be more generally classified as a WEIRD (Western, Educated, Industrialized, Rich, and Democratic) background (Krys et al., 2015). As facial expressions have different functions and meanings for humans depending on their background, response bias and diversity of identifications are likely to differ across populations with a different cultural background (Jack et al., 2012; Krys et al., 2015; Smith, Fischer, Vignoles, & Bond, 2013). This means that the found age differences in these measures in the current study cannot be generalized to children with a different cultural background. It would be interesting to conduct this research across humans with differing cultural backgrounds that might elucidate the age-related differences in cultural-specific functional relations between the perception of facial expressions as they occur in (social) contexts as and the adaptive responses to them.

Next, we found a negative relation between age group and average reaction times on the FEP task, meaning that as children grow older, they need less time to identify facial expressions as belonging to their emotion categories. Gao and Maurer (2009, 2010) have shown that in comparison to five-year-old children, children of nine and ten years of age also need less time to identify facial expressions and are better able to identify facial expressions that are presented with low intensities such as a mild expression of anger (for comparable results and methods, see Rodger, Lao, & Caldara, 2018; Thomas et al., 2007). Together, their results point to the idea that as children grow older, they require less visual information to identify facial expressions and identify them as belonging to distinct emotion categories. From a perceptual learning account, these results could be explained by the idea that with experience, children become more perceptually and selectively attuned (i.e. sensitive) to the invariant and distinctive features of each facial expression and the context in which they are presented (see Gibson, 1979; Goldstone, 1998; Hellendoorn, Wijnroks, & Leseman, 2015; Kellman & Massey, 2013; Leppänen & Nelson, 2009; Pollak & Kistler, 2002; Pollak, Messner, Kistler, &

Cohn, 2009; Zebrowitz, 2011). This form of learning is referred to as 'differentiation' (Adolph & Kretch, 2015; Genesee, Nicoladis, & Paradis, 1995; Gibson & Gibson, 1955; Gibson & Pick, 2000; Goldstone, 1998; Goldstone & Steyvers, 2001; also see Honey, Close & Lin, 2010 for a comparable notion of differentiation called 'acquired distinctiveness'). However, the age-related decrease in reaction time found in the current study and others could also be due to other factors such as an enhanced (oculo-) motor control. Therefore, although the idea that increases in perceptual sensitivity for the categories of facial expressions is due to differentiation seems plausible in our view, it remains speculative based on the current data and should be investigated further in future studies, using for instance eye-tracking technology.

Finally, we restricted the FEP task in terms of duration as to prevent effects of fatigue for the youngest age groups and in order to comply with the regulations of the ethical board of the university concerning the task demands for young children. In effect, each condition had twelve trials, which was too little data for reliable computations of the Shannon's Diversity Index of the incorrect identifications on the individual level and on the level of conditions. Therefore, Shannon's Diversity Index was only computed on the group level. With older participants or at least without the two youngest age groups, it would be more feasible to include more trials per condition, allowing to compute Shannon's Diversity Index of the incorrect identifications on the individual level. This could be used to investigate factors related to individual differences in Shannon's Diversity Index of the incorrect identifications. For instance, as recent literature suggests that children with ASD have trouble in perceptually differentiating (social) information (Hellendoorn et al., 2015), this deficiency is likely to be evidenced in a rather high diversity of incorrect identifications compared to their normally developing peers. Extending the FEP task in terms of adding more trials and hence the duration of the task might affect age-related memory demands of the task which needs to be controlled for.

Acknowledgements

We would like to thank Susan van der Baan, Kayla Bartels, Nina Chmielowice-Szymanski, Lidewij Hardenborg, and Julia Janssen for helping with the data collection, Tom ter Bogt and Paul van Geert for supporting this study and Jan Boom, Ludger van Dijk, Roy Hessels, Tim Walther, and the anonymous reviewers on providing helpful suggestions during the preparation of this paper. Finally, we would like to thank all the children that participated in this study.

Chapter 5

Modelling children's Gear task strategy use with the Dynamic Overlapping Waves Model

P. F. de Bordes
J. Boom
W. D. Schot
M. van den Heuvel-Panhuizen
P. P. Leseman

Published as:
de Bordes, P. F., Boom, J., Schot, W. D., van den Heuvel-Panhuizen, M., & Leseman, P. P. (2019). Modelling children's Gear task strategy use with the Dynamic Overlapping Waves Model. *Cognitive Development*, *50*, 237-247. https://doi.org/10.1016/j.cogdev.2019.01.001

Abstract

The Dynamic Overlapping Waves Model (DOWM) can model strategy use in problem-solving tasks for strategies that can be construed as developmentally and hierarchically ordered (Boom, 2015). We observed children's (M_{age} = 11 years, *SD* = 6 months) strategy use during a task in which they had to find the rotation direction of the last gear in a series of connected gear chains, given the rotation direction of the first gear. Using DOWM, we found that strategy use was ordered as expected, from unskilled sensorimotor strategies to abstract strategies, and from less to more efficient in terms of speed and accuracy. This order aligns with the idea that perceptual learning is central to the emergence of abstract conceptual knowledge. Moreover, the current study shows that the DOWM does not preclude forward and backward transitions and even occasional transitions that skip certain strategies in the ordering. The DOWM seems a promising tool to developmentally capture the breadth of behavioral repertoire children display when they adopt new strategies for various problem-solving tasks.

Keywords: Problem Solving; Dynamic Overlapping Waves Model; Perceptual learning; Embodied Experience

Introduction

Many studies have investigated how children learn Science, Technology, Engineering and Mathematics (STEM) topics (Bolger, Kobiela, Weinberg, & Lehrer, 2010; Lehrer & Schaube, 1998 Martin & Schwartz, 2005; National Research Council, 2014). In tailoring learning situations in which children could learn these topics, an important issue is how to support children's learning processes. While working on tasks, the students' strategy use reveals their progress in understanding the underlying STEM principles; see, for example, the torque principle in Siegler's (1976) balance scale task. One of the difficulties in describing and modelling children's learning processes during problem solving is that this development is not always continuous or step-wise and does not occur in similar fashion across children (Fischer & Bidell, 2006; Siegler, 1996; 2006; Van der Ven, Boom, Kroesbergen, & Leseman, 2012). In addition, children might vacillate between strategies (Boncoddo, Dixon, & Kelley, 2010), sometimes relapsing to less efficient strategies or make sudden jumps to more efficient strategies. This variability could be the result of relevant previous experience, different developmental pathways, or measurement error.

The variability in progress of attaining different strategies is commonly treated as measurement error in developmental theory, supposedly occluding the 'true' development or learning that takes place (Fischer & Bidell, 2006; Siegler, 2006; Zheng & Fischer, 2002). In addition, this variability poses difficulties for using standard statistical techniques that rely on linearity of strong order effects of the data (Boom & Ter Laak, 2007; Van der Ven, Boom, Kroesbergen, & Leseman, 2012). However, close inspection of the variable nature of children's microgenetic development in applying strategies in Siegler's Balance Beam (Boom & Ter Laak, 2007) and children's developing numeracy skills (Van der Ven, Boom, Kroesbergen, & Leseman, 2012) suggests that this variability is rather an intrinsic part of learning and can be taken into account when modelling microgenetic development (see van Geert & Van Dijk, 2002; Shrager & Siegler, 1998; Zheng & Fischer, 2002).

The Dynamic Overlapping Waves Model

According to Siegler (1996; 2006) children typically know and use several strategies to solve a given task at a given timepoint. With experience, the relative frequency in which certain strategies are used change, leading to increased or decreased use of certain strategies, abandoning old strategies, and discovering new ones (Siegler, 2006). So rather than conceptualizing children's developing strategy use in an all-or-non-fashion, children's strategy use can be seen as a competitive adaption of strategies at any given timepoint. Siegler (1996) conceptualized this kind of development of strategy use as overlapping waves (Figure

5.1). In his Overlapping Waves Model, the development of each strategy use is depicted in the form of a wave along a dimension of increasing maturation or sophistication (Siegler, 1996). Upon discovery of a strategy, children might use this strategy increasingly in respect to other strategies (depicted by a rise of the wave). But next, children might discover a newer strategy, slowly abandoning the now older one (depicted by a fall of the wave).

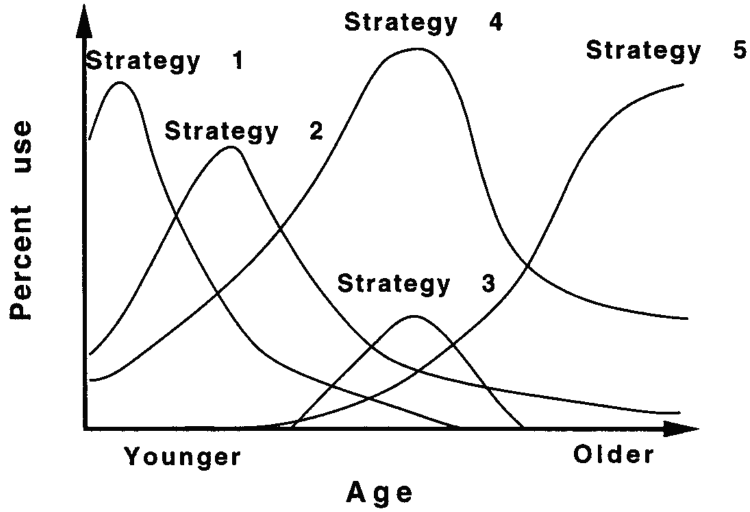

Figure 5.1. Overlapping waves depiction of cognitive development from "Emerging Minds," by Siegler (1996). This well-known picture is just a heuristic model, for more realistic distributions see our Figure 5.4.

Boom and colleagues have extended Siegler's (1996) Overlapping Waves Model statistically (Boom, 2015; Boom & Ter Laak, 2007; Van der Ven, Boom, Kroesbergen, & Leseman, 2012) by combining Latent variable Growth curve Modelling (LGM) and Item Response Theory (IRT) modelling. The resulting Dynamic Overlapping Waves Model (DOWM) promises several advantages in comparison to step-wise models or continuous models. First, it can address the frequently observed variability in children's strategy use, incorporating observed variability into the model instead of ignoring it. Second, it can model the use of several qualitatively different strategies in one model, by assuming a latent single continuous dimension that is nonlinearly and probabilistically related to the relative frequency of use of each strategy, instead of just averaging over different strategies. Thirdly, it can model the microgenetic development of strategies as a dynamic process, meaning that strategy use on time $t + 1$ is at least based on strategy use on time t. Finally, the model can deal with (large) inter-individual variation in timing of the learning process, meaning that it can easily incorporate data from participants who are on beginners, intermediate or advanced level regarding their proficiency on the learning task, and/or who differ in amount of change or learning.

Development of strategy use during the Gear task

In the current study, we applied the DOWM to capture children's development in strategy use on a problem-solving task involving force transmission through gear-chains, known as the Gear task (see, e.g., Dixon & Bangert, 2002). For this task, participants are typically asked to predict the direction of movement of a final gear (given the direction of the first gear) within a visually presented static chain of gears (Figure 5.2) and are asked how they obtained their answer (e.g., Dixon & Dohn, 2003; Dixon & Kelley, 2007).The different strategies to solve the Gear task and transitions between them have been well documented and defined (Boncoddo, Dixon, & Kelley, 2010; Dixon & Bangert, 2002, 2004; Dixon & Dohn, 2003; Dixon, Holden, Mirman, & Stephen, 2012; Dixon & Kelley, 2007; Dixon, Stephen, Boncoddo, & Anastas, 2010; Stephen, Boncoddo, Magnuson, & Dixon, 2009; Trudeau & Dixon, 2007), allowing these strategies to be used for complex modelling.

The results of the aforementioned studies using the Gear task typically show that most of the participants start out by literally tracing presumed local force transmission across gears by means of eye movements and/or gestures, known as the Force Tracing strategy (e.g., Dixon & Bangert, 2002). This suggests that participants use sensorimotor resources in order to simulate the local force transmission from each gear to the next (Alibali, Spencer, Knox, & Kita, 2011; Boncoddo, Dixon, & Kelley, 2010; Dixon & Kelley, 2007; Trudeau & Dixon, 2007). According to Dixon & Kelley (2007) classifying each subsequent gear as either left turning or right turning (i.e., Classification strategy) is usually developed directly out of the Force Tracing strategy. This abstraction occurs, according to them, because the actions that coincide with force tracing (i.e., moving hands or eyes along the gear chain from left to right and vice versa) contain alternation information. In support of this claim, Boncoddo, Dixon and Kelley (2010) found that the number of children's gestures associated with force tracing predicted later emergence of the Classification strategy (also see Trudeau and Dixon, 2007, for similar findings). This illustrates how finding an abstract strategy like the Classification strategy is perhaps ultimately grounded in sensorimotor actions whereby the movement of a gear is abstracted in terms of *directionality* (either left or right) of each gear along the chain.

A more efficient strategy that is seldom used by both adults and children is the Skipping strategy and consists of classifying only every other gear as turning into the same direction, starting from the first gear (e.g., Dixon & Bangert, 2002). This strategy eliminates the need to label (including pointing towards) each gear individually, making it more efficient in its execution. The directionality of the gears seems to be central to this strategy as with the Classification strategy and it already seems to hint towards the mathematical notion of parity. However, in contrast to the Parity strategy, the gears do not have to be counted in order to apply the Skipping strategy. Using the Parity strategy entails applying the rule that the first and last gears turn in opposite direction when the number of gears is even (even parity) and in the same direction otherwise (uneven parity). When using this strategy, the movement of each gear and directionality of the gears along the chain, central to the Force Tracing strategy and the Classification strategy, respectively, seem to be inherent to the *number of gears* within the chain that is used to apply the Parity strategy. Indeed, participants usually apply this strategy only after using the Force Tracing strategy, the Classification strategy, the Skipping strategy or all of them (Dixon & Bangert, 2002, 2004; Stephen, Boncoddo, Magnuson, & Dixon, 2009). Like the shift from the Force Tracing strategy to the Classification strategy described earlier, applying the Classification strategy and/or the Skipping strategy might yield information about the parity (i.e., left and right alternate along the chain of gears) that participants can use to discover the Parity strategy.

Embodied experience

Children up to 12 years of age rarely use the Parity strategy spontaneously within the Gear task (Dixon & Bangert, 2002). Because sensorimotor actions seem to play a pivotal role in developing abstract strategies for solving the Gear task (see Dixon et al., 2010) and learning abstract concepts in general (Rambusch & Ziemke, 2005), the use of the Parity strategy might be increased for this age group if they can manually explore energy transmission through gears prior to the Gear task. This would stress the importance of embodied experience in learning abstract strategies. In addition, it might allow us to include the Parity strategy in the overlapping waves model, elucidating the full range of development possible for this task. To this end, we devised an exploration task in which children could make either curved or straight gear tracks, using plastic gears. The alternating rotational direction of gears along the gear chain would be clearly visible when they would make straight tracks compared to curved tracks. Making the straight tracks therefore might lead children to already discover something about the functional relation of interlocked gears in terms of alternating rotational movements, possibly leading them to discover and apply more abstract strategies during the Gear task than children who made curved tracks.

Hypotheses

Taken together, we expect strategy use for the Gear task to be ordered developmentally in which the discovery of abstract strategies is ultimately grounded in embodied experience and specifically the sensorimotor actions that coincide with executing less advanced strategies. Our first hypothesis is therefore that the strategies can be ordered in terms of becoming more efficient, meaning that the more abstract the applied strategy is, the less sensorimotor input is required (as indicated by shorter response times) and the more trials are correct. Participants attune to less (and possibly different) perceptual information in order to perform well when abstract strategies are used compared to sensorimotor strategies, making them more efficient to solve the Gear task trials. Our second hypothesis is that there is a specific order of microgenetic development in which these strategies (i.e., the Force Tracing strategy, Classification strategy, Skipping strategy, Parity strategy, respectively) are used. Our third hypothesis is that children learn from embodied experience with energy transmission through gears by which they can discover and apply more abstract strategies on the Gear task afterwards. Specifically, we expect that if they would make straight tracks, they will apply more abstract strategies on the subsequent Gear task in comparison to those that made curved gear tracks.

Methods

Design

The present study has a microgenetic design. First, children's strategy use during the Gear task was assessed repeatedly (by means of interviews) over a short period of time. Second, we measured the number of (in)correct trials and the response time per trial of the Gear task. Third, we assessed whether making straight gear tracks in the exploration task would lead children to make more use of abstract strategies in the subsequent Gear task than if they would make curved tracks during the exploration task.

Participants

The sample consisted of 69 primary school children (M_{age} = 11 years; 0 months, *SD* = 6 months) from two different schools in the Netherlands. We excluded the data of one participant from the analysis since she had an arm in a cast. In addition, we excluded the data of two participants from the analysis because one child was not speaking loud enough for coding, and one child did not understand the questions during the Gear task (answers were about whether the child correctly solved trials or not, instead of *how* he solved it). Excluding these participants left 66 participants in the sample of which 29 were boys (M_{age} = 11 years; 1 month, *SD* = 6 months and 37 were girls (M_{age} = 10 years; 11 months, *SD* = 5 months).

Materials

Gear task

A total of 30 trials were displayed on a 17” computer screen using E-prime 2 (Schneider, Eschman, & Zuccolotto, 2002). Each trial contained a gear-chain with a purple gear with a yellow arrow pointing clock-wise, a yellow gear with a black, bidirectional arrow, and one to six black gears (5 of each, presented in random order across trials and across children) in between (Figure 5.2). Children indicated whether they thought the yellow gear would turn clockwise or counterclockwise by pressing one of two response buttons on the computer keyboard that were labelled with a symbol of a rotating arrow (one with a counter clock-wise arrow on the left and one with a clock-wise arrow on the right). After each response, auditory and visual feedback was provided for 2000 milliseconds. If the trial was solved correctly, a smiling mouse would appear saying “*very good*”. If the trial was solved incorrectly, a mouse with a neutral expression would appear saying “*incorrect*”. After each third trial, a mouse and question mark were presented, indicating a verbal response by the child was required in which the child explained how he or she solved the previous trials.

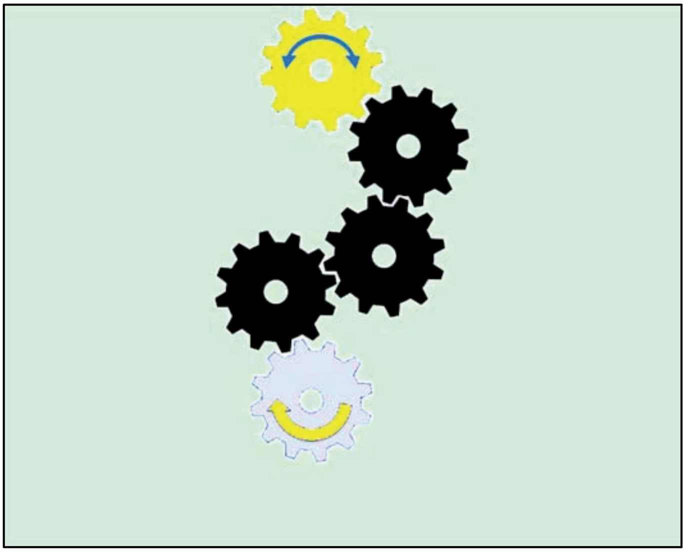

Figure 5.2. An example of a trial during the Gear task.

Exploration task

The materials for this task consisted of yellow, plastic gears, colored pegs (blue, yellow, and red) and plastic white pegboards in which yellow pegs could be placed as to attach the gears.

Procedure

General procedure

Testing took place in a quiet room in a school building. As a cover story, children were told that they would play two different games in succession, after which they had to decide which game they favored most.

Procedure Gear task

The Gear task was in part inspired by the task involving gears as used in several studies by Dixon and colleagues (e.g., Boncoddo, Dixon & Kelley, 2010; Dixon & Bangert, 2002, 2004; Dixon & Dohn, 2003). Children were seated, facing the computer screen at a distance of 70 to 80 cm, while a test trial displayed two adjacent gears; one purple and one yellow. The researcher explained that on each trial, there would be one purple gear (with an arrow always pointing clockwise) and one yellow gear (with bidirectional arrows) that were interlocked by one or more gears. Next, the researcher asked the child to which direction the yellow gear would rotate. Children usually answered this question correctly (three children gave an initial incorrect answer but revised their answer after a simple "*are you sure?*" question) by indicating (verbally or by means of pointing) that the yellow gear would go counterclockwise. Next, the response buttons were shown on the computer keyboard. The researcher explained that after every *three* trials, a screen with a mouse with the question mark would appear and the children had to state how they solved the last three trials. If children had no further questions, the task started. Questions were for example: "How did you solve the puzzle?" or "How did you know to what way the final gear was rotating", or after having asked this "How did you solve it this time?" Whenever the children's answer was vague or incomplete, follow-up questions were: "Can you explain that a bit more?" or "Can you explain that in different words?"

Procedure Exploration task

The experimenter introduced children to the task by presenting a plastic toy pegboard with five interlocking gears attached to it with pegs. The experimenter replaced the middle gear of the gear chain as to demonstrate how the gears can be placed in and out of the pegboard. The experimenter instructed children to make a gear-chain as such, that the first and last gear would rotate into the similar direction. While rotating the gears, the researcher stressed that the first and last gears (attached with the blue pegs) were rotating in the same direction to illustrate what the end state of each trial should look like. In addition, they were told not to touch the red pegs already present on the board with the gears. By using specific configurations of the red pegs, children were constrained to make either straight tracks or curved tracks in three different trials in a fixed order. Finally, the children were told that they had to indicate when the gear track was finished. After this instruction, the first pegboard with two gears on it (attached with blue pins), yellow pegs, and gears were placed in front of the child and the Exploration task started. Whenever children finished the gear track, they typically turned one of the gears that would cause for all the gears within the chain to move. We then asked whether the gears attached with the blue pegs rotated in the same direction. If this was the case and if the child confirmed this, we would praise the child and replace the board with the next one until all three parts of the task were finished.

Measures and coding data

We videotaped all ten answer fragments for subsequent coding of verbal and nonverbal behavior in terms of strategy use and we assigned a code of one to five for Guessing/ unknown strategy, Force Tracing strategy, Classification strategy, Skipping, and Parity strategy, respectively (see Table 5.1). To assure inter-rater reliability, we recoded 20% of the data on strategy use that resulted in a kappa of .84, which is high (Landis & Koch, 1977). So, a participant was credited with using one out of five mutually exclusive strategies for each of our ten measurement points (items) or was set to missing when the response could not be classified according to the strategy codes (a total of 5.09% had to be set to missing), resulting in 10 *item* codes per child. Next, we calculated the modus of strategy use on the gear task over the 10 measurement occasions per child in order to analyse the relation between the most frequently used strategy, amount of trials correct, and average reaction time per trial.

We registered responses as correct ‘1’ or incorrect ‘0’ and summed them, resulting in scores that could range from 0 to 30. In addition, we counted the number of trials correct per set of three trials after the children’s strategy use was inquired and transformed this number into percentages of items correct per used strategy across children. With this, we could analyze the relation between the percentage of trials correct per strategy use. Accompanying response times (in milliseconds) on each trial were also registered, using E-prime 2 (Schneider, Eschman, & Zuccolotto, 2002). The responses and response times were averaged within children for each set size (number of gears in trials). Of this dataset, two values were omitted from further analysis as they were more than 3 standard deviations from the average response time of children for that set size. Finally, no behavioral measures were used during the exploration task as the aim of this task was to test whether the condition contrast would affect strategy progress in the subsequent gear task.

Table 5.1. Strategy codes, definitions and examples

Strategy	Definition	Example
1. Guessing/other	The child either used an unknown strategy or stated that he/she guessed.	*"The curve in the gear chain is causing all the gears to go left" or "I don't know!"*
2. Force Tracing	Enacting local force transmission across gears by means of eye movements and/ or gestures.	*"This gear is going that way and it pushes the next gear that way"*
3. Classification	Classifying each subsequent gear as either left turning or right turning.	*"If this gear goes left, the next gear goes right and the one after goes left"*
4. Skipping	Classifying every other gear as turning into the same direction until the last gear is reached, starting with the first gear.	*"This gear turns the same way as that gear because there is one in between"*
5. Parity	Applying the rule that the first and last gears turn in opposite direction when the number of gears is even (even parity) and in the same direction otherwise (uneven parity).	*"It's an uneven number so the first and last go the same way" or "If the first gear goes left, the third gear also goes left"*

Note: Strategies were coded with both verbal and non-verbal behavior considered.

Analysis

Participants were asked the same questions ten times, with three trials of gear making in between questions. The coding system described above was used for each to these ten repeated measurements (the items). The core of our DOWM model is a hypothetical latent ability that is different for each participant (inter-individual differences), but also possibly changing within participant (intra-individual differences) over the ten measurement occasions. To examine level and change in this ability a Latent Growth Modelling (LGM) approach was used. An LGM presupposes a steady increase or decrease in such an ability over a small number of regularly spaced measurement occasions for each person, with the baseline or average over all measurement occasions denoted as the intercept, and the change (if any) from one measurement occasion to the next denoted as the slope. More on LGM, as a variety of longitudinal Structural Equation Modeling (SEM), can be found in Bollen and Curran (2006).

However, this ability itself cannot be directly observed: it is an underlying, hypothetical, hence latent ability. What only can be observed are the ten responses of the children, scored as Guessing/ unknown strategy, Force Tracing strategy, Classification strategy, Skipping strategy, and Parity strategy (coded as 1 to 5). So, we need to map the manifest display of these five strategies, in the participants recorded answers, to the latent ability of these participants.

Fortunately, Item Response Theory models can estimate the probability of choosing a certain strategy as a function of the properties of that strategy and the latent ability of the respondent, both expressed on a common scale, with interesting properties. With increasing ability, the likelihood of using the strategy 1 will decrease, the likelihood of using the strategy 5 will increase and the likelihood of using in-between strategies will rise and fall (depending on where you start; see Figure 5.4). The properties of that strategy are represented in Figure 5.4 by the location (difficulty) and by the relative flatness/steepness (discrimination) of the three shapes. The difficulty of the strategy is characterized by a parameter (the threshold) that defines the border between two subsequent categories (strategies), the discrimination (sensitivity) is characterized by the residual variance for the category. The larger the positive difference between the ability of the participant and the difficulty of the strategy, the larger the likelihood of using that strategy. The larger the negative difference between the ability and the difficulty of the strategy, the smaller the likelihood of using that strategy. Figure 5.4 illustrates the distributions of each of the five strategies so that the likelihood of responding according to one of the included strategies can be read off easily: e.g., with latent ability 1.0 (the scale has an arbitrary origin or zero point, so the absolute value is not important, but relative distances are) the likelihood of using strategy 3 and strategy 2 is almost the same, while the other strategies are not likely to be used by a hypothetical child with this level of ability. However, with a latent ability of e.g. 4.0, the highest three strategies (3, 4, and 5) have an almost equal chance of being used.

Figure 5.4 applies to all 10 repeated items equally because analyzing such strategy change requires measurement invariance. This means that all items are supposed to have the same properties, and only respondents can change over the 10 measurement occasions/items. Therefore, in the analysis, categories were restricted to have the same difficulty and discrimination[1] over the 10 measurement occasions/items. A general increase in how well participants do, would not be surprising, as would be individual differences in both starting point (position on the latent ability x-axis in Figure 5.4) and in change (how much a child would move to the right along this latent ability axis) over the 10 occasions/items. We used a Graded Response Model (GRM; Ostini and Nering, 2006; Samejima, 1969) appropriate for multi-category items. An important assumption of such a model is that strategies are orderd as a series of steps that are mastered in sequence when solving trials on the gear task. Overall fit will indicate whether this assumption is tenable. The model reveals furthermore whether the strategies are equally difficult or not, and how much overlap there is, as can be seen in Figure 5.4. The model thus provides important information about the strategies.

[1] Implemented by requiring that thresholds are the same across occasions/items and that residual variances are the same by fixing them to 1 for all items. We also fixed the intercept mean to zero, as this was a convenient way to anchor the scale which is needed for model identification. More detailed information can be obtained from the second author.

In sum: We combined and integrated Item Response Theory (IRT) modeling and Latent Growth Modelling (LGM) to model the responses of the participants on the ten questions. The IRT part is not only needed because the data are categorical and the LGM part is not only needed to account for the repeated nature of the measurements[2]. The combined and integrated analysis offers more: the model allows to relate differences in latent ability of participants to differences in the likelihood of using the particular strategies (in an inter-individual sense) but also to model increases in ability to shifts in using the particular strategies (in an intra-individual sense). Learning in this sense implies a shift in the probability distribution of strategy use: chances increase of using a more advanced strategy while chances decrease of using a less advanced strategy.

Although we conceptually distinguished two components in the model, the entire analysis was done as one statistical model in Mplus 8, with the Weighted Least Squares Means and Variances adjusted estimator (WLS-MV), a PROBIT link, and Theta parameterization (Muthén & Muthén, 1998-2017). An advantage of using the WLS-MV estimator in Mplus is that absolute overall fit measures are available. This overall fit of the model was evaluated in terms of the indices Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker Lewis Index (TLI), following commonly applied cut-off criteria. Specifically, model fit was considered good if RMSEA was below .05, and CFI and TLI were above .90 (Little, 2013).

Results

Descriptive statistics

On the Gear task, children had on average 23 trials (SD = 5.6) out of 30 trials correct and spend on average 10 seconds (SD = 5.25 sec.) per trial. Making straight or curved gear tracks during the exploration task did not affect performance on the Gear task in terms of number of trials correct ($M_{difference}$ = -2.52, p = .07), modus of strategy use ($M_{difference}$ = -.015, p = .95), and average response times ($M_{difference}$ = 523.80 ms, p = .64). The average response time per child per trial was related to the set size (number of gears visible trials), r = .45, $p < .001$. Gender did not affect this relation as split sample correlations did not deviate from this outcome or from each other. There was no relation between the set size and whether it was correctly solved, r = -.003, p = .88. In sum: while accuracy *did not* increase with increasing set size, the average response time and the standard deviation of the response time *did* increase with increasing set size (see Table 5.2).

5

[2] Because measurements are repeated, correlations between items are to be expected, which would violate IRT assumptions. However, because LGM accounts for this correlational structure between items, the IRT model can still be used. See example 6.5 in the Mplus users manual (Muthén & Muthén, 1998-2017).

Table 5.2. Group results of percentage of trials correct, average response time (in ms) and standard deviations of the response time (in ms) for each set size (amount of gears displayed in a trial) of the Problem-solving task

	Set size					
	3	4	5	6	7	8
Percentage trials correct	0.78	0.80	0.79	0.80	0.77	0.78
Average response time	5738	7687	8848	10240	11412	13412
SD response time	2765	4301	5171	5423	5396	7530

Frequencies of strategy use per measurement point during the Gear task are shown in Figure 5.3. A total of 41 11-year-olds (62.12%) used the Force Tracing strategy at least once, 38 children (57.58%) used the Classification strategy at least once, 16 children (24.24%) used the Skipping strategy at least once, and 6 children (9.10%) used the Parity strategy at least once.

As can be seen in Figure 5.3, the frequency of children using the Force-tracing and Classification strategies slightly decreases over the course of measurements whereas the frequency of children using the Skipping and Parity strategies increases slightly. However, it is unclear how all these strategies relate to each other intra-individually in terms of developmental sequence: note that if the frequency at a next timepoint is the same (or higher) it does not imply that the same children are involved. Therefore, in order to test whether the strategies have the developmental order as hypothesized in the introduction, we modelled strategy use in a manner that takes the individual developmental trajectories into account (see section 3.3).

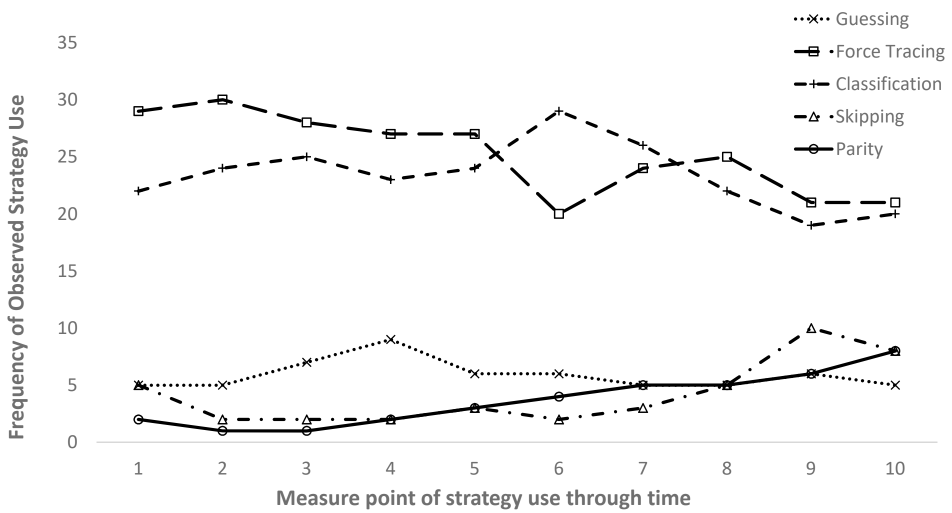

Figure 5.3. Frequency of observed strategy use per measure point during the Gear task.

Number of trials correct and response time

Interindividual correlations between gender, age, number of trials correct, average response time, and modus of strategy use are shown in Table 5.3. As can be seen, there was a positive correlation between the modus of strategy use and number of trials solved correctly ($r = .37$, $p < .001$), meaning that children who applied more abstract strategies yielded more correct answers. Table 5.4 further reveals that more abstract strategies yield a larger percentage of trials correct. In addition, there was a negative correlation between modus of strategy use and response time ($r = -.44$, $p < .001$), meaning that children who applied more abstract strategies solved the task faster. Whether strategy use is related to number correct and response time on the intraindividual level, is addressed in the next section.

Table 5.3. Correlation matrix between gender, age, number of trials correct, average response time, and modus of strategy use

	1	2	3	4	5
1. Gender	–	-.09	-.16	-.03	-.13
2. Age		–	-.01	-.04	.16
3. Number of trials correct			–	-.11	.37**
4. Average response time				–	-.44**
5. Modus of strategy use					–

Strategy	Proportion correct
Guessing	0.57
Force Tracing	0.73
Classification	0.86
Skipping	0.94
Parity	0.94

5

Dynamic Overlapping Waves Model

Using the DOWM as analysis tool revealed several interesting results. In Figure 5.4 the Category Characteristic Curves (CCCs) of the final model are given. As explained in the method section, the likelihood of responding according to one of the strategies can be read off easily: e.g., with latent ability 1.0 the likelihood of using strategy 3 and strategy 2 is almost the same, while with a latent ability of e.g. 4.0, the highest three strategies (3, 4, and 5) have an almost equal chance of being used.

The scale on the X-axis represents latent ability and learning and pertains to the children. Whereas such a figure is customarily used for charting inter-individual ability differences, and that still is possible, it can also be used for charting intra-individual learning differences. With increasing learning level (or ability, or maturity, or development, or moving to the right) the likelihood of responding according to a particular strategy changes: for the guessing strategy it clearly goes down, for the parity strategy it increases, while for the in-between strategies it first rises and then falls. The thresholds determine the position of the CCCs as they demarcate where a next strategy becomes dominant. The thresholds and curves are invariant and fixed over items, and thus over time, and define the ruler that makes it possible to gauge intra-individual change and inter-individual differences. When it comes to inter-individual differences within our sample of 11-year-olds, it seems that most children with the average ability (0 on the x-axis of Figure 5.4) apply the Force Tracing strategy. In addition, children in this sample were least likely to use the Skipping strategy (indicated by the lowest peak on the y-axis, compared to the other strategies).

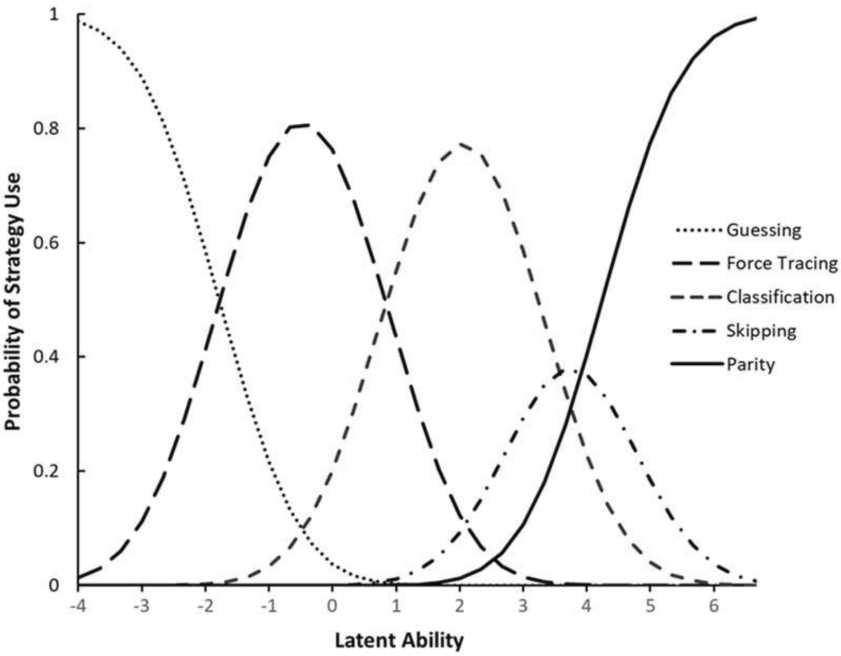

Figure 5.4. Category Characteristic Curves (CCCs) with Probability of Strategy Use on the Y-axis and Latent Ability on the X-axis. The Latent Ability scale may stand for inter-individual differences, but also for intra-individual learning differences.

The fit of initial model was not acceptable (RMSEA 0.111, CFI 0.958, and TLI 0.967), but because the sample size is small and the results seemed promising, we accepted two modifications to improve overall fit: item two has been discarded entirely due to improper behavior, such as a negative residual variance, and item 4 and 6 were allowed to co-vary (based on the modification indices). The improved model had acceptable to good fit (RMSEA 0.076, CFI 0.979, and TLI 0.984). The average slope increased 0.134 units per occasion[3]. So, the average increase over all 9 measurement occasions/items is (8*.134) 1.072 units on the X-axis scale in Figure 5.4. There was a significant variance for both the intercept (1.87 units) and the slope (0.062 units). In addition, the individual variance of the slope could be predicted from school-type (school A or B; beta 0.54), but not by Condition (straight or curved gear-chains in the exploration phase; ns). However, since their interaction had a considerable effect (beta -0.423) we kept Condition in the model (fixed at zero). We centered the school and condition indicators and used these for creating an interaction term. The intercept variance could not be explained by these same predictors.

There was a positive correlation between the level of ability (intercept) and number correct ($r = .50$, $p < .001$) and a negative correlation between ability (intercept) and response time ($r = -.565$, $p < .001$) confirming and accentuating what was reported in 3.2. In addition, there is a positive correlation between the change in ability (slope) and number correct ($r = .30$, $p < .001$). A negative correlation between change in ability (slope) and response time was not significant ($r = -.147$, $p = .303$).

Discussion

In this study we measured children's reasoning and performance on a problem-solving task involving gears. We predicted that the strategies can be ordered in terms of becoming more efficient as using more abstract strategies yielded shorter response times per trial and a larger number of trials correct. We found that children using a more abstract strategy had more trials correct. One explanation for this might be that more abstract strategies like the Parity strategy relies less on sensorimotor actions and different perceptual information, like when applying the Force Tracing strategy, that is prone for errors. For example, when children apply the Force Tracing strategy, they visually track the rotational direction of each individual gear and the direction the gears interlock along the whole gear chain. When applying the Parity strategy, however, they simply count each gear along the chain without tracking the movement of each gear or the way the gears interlock. In support of the claim that abstract strategies require fewer sensorimotor actions, we found that more abstract strategies require less time to solve the Gear task trials. Finally, although the set size of the trials was positively related with the time needed to solve those trials, set size did not affect whether children found the correct answer.

[3] For the initial model all the parameters had very similar estimates.

Next, we assumed that the different strategies for solving this task would have a specific order in microgenetic development, running from strategies that draw heavily on sensorimotor actions (i.e., Force Tracing strategy) to strategies that are more abstract (Parity strategy). That is to say, we expected the probability of using each of the strategies to rise and fall as experience with the task increases, with abstract strategies gaining probability to be used and more rudimentary strategies losing probability to be used. This is important in that it is consistent with the idea that strategy use can be seen as a competitive adaption of strategies at any given timepoint and as experience on the task increases, more abstract strategies will be likely applied. We tried to fit the five strategies into the model in this presupposed order. The order itself therefore, strictly speaking, cannot be tested, but the good fit we found supports this presupposed order as at least plausible. Although the model presupposes that the strategies are developmentally ordered in general, the, strategy use on the individual level is not always sequential or even progressive. This is often considered to be error variance but here we entertain the possibility that ability itself on the task can deteriorate (e.g., by inattentiveness) or make sudden jumps (e.g., after experiencing a sudden flash of insight). This natural variability could be seen as a process of repeated building up and collapse of task-ability (i.e., "scalloping", Zheng & Fischer, 2002) and is perceived to be crucial for learning to occur (Siegler, 2006; Zheng & Fischer, 2002). The DOWM allows for a limited amount of such backward transitions and even transitions that skip certain strategies on the individual level.

Taken together, we perceive this microgenetic development as a result of perceptual learning in which children are able to differentiate useful information for a more advanced strategy while working on the task with a certain strategy (see Adolph & Kretch, 2015; Adolph, Joh, Franchak, Ishak, & Gill, 2008; Gibson & Pick, 2000). For example, the Classification strategy is usually found after applying the Force Tracing strategy because alternation information that is used for the Classification strategy is nested within the actions that coincide with force tracing (Dixon & Kelley, 2007). From a perceptual learning account, using a more abstract strategy in this sense does not mean that it is based outside a physical realm (e.g., mental representation) but rather that it is nested within the prolonged interaction with the task materials (also see Dixon et al., 2009; Trudeau & Dixon, 2007 for a similar account).

Finally, manual exploration of gears in the exploration task might have helped children to discover and apply abstract strategies in the Gear task later on as a considerable number of children in the current study used the Skipping strategy (24.24%) and the Parity strategy (9.10%), in comparison to children of comparable ages in the study of Dixon and Bangert (2002). This allowed us to adopt the Parity strategy in DOWM. In addition, this seems to stress the importance of embodied experience in acquiring an advanced understanding of STEM related topics (Abrahamson & Lindgren, 2014; Weisberg & Newcombe, 2017). However, the effect of other differences in task design between our study Dixon and Bangert's (2002) study cannot be ruled out. Hence, the effect of embodied experience with gears in the exploration task on the discovery and application of abstract strategies in the subsequent Gear task is solely speculative. We did not find any differences in strategy use on the Gear task between children constructing curved or straight gear tracks during the exploration phase, possibly because the difference between the conditions might have been too subtle or because the sample was too small as the found difference was close to significance. This still leaves the question open of what kind of sensorimotor input is most beneficial to acquiring an abstract notion of the system at hand (see Pouw, Van Gog & Paas, 2014 for an extended discussion on this issue). For further investigation of this matter, it might be helpful to increase the salience of the alternation in rotational movements along the gear chain by assigning each alternating gear a specific color (see Dixon & Dohn, 2003).

Limitations of the study

A limitation of our study is that some children had difficulties initiating the Gear task. This was evidenced by a large variation in both response times as performance (correct/incorrect) of the second trial and difficulty in resuming the task at the fourth trial, just after they received the first question on how they resolved the first three trials. This seems to indicate that the training trail was insufficient for children to get properly accustomed to the task demands and that the first items might have measured task understanding, rather than task proficiency. Ideally, we would have preferred to drop the first 6 trials (= 2 items) from the analysis to be assured that task proficiency instead of task understanding was measured. However, given the limited number of trials, this was not feasible. For future research, the training session should be extended and should include the interview questions that were used during the Gear task to get children accustomed to these questions. Next, it is known that younger children, unlike adults, hardly arrive at using the parity strategy (see Dixon and Bangert, 2002), meaning that the frequencies of strategy use are age dependent. However, including younger and older children would have been interesting as that would make it possible to analyze whether development of strategy use is affected by age, which is feasible using DOWM as age can simply be added to the model as a predictor. Finally, it would have been interesting to perform an analysis that could reveal around which point(s) of measurement children are likely to progress in strategy use. However, this analysis would not yield robust results because of the small sample used in this study.

Strengths of the DOWM

As the DOWM can deal with (large) inter-individual variation in timing of the learning process, it can easily incorporate data from participants who are on beginners, intermediate or advanced level within the proposed microgenetic development. DOWM takes the observed variability of strategy use of all participants into account in constructing the Category Characteristic Curves for each strategy separately. Importantly, these Category Characteristic Curves are placed along one latent single continuous dimension that is nonlinearly and probabilistically related to the relative frequency of use of each strategy, instead of just averaging over different strategies. Therefore, it can display the full breath of how a *repertoire* of strategy use can develop across children within a single model instead of proposing a single possible developmental pathway along the different strategies or the development of a single strategy. With typical item response functions for polytomous items the personal latent ability can be linked to a profile of responding in one of the categories. Therefore, the most important advantage of applying the DOWM might be that the relations between the use of the strategies in tasks like this, that have always been difficult to conceptualize, are now easily accessible both conceptually as well as computationally.

We have found an effect of school on strategy use in the DOWM, meaning that children from one school seemed to apply more abstract strategies during the Gear task than children going to the other school. We did not predict this difference and to our knowledge, neither schools have taught children anything specifically on force transmission by means of gear chains. Although this unexplained variability can be perceived as a weakness of the current study, it can also be perceived as a strength of the DOWM in that it can incorporate data of children from different schools in modelling children's microgenetic development.

For example, it can be used to test hypotheses on the effectiveness of different teaching methods in terms of what children gain in understanding STEM topics over time. In addition, it could assist teachers in designing educational activities that provide children the right challenges at the right time, in the right sequence and at the proper pace. Finally, it can be used to test predictions from theories on the amount and developmental order of discrete steps in which children can learn such as strategy use (e.g., for the Balance Beam, see Siegler & Chen, 2002) or skill level (Fischer & Bidell, 2006; Van Der Steen, Steenbeek, Van Dijk & Van Geert, 2014). Therefore, we suggest that next to modelling strategy progression, DOWM is a promising tool that can be used on a wider scope in education and research. To further discover the applicability of DOWM and exploit its merits, we suggest that for future research the DOWM is applied on children's strategy use on different tasks and with larger populations.

Acknowledgements

We would like to thank Suzanne van Vliet, Nelleke Doelman, and Lineke Spiele for helping with the data-collection and the Curious Minds platform, Platform Beta-techniek, and professor Tom ter Bogt for supporting this study.

Chapter 6

General conclusions and epilogue

Introduction

The general aim of this dissertation was to show the potential of the ecological view on developmental psychology in accounting for how infants and children are able to develop and demonstrate skilful behaviours such as gaze following, imitation, facial expression recognition and understanding physical mechanisms. Throughout the dissertation, it was argued and shown that children's learning and mastering of these skills occurs as a result of perceptual learning which involves the attunement of their perception-action system to the relevant information (i.e., affordances) in the environment. This attunement was conceived as a developmental process through which they perceive order amid change and change amid order in the service of adaptive action. In several chapters, we offered this ecological description of how infants and children are able to develop and demonstrate skilful behaviours as an alternative to the cognitivist explanations, as contrasted in the introduction of the dissertation. In this final chapter, conclusions of each chapter are presented in relation to future research and theory building, and applications within the domain of education and therapy. This chapter concludes with an epilogue in which final remarks are given on how research on developmental psychology can move forwards by leaving the cognitivist view behind in favour of adopting the ecological view.

Chapter conclusions

Perceptual basis for gaze following

In Chapter 2, we have shown that infants are able to follow the eye gaze of another person whenever they visually attune to the eye region of that person and without the necessity of interpreting the intent of that person. We demonstrated that this attenuation could be modulated and enhanced by drawing their visual attention towards the eyes of someone just before the eyes moved from a frontal position to a lateral position. This modulation worked irrespective of whether that person displayed social cues, such as eye contact that infants could interpret as a communicative intent, or non-social cues, such as blinking and moving dots. In our view, these results render the idea that infants need to interpret someone's communicative intent (as cued by eye contact) in order to follow gaze, as obsolete. This invites a closer look at the perceptual processes that might elucidate how infants are able to follow gaze. To that end, we proposed in Chapter 2 that infants have a tendency to look at the face and especially the eye region of a frontally presented face, and that they tend to orient towards the direction of subsequent movements of that face and eyes, resulting in gaze following.

Since the publication of Chapter 2, other researchers have drawn similar conclusions, based on studies using variations on the paradigm as described in Chapter 2 (see Gredebäck et al., 2018; Szufnarowska, Rohlfing, Fawcett, & Gredebäck, 2014). All these studies seem to agree on the idea that "young infants learn from what we do, not from what we intend to communicate" (Gredebäck et al., 2018, p. 2097). Following this line of thought, researchers have recently proposed that gaze following is a skill that can dynamically emerge over the first years of life across different contexts in which the perceptual system is progressively attuned to perceiving the eyes of other people, their eye movements and the accompanying line of sight (Del Bianco, Falck-Ytter, Thorup, & Gredebäck, 2019). Using longitudinal research designs, this perceptual learning account on gaze following could be further substantiated in future research in order to find out which kinds of experiences facilitate this perceptual attunement across different contexts and which (perceptual aspects of these) contexts can elicit gaze following (Astor et al., 2020; Del Bianco et al., 2019). Recordings of naturalistic social interactions such as diaper changes and feeding might serve as an interesting starting point for this type of research (see Nomikou, Leonardi, Radkowska, Rączaszek-Leonardi, & Rohlfing, 2017). This is because these types of interactions are reoccurring and repetitive, making it relatively easy to detect developmental changes in the manner infants act upon the eye gaze of their caretaker. Ultimately, this line of research can also shed light on how to teach children to follow gaze when they experience difficulties in doing so, in which situations they make use of it effectively, and how it can be provoked situationally when deemed necessary.

Following gaze towards affordances

In Chapter 3, we further investigated gaze following and we asked whether gaze following might assist children in detecting object affordances in a social learning context. An affordance was defined as an action possibility provided by the social and/or material environment and the action capabilities of the infant (Gibson, 1979). We ran an experiment in which we experimentally either facilitated or hindered infant's attunement towards the eyes of a person presented on a video screen, after which that person acted out a failed attempt to realize a target affordance with two objects (e.g., failing to put a peg in a hole). We found that when attunement towards the eyes was facilitated, infants matched the eye gaze of the person more often, had more success in realizing the target affordance, and needed less exploration time to realize the target affordance than infants that were hindered to attune towards the eyes. This suggests that the gaze of the person, in addition to the person's object movements, guided (or educated) the perceptual system of infants by which they attuned to critical aspects of the failed attempt through which they learned a certain affordance. In other words, gaze following seems to have influenced the manner in which infants perceived object manipulations carried out by another person, helping them to discover something novel.

Although gaze following might have helped infants to realize a novel object affordance, it is not a crucial component of the learning context in imitation learning. For instance, almost half of the infants that were hindered to see the eyes of a person before she acted out the failed attempt demonstration were still able to realize the target affordance, perhaps in part by means of manual exploration of the objects as they needed more time to establish the target affordance. This suggests that in a learning context as studied in Chapter 3, some infants might benefit from gaze following, whereas other infants might benefit more from perceiving the object movements alone, or manually exploring the objects. This multicausality in learning and development is easily overlooked when looking for main effects of study manipulations as was done in Chapter 3. Therefore, the complex interplay of contextual and personal factors that causes the observed variability of learning success across infants should be the topic of future investigation (see Thelen et al., 2001; Yu & Smith, 2017). A possible way to further investigate this is by studying the spatiotemporal structure of the naturally occurring and emergent dynamics of interaction and coordination of gaze and behaviour between the infant and a caregiver across various learning contexts (for recent examples, see Monroy, Chen, Houston, & Yu, 2020; Northrup & Iverson, 2020). For example, indices of coupling and coordination of behaviours such as eye gaze in dyadic settings have been linked to the quality of collaboration and mutual understanding in adult studies (e.g., Richardson & Dale, 2005). Measuring these types of indices of interactions between infants and their caregivers might elucidate how they adapt and attune their behaviour dynamically in ways that allow infants to learn from their caregiver effectively.

Perceptual differentiation of facial expressions

In Chapter 4, we used a perceptual learning account in describing how children develop their abilities to perceive facial expressions as categorically distinct. We discussed studies that show how cultural background and specific experiences can shape the manner in which facial expressions are perceived and distinguished. With this, we argued that with accumulating experiences of perceiving facial expressions across different contexts, facial expressions gradually become more differentiated in terms of the contexts in which they are typically perceived. This differentiation occurs at the level of perception, which means that children are increasingly sensitive to the perceptual information of the facial expressions and the context in which they are perceived. This in turn can specify which types of actions are adaptive in response to it (i.e., affordances). For example, an adult perceiving a person with a sad facial expression on a funeral might comfort that person in specific ways, depending on factors like proximity, intensity of the expression and the presence of other people. Alternatively, a bullying young child with Conduct Disorder might not perceive the sad face of his/her victim as a sign to stop bullying (see Blair & Coles, 2000).

In order to study children's development of differentiating facial expressions, we presented six-to-nine-year-olds six conditions which each started with a depiction of a prototypical context for a particular emotion category, such as a depiction of infested food and dog faeces to which most people respond to by expressing disgust. We then asked them to identify whether different kinds of facial expressions belonged to the previously presented prototypical context, or not. We found that with increasing age, children correctly identified more facial expressions as belonging to their prototypical contexts. In addition, they were also better at correctly identifying facial expressions as not belonging to prototypical contexts of other facial expressions. This shows that with age and experience, children perceive facial expressions as increasingly specific to the context in which they typically occur. Finally, with increasing age, children identified less kinds of facial expressions as belonging to a context in which they are unlikely to occur. This reveals that older children are better at perceiving the differences between facial expressions than younger children. Together, these results are in line with the idea of perceptual learning that with increasing age, children's perception of facial expressions becomes more differentiated.

In the current study, we predefined facial expressions into separate and non-overlapping categories and as belonging to specific prototypical contexts which we believed to be specific to those facial expressions. We did this in order to study age differences in how well children differentiate these categories of facial expressions as belonging to the prototypical contexts. In reality however, the appearance of category boundaries might be much more flexibly related to the contexts in which they are perceived, as has been shown to be the case for the speech sound categories (Case, Tuller, Ding, & Kelso, 1995; Hasselman, 2015). For example, in bullying situations, a perceived bully that expresses emotions of joy and happiness in reaction to the distress of a victim might call for a different reaction from the observing child (e.g., scolding the bully) than if a perceived person displays emotions of happiness after winning a contest (e.g., congratulating the winner). Although in both contexts, children might be able to correctly categorize the perceived facial expression as 'happy', it is likely that they do so by attuning to different perceptual information, leading to different responses. Further investigation is required in order to find out how children are able to differentiate facial expressions as a result of perceiving them across different contexts and how this relates to adaptive responses to this information.

The ecological view and domain-specificity

In Chapter 2, Chapter 3, and Chapter 4 studies were reported on the social-cognitive abilities of infants and young children, respectively, following the eye gaze of another person, using gaze direction to learn specific object affordances and identifying facial expressions of others. The combined results of these chapters point to the idea that social information like gaze direction or a facial expression does not require a specific (potentially inborn) social capacity to read other people's intentions or to interpret the emotional valence of facial expressions. Instead, they show that social information is picked up and used similarly as physical information in the environment. That is, by attuning the perceptual system to invariant structures that optimally specify certain affordances (Gibson & Pick, 2000; Ingold, 2001; also see Chapter 1). For instance, the differentiation process by which children (learn to) detect the invariant relation between a facial display of disgust and the likely presence of something repellent such as rotten food that should not be eaten (see Chapter 4) might not be fundamentally different from discovering a new strategy to solve a problem-solving task that is based on invariant information within the execution of a current strategy (see Chapter 5). That is because in both situations, attunement of the perceptual system can lead to specific information in terms of affordances that can guide further perception and action in an adaptive manner.

On a wider scope, this conclusion can be seen as an invitation to reflect on the viability of the manner in which scholars order the human behavioural repertoire in distinct capacities or domains, like social, perceptual, and mathematical capacities, that humans develop independently from each other and can call upon when needed (e.g., Spelke & Kinzler, 2007). This ordering potentially leads and has led to confining specific human capacities (e.g., social capacity) within certain behavioural contingencies (e.g., social interaction) that are unwarranted and counterproductive. For instance, children diagnosed with Autism Spectrum Disorder (ASD) are mainly characterized by having "persistent deficits in social communication and social interaction across multiple contexts" (American Psychiatric Association, 2013, p. 50). Therefore, most clinical interventions and therapies aimed at helping and educating children with ASD are directed at improving their social skills and social understanding (e.g., Pickles et al., 2016). However, it is also known that ASD is a pervasive disorder, meaning that children with ASD show general deficits in behaviour beyond the social domain such as the display of repetitive behaviours (Kirby, Boyd, Williams, Faldowski, & Baranek, 2017). According to the research of Hellendoorn and colleagues, these general deficits can be summarized as consisting of atypical patterns of explorative behaviours and invariant detection, leading young children to detect different affordances in comparison to normally developing children (Hellendoorn et al., 2014; Hellendoorn, Wijnroks, & Leseman, 2015). More specifically, the repetitive and invariant behaviours that children with ASD often display might possibly hinder them in detecting invariant structures within the relation between them and their environment. Variability within this relation might play a key role for attuning the perceptual system towards information that specifies affordances in an adaptive manner, as will be argued in the next sections of this chapter. In line with the conclusions drawn in Chapter 2, Chapter 3, and Chapter 4, Hellendoorn et al. (2015) conceive perceptual learning and invariant detection as the basis of learning and development in general, including the development of social capacities. If so, atypical patterns of explorative behaviours and invariant detection during infancy can have cascading developmental effects within and across several (supposed distinct) behavioural domains, including the social domain. Therefore, it seems sensible to consider the use of clinical interventions and therapies directed at their perceptual learning skills that are not solely confined to the social domain, but most certainly affect it. Interventions aimed at increasing the behavioural variability of young children with ASD might be a good start for this (e.g., Rodriguez & Thompson, 2015).

Emerging insight through attuning action and perception.

In Chapter 5, we investigated children's development in strategy use microgenetically on a problem-solving task involving a physical mechanism with interconnected gears. In this task, children viewed tracks of interconnected gears and had to predict the rotational direction of the final gear, given the rotational direction of the first gear (see Figure 5.2). Using a perceptual learning account, we described how children are able to create and subsequently attune to critical perceptual information, while working on a problem-solving task. In other words, we suggested that within the interaction with the task materials, perceptual information becomes available for children that can be coupled to novel behaviours in adaptive ways. Attunement to perceptual information in this manner might enable children to discover novel and advanced strategies by which they could execute the task more efficiently. For instance, most children start working on the task by using their fingers to literally trace the rotational movements of each gear along the gear track (i.e., force-tracing strategy). These movements running from left to right and back along each gear contains specifying and invariant information, namely that the direction of rotational movement alternates along each gear within the gear chain. This information can then be used to execute a more advanced strategy, which is to classify each gear as either left-turning or right-turning (i.e., 'classification strategy', also see Trudeau & Dixon, 2007). Children's developing strategy use was modelled with the Dynamic Overlapping Waves Model (DOWM). In this model, strategy use was conceived as a multi-stable behaviour, meaning that strategy use was depicted as a probability distribution which represents the likelihood of using particular strategies on any given time point during the task. By using this model, we found that strategy use was ordered as expected from the perceptual learning account, running from unskilled sensorimotor strategies to abstract strategies, and from less to more efficient in terms of speed and accuracy. The term '*abstract*' strategy here does not refer to the idea that it is based outside a physical realm (e.g., mental representation). Instead, it refers to the idea that it is nested within the execution of a more rudimentary strategy in which the perception-action system could attune to the specifying information needed to execute a more advanced strategy.

In Chapter 5, the merits of DOWM were discussed in terms of its applications within the context of education. For example, it can be used to track down the frequency of strategy use and attained levels of skill by children on a learning task and the order in which these strategies and levels of skill develop. This could be used for assessing the quality of a learning task as it can reveal what children could potentially learn from it, and in which pace they could do so. In addition, DOWM can also be used to compare learning tasks in order to asses which learning task can accommodate learning by children with a specific range of entry level of strategy use or levels of skill. In this manner for instance, learning tasks can be created in order to accommodate learning by children that have a low entry level, high entry level, or a diverse entry level. Next, as DOWM can identify multi-stable behaviours in the observed developmental order, educators and researchers could use these momenta of instability to provide children with specific perceptual information that can trigger a transition to a more advanced skill level or strategy use (as will be discussed in the next section). Taken together, DOWM can inform educators to provide children the right challenges at the right time, in the right sequence, and at the proper pace.

Children's education of attention

In Chapter 5, we discussed the possibility of changing the used problem-task in ways that might help children in attuning to information that allows them to discover and use abstract strategies earlier during the execution of the problem-solving task. This idea is very similar to the practice of scaffolding that can boost the performance of children in educational settings (e.g., Belland, 2017). However, enhancing performance might actually hinder children from learning in some cases, as Dixon and Dohn (2003) have illustrated. Dixon and Dohn (2003) performed an experiment in which they structured a problem-solving task in such a way that participants were stimulated to discover and use an advanced strategy. After this task, participants engaged in another learning task in which the same strategies could be applied. Notably, only a few of these participants were able to apply the advanced strategy that they learned on the first task on the second task. For a second group of participants, the first learning task was not structured. Interestingly, participants of this second group, who also discovered the advanced strategy during the first task, were able to apply this strategy on the second task. This suggests that if a learning task is structured as such that it channels students towards optimal performance, attunement of perception and action towards the relevant and specifying information seems to be redundant and consequently absent.

For learning to be effective in a continuously changing environment, the ultimate goal of learning should not be the retention of information to deal with specific environmental constraints in the here and now. Rather, it should involve the attunement of perception and action by which children can attend and react to the relevant information and ignore the irrelevant information in varying contexts (Adolph et al., 2009). In this way, newly acquired action-perception couplings function to be adaptive across various contexts (e.g., outside the initial learning context). To this aim, future studies should be directed at finding ways in which children can engage in learning tasks in such a manner that it promotes the attunement of perception and action towards the relevant and specifying information. Within the context of instructional learning, this attunement can perhaps be promoted by making use of the teachers' gaze (see Chapter 2), goal-directed object movements (see Chapter 3) and facial expressions (see Chapter 4). In addition, inducing variability in learning tasks might also be important to facilitate this process of attunement. That is because invariant information can only be detected within the human-environment system when the relation between humans and its environment is variable and continuously changing. For example, infants that manually explore an object can create a high moment-to-moment variability of visual instances of that object (e.g., perceiving an object from different angles). This experience allows them to discover its invariant properties (such as constancy in shape and size) which can subsequently be used to recognize and name the object unequivocally in future instances. In this manner, variability of visual instances of objects can foster language development (Slone, Smith, & Yu, 2019; for many similar examples, see Dixon & Bangert, 2002; Gibson, 1979; de Weerth & van Geert, 2002).

Bolstering perceptual learning by inducing variability.

Taking the idea seriously that variability in the human-environment system can lead to the detection of invariants and with that the discovery of new affordances, it seems that strategically inducing variability in the system (also referred to as 'perturbation', see Steenbeek & van Geert, 2020) could foster the learning process in terms of detecting invariant information. Research into inducing variability in this manner as to foster learning processes has just started to take place (e.g., Huet et al., 2011). For example, Stephen, Dixon, and Isenhower (2009) presented participants with items displaying gear systems (as used in Chapter 5) that moved unpredictably across the screen, forcing participants that used the force-tracing strategy to reorient their motions to the new location to continue tracing the force across the system. In comparison to a condition with stationary depictions of the gear systems, participants discovered that the directional movements of the gears along the gear system alternate much earlier. This perturbation of the relation between participants and the task at hand efficiently led participants to new invariant information (i.e., alternation) that they could use to apply a new strategy (i.e., classification). Ultimately, the idea of inducing variability as to enhance learning processes might be very useful in designing personalized educational activities (such as the 'Rekentuin', see van der Ven, van der Maas, Straatemeier, & Jansen, 2013; also see Steenbeek & van Geert, 2020). To this end, the Dynamic Overlapping Waves Model as presented in Chapter 5 can be used in order to elucidate when children are likely to profit from these kinds of perturbations by revealing when their behaviour is multi-stable (demarcated by the overlap in waves, see Figure 5.4) and therefore likely to transit towards a new skill or strategy. However, more research is required in order to find out which aspects of a learning task need to vary as to optimally support children's perceptual attunement.

Epilogue: Moving forwards with an ecological view on behaviour and development

The proficiency of infants to interact socially and to learn from social situations are often interpreted in a rich way by assuming that infants interpret the behaviour of others as intentional (see Chapter 2 and Chapter 3) and that they heavily rely on mental constructs in assessing social information (see Chapter 4). This manner of explaining of how infants perceive the behaviour of others reflects the normative cognitivist view on human behaviour and development (see Chapter 1). This view seems to be deeply rooted in our sociocultural practices, considering the tendency of humans in everyday life to predict and explain the behaviours of others by offering descriptions of their mental states as causing them (e.g., Tomasello, 1999). However, this tendency does not justify the idea that infants actively seek a causal connection between the mental state of others and their behaviour, as was contented in Chapter 2 and Chapter 3 (also see Haith, 1998). Instead, these explanations rather reflect the way in which scientists tend to adopt a sociocultural context in explaining of infant behaviour (Churchland, 1981). This is nicely illustrated in the following quote in which Meltzoff describes the aim of one of his experiments: "The re-enactment procedure uses infants' nonverbal reconstructions of events to investigate *their* interpretive structures, here to explore *their* folk psychological framework" (Meltzoff, 1995, p. 845, italics added). William James warned psychological scientists about this fallacy over a century earlier: "The great snare of the psychologist *is the confusion of his own standpoint with that of the mental fact* about which he is making his report. I shall hereafter call this the 'psychologist's fallacy' *par excellence*" (James, 1890, p. 196, original italics). In other words, the mentalistic nature infants (and humans in general) tend to display when interacting socially with others might merely exist in the eye of the beholder, that is, the psychologist, and *not* the infant.

In this dissertation an attempt has been made to refute the *necessity* of inferring mental representational abilities in causally affecting behaviour and development (see Chapters 2, 3, and 4). However, it will remain impossible to settle the dispute as to the precise role mental representations (e.g., attitudes, concepts, plans) play in human behaviour on empirical grounds. Neither their existence nor their inexistence can be shown definitively through empirical science, because mental representations are non-physical and non-observable by definition. Therefore, "*If* its very object of study is, by definition, beyond the realm of science, then psychology becomes a paradox: the science of the 'unscientific.'" (Costall, 2011, p. 247). How then, is it possible that mental representations are so much a part of the psychological reality that cognitive psychologists aim to study? It seems that this is due to the act of reification in which inferences or abstractions of the observed reality are treated as if they are a part of the observed reality (van Dijk, 2016). For example, and in relation to Chapter 2, Senju and Csibra (2008) wrote: "We found that 6-month-old infants followed the adult's gaze (a potential communicative-referential signal) toward an object only when such an act is preceded by ostensive cues such as direct gaze" (Senju & Csibra, 2008, p. 668). In their study, they *infer* that the empirically observable 'direct gaze' operates as an 'ostensive cue', but they seem to describe ostensive cues as a part of the empirical reality that can be observed by the infants in their study. This type of fallacy seems omnipresent within the cognitive view (see Hasselman, Seevinck, & Cox, 2019) which turns an inference such as the existence of an internal mental representation into an invisible yet material ghost that taunts the scientific empirical endeavour of social sciences. On logical grounds, the mental realm cannot be refuted from the ecological view based on the argument known as incommensurability (Kuhn, 1962). More specifically, refuting the existence of mental representations requires at least a provisional definition and acceptance of it for which there are no logical foundations in an ecological view, rendering its definition incoherent to argue for and against it. Instead, the ecological view mostly gains territory within the field of psychology by being progressive in the Lakatosian sense. That is, by its ability to generate novel predictions and descriptions of human behaviour and development that can be tested empirically and advance our understanding of it in a parsimonious way (see Chemero, 2011; Ketelaar & Ellis, 2000).

With the aim of parsimony, theoretical considerations in the current dissertation that led to the study predictions of the chapters were mainly founded on two assumptions from the ecological view that were addressed in Chapter 1. First, humans are endowed with the ability to detect invariant structures in the sensory array. Second, humans have a tendency to actively explore their environment. Together, these assumptions were used to formulate a perceptual learning account in which learning and development was described as a process in which perception is increasingly attuned to invariant information within the environment through which affordances surface for adaptive action to occur. In the current dissertation, this account was used to describe a wide scope of behaviours. These included infant gaze following (Chapter 2) and imitation (Chapter 3), children's perception of facial expressions (Chapter 4) and children's understanding of physical mechanisms (Chapter 5). Based on the results of the studies presented in this chapter, several novel research predictions and practical implications were advanced in Chapter 6, showing that the ecological view on human behaviour and development has much to offer.

References

Abrahamson, D., & Lindgren, R. (2014). Embodiment and embodied design. *The Cambridge Handbook of the Learning Sciences*, *2*, 358-376. https://doi.org/10.1017/CBO9781139519526

Adam, J. J., Buetti, S., & Kerzel, D. (2012). Coordinated flexibility: How initial gaze position modulates eye-hand coordination and reaching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 891. https://doi.org/10.1037/a0027592

Adolph, K. E. (2019). An ecological approach to learning in (not and) development. *Human Development*, *63*(3-4), 180-201. https://doi.org/10.1159/000503823

Adolph, K. E., & Berger, S. A. (2007). Motor development. In W. Damon & R. Lerner (Series Eds.), D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol 2: Cognition, perception, and language* (6th ed., pp. 161-213). New York: Wiley. https://doi.org/10.1002/9780470147658.chpsy0204

Adolph, K. E., & Kretch, K. S. (2015). Gibson's theory of perceptual learning. *International encyclopedia of the social and behavioral sciences*, *10*, 127-134. https://doi.org/10.1016/b978-0-08-097086-8.23096-1

Adolph, K. E., Joh, A. S., Franchak, J. M., Ishak, S., & Gill, S. V. (2008). Flexibility in the development of action. In J. Bargh, P. Gollwitzer, & E. Morsella (Eds.), *Oxford Handbook of Human Action*. New York: Oxford University Press.

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley. https://doi.org/10.1002/0471249688

Agresti, A., & Coull, B. A. (1996). Order-restricted tests for stratified comparisons of binomial proportions. *Biometrics*, 1103-1111. https://doi.org/10.2307/2533072

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous gestures influences strategy choices in problem solving. *Psychological Science, 22,* 1138-1144. https://doi.org/10.1177/0956797611417722

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington DC: American Psychiatric Association.

Ardizzi, M., Martini, F., Umiltà, M. A., Evangelista, V., Ravera, R., & Gallese, V. (2015). Impact of Childhood Maltreatment on the Recognition of Facial Expressions of Emotions. *PLOS ONE*, *10*(10), e0141732. https://doi.org/10.1371/journal.pone.0141732

Astor, K., Lindskog, M., Forssman, L., Kenward, B., Fransson, M., Skalkidou, A., … Gredebäck, G. (2020). Social and emotional contexts predict the development of gaze following in early infancy. *Royal Society open science*, *7*(9), 201178. https://doi.org/10.1098/rsos.201178

Atkinson, J., Hood, B., Wattam-Bell, J., & Braddick, O. (1992). Changes in infants' ability to switch attention in the first three months of life. *Perception, 21*, 643-653. https://doi.org/10.1068/p210643

Atkinson, R. C., & Shiffrin, R.M. (1968). Human memory: a proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829-839. https://doi.org/10.1038/nrn1201

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47-89. https://doi.org/10.1016/S0079-7421(08)60452-1

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders*, *40*(3), 358-370. https://doi.org/10.1007/s10803-009-0884-3

Bar-Yam, Y. (2004). A mathematical theory of strong emergence using multiscale variety. *Complexity*, *9*(6), 15-24. https://doi.org/10.1002/cplx.20029

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology, 13*(4), 379-398. https://doi.org/10.1111/j.2044-835X.1995.tb00687.x

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20-46. https://doi.org/10.1207/s15327957pspr1001_2

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior & Development, 23*(2), 223-229. https://doi.org/10.1016/S0163-6383(01)00037-6

Bauchau, V. (2006). Emergence and reductionism: from the game of life to science of life. In *Self-organization and emergence in life sciences* (pp. 29-40). Dordrecht: Springer.

Begon, M., Harper, J. L., & Townsend, C. R. (1996). *Ecology. Individuals, populations and communities* (3rd ed.). Cambridge, MA: Blackwell scientific publications. https://doi.org/10.4236/eng.2011.34046

Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: Comparing 12-and 18-month-olds. *Infant Behavior and Development*, *22*, 277-282. https://doi.org/10.1016/S0163-6383(99)00002-8

Belland, B. R. (2017). *Instructional scaffolding in STEM education: Strategies and efficacy evidence* (p. 144). Springer Nature.

Beris, A. N., & Giacomin, A. J. (2014). πάντα ῥεϊ: Everything flows. *Applied Rheology*, *24*(5), 11-23. https://doi.org/10.3933/applrheol-24-52918

Bimler, D., & Kirkland, J. (2001). Categorical perception of facial expressions of emotion: Evidence from multidimensional scaling. *Cognition and Emotion*, *15*(5), 633-658. https://doi.org/10.1080/02699930143000077

Biro, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science*, *10*(3), 379-398. https://doi.org/10.1111/j.1467-7687.2006.00544.x

Blair, R. J. R., & Coles, M. (2000). Expression recognition and behavioural problems in early adolescence. *Cognitive Development, 15*(4), 421-434. https://doi.org/10.1016/s0885-2014(01)00039-9

Bolger, M. S., Kobiela, M., Weinberg, P. J., & Lehrer, R. (2010). Embodied experiences within an engineering curriculum. *Proceedings of the 9th International Conference of the Learning Sciences (ICLS, 2010), 1,* 706-713. International Society of the Learning Sciences: Chicago, IL.

Bollen, K. A., & Curran, P. J. (2006). *Latent Curve Models.* Hoboken, NJ: Wiley. https://doi.org/10.1002/0471746096

Boncoddo, R., Dixon, J. A., & Kelley, E. (2010). The emergence of a novel representation from action: Evidence from preschoolers. *Developmental Science, 13*, 370-377. https://doi.org/10.1111/j.1467-7687.2009.00905.x

Boom, J. (2009). Piaget on equilibration. In U. Müller, J. I. M. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 132-149). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CCOL9780521898584.006

Boom, J. (2015). A new visualization and conceptualization of categorical longitudinal development: measurement invariance and change. *Frontiers in Psychology*, *6*, *289*. https://doi.org/10.3389/fpsyg.2015.00289

Boom, J., & ter Laak, J. (2007). Classes in the balance: Latent class analysis and the balance scale task. *Developmental Review, 27*, 127-149. https://doi.org/10.1016/j.dr.2006.06.001

Bowlby, J. (1958). The nature of the child's tie to his mother. *International Journal of Psychoanalysis, 39*, 350-371.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*, 535-543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38*(6), 958-966. https://doi.org/10.1037/0012-1649.38.6.958

Bruineberg, J., Chemero, A., & Rietveld, E. (2019). General ecological information supports engagement with affordances for 'higher' cognition. *Synthese*, *196*(12), 5231-5251. https://doi.org/10.1146/10.1007/s11229-018-1716-9

Bruinsma, Y., Koegel, R., & Koegel, L. K. (2004). Joint attention and children with autism: A review of the literature. *Mental Retardation and Developmental Disabilities Reviews, 10*, 169-175. https://doi.org/10.1002/mrdd.20036

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9*(1), 55-72. https://doi.org/10.1111/j.2044-835X.1991.tb00862.x

Calvo, M. G., & Marrero, H. (2009). Visual search of emotional faces: The role of affective content and featural distinctiveness. *Cognition and Emotion*, *23*(4), 782-806. https://doi.org/10.1080/02699930802151654

Camras, L. A. (2000). Surprise! Facial Expressions Can Be Coordinative Motor Structures. In A. T. W. Freeman, W. J. Freeman, M. D. Lewis, & I. Granic (Eds.), *Emotion, Development, and Self-Organization Dynamic Systems Approaches to Emotional Development* (pp. 100-124). Cambridge, UK: Cambridge University Press.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4).

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology*, *70*(2), 205. https://doi.org/10.1037//0022-3514.70.2.205

Case, P., Tuller, B., Ding, M., & Kelso, J. A. (1995). Evaluation of a dynamical model of speech perception. *Perception & Psychophysics, 57*(7), 977-988. https://doi.org/10.3758/bf03205457

Charman, T., & Huang, C. T. (2002). Delineating the role of stimulus enhancement and emulation learning in the behavioural re-enactment paradigm. *Developmental Science*, *5*(1), 25-27. https://doi.org/10.1111/1467-7687.00200

Chemero, A. (2011). *Radical embodied cognitive science*. MIT press.

Chomsky, N. (1980). Rules and representations. *Behavioral and brain sciences*, *3*(1), 1-15. https://doi.org/10.1017/S0140525X00001515

Chow, J. Y., Davids, K., Hristovski, R., Araujo, D., & Passos, P. (2011). Nonlinear pedagogy: Learning design for self-organising neurobiological systems. *New Ideas in Psychology*, *29*, 189-200. https://doi.org/10.1016/j.newideapsych.2010.10.001

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy, 78,* 67-90. https://doi.org/10.5840/jphil198178268

Cong, Y. Q., Junge, C., Aktar, E., Raijmakers, M., Franklin, A., & Sauter, D. (2019). Preverbal infants perceive emotional facial expressions categorically. *Cognition and Emotion*, *33*(3), 391-403. https://doi.org/10.1080/02699931.2018.1455640

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61-83). Hillsdale, NJ: Lawrence Erlbaum.

Costall, A. (1995). Socializing affordances. *Theory & Psychology*, *5*(4), 467-481. https://doi.org/10.1177/0959354395054001

Costall, A. (2011). Against representationalism: James Gibson's secret intellectual debt to E. B. Holt. In E. P. Charles (Ed.), *A new look at New Realism: The psychology and philosophy of E. B. Holt* (pp. 243-261). Transaction Publishers.

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, *11*(2), 239-264. https://doi.org/10.1177/1745691615621279

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25*, 141-168. https://doi.org/10.1111/j.1468-0017.2009.01384.x

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148-153. https://doi.org/10.1016/j.tics.2009.01.005

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148-153. https://doi.org/10.1016/j.tics.2009.01.005

Danish, D., & Russell, J. (2007). The role of 'action-effects' and agency in toddlers' imitation. *Cognitive Development*, *22*, 69-76. https://doi.org/10.1016/j.cogdev.2006.08.001

Danovitch, J., & Bloom, P. (2009). Children's extension of disgust to physical and moral events. *Emotion, 9,* 107-112. https://doi.org/10.1037/a0014113

Darwin, C. (1872). *The expression of the emotions in man and animals.* Oxford, UK: Oxford University Press. https://doi.org/10.1037/10001-000

de Bordes, P. F., Boom, J., Schot, W. D., van den Heuvel-Panhuizen, M., & Leseman, P. P. (2019). Modelling children's Gear task strategy use with the Dynamic Overlapping Waves Model. *Cognitive Development*, *50*, 237-247. https://doi.org/10.1016/j.cogdev.2019.01.001

de Bordes, P. F., Cox, R. F.A., Hasselman, F., & Cillessen, A. H. (2013). Toddlers' gaze following through attention modulation: Intention is in the eye of the beholder. *Journal of Experimental Child Psychology*, *116*, 443-452. https://doi.org/10.1016/j.jecp.2012.09.008

de Bordes, P. F., Hasselman, F., & Cox, R. F. (2019). Attunement and affordance learning in infants. *Journal of Cognition and Development*, *20*(4), 534-554. https://doi.org/10.1080/15248372.2019.1626398

de Bordes, P. F., Hasselman, F., & Cox, R. F. (Submitted). Children's Perception of Facial Expressions. *Child Development.*

de Jonge-Hoekstra, L., van der Steen, S., & Cox, R. F. (2020). Movers and shakers: Hand movements, speech, task properties, and variability. *Acta Psychologica*, *211*, 103187. https://doi.org/10.31234/osf.io/tfu47

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology*, *24*(2), 200-213. https://doi.org/10.1076/jcen.24.2.200.989

de Weerth, C., & van Geert, P. (2002). A longitudinal study of basal cortisol in infants: Intra-individual variability, circadian rhythm and developmental trends. *Infant Behavior & Development, 25*(4), 375-398. https://doi.org/10.1016/S0163-6383(02)00141-8

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology, 36*(4), 511-523. https://doi.org/10.1037/0012-1649.36.4.511

Deák, G. O., Krasno, A. M., Jasso, H., & Triesch, J. (2018). What Leads To Shared Attention? Maternal Cues and Infant Responses During Object Play. *Infancy*, *23*(1), 4-28. https://doi.org/10.1111/infa.12204

Dejonckheere, P. J., Smitsman, A. W., & Verhofstadt-Denève, L. (2007). Occluding the trajectory of an object: Effects on predicting containment with 9-, 12- and 16-month-olds. *Infant Behavior and Development*, *30*(3), 523-528. https://doi.org/10.1016/j.infbeh.2006.12.002

Del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2019). The Developmental Origins of Gaze-Following in Human Infants. *Infancy*, *24*(3), 433-454. https://doi.org/10.1111/infa.12276

Demetriou, A., Mouyi, A., & Spanoudis, G. (2010). The development of mental processing. In R. M. Lerner (Ed.), *Biology, cognition and methods across the life-span. Volume 1 of the Handbook of life-span development* (pp. 36-55). Hoboken, New Jersey: Wiley.

Den Hartigh, R. J. R., Cox, R. F. A., & Van Geert, P. (2017). Complex versus complicated models of cognition. In L. Magnani & T. Bertolotti, (Eds.), *Springer Handbook of Model-Based Science* (pp. 657-669). Cham, Switserland: Springer International Publishing.

Dewey, J., & Bentley, A. F. (1946). Transactions as known and named. *The Journal of Philosophy*, *43*(20), 533-551. https://doi.org/10.2307/2019334

Dixon, J. A., & Bangert, A. S. (2002). The prehistory of discovery: Precursors of representational change in solving gear-systems problems. *Developmental Psychology, 38*, 918-933. https://doi.org/10.1037/0012-1649.38.6.918

Dixon, J. A., & Bangert, A. S. (2004). On the spontaneous discovery of a mathematical relation during problem solving. *Cognitive Science, 28*, 433-449. https://doi.org/10.1207/s15516709cog2803_6

Dixon, J. A., & Dohn, M. C. (2003). Redescription disembeds relations: Evidence from relational transfer and use in problem solving. *Memory & Cognition, 31*, 1082-1093. https://doi.org/10.3758/BF03196129

Dixon, J. A., & Kelley, E. (2007). Theory revision and redescription: Complementary processes in knowledge acquisition. *Current Directions in Psychological Science, 16*, 111-115. https://doi.org/10.1111/j.1467-8721.2007.00486.x

Dixon, J. A., Holden, J. G., Mirman, D., & Stephen, D. G. (2012). Multifractal dynamics in the emergence of cognitive structure. *Topics in Cognitive Science, 4*, 51-62. https://doi.org/10.1111/j.1756-8765.2011.01162.x

Dixon, J. A., Stephen, D. G., Boncoddo, R., & Anastas, J. (2010). The self-organization of cognitive structure. *Psychology of Learning and Motivation, 52*, 343-384. https://doi.org/10.1016/S0079-7421(10)52009-7

Doherty, M. J. (2006). The development of mentalistic gaze understanding. *Infant and Child Development*, *15*(2), 179-186. https://doi.org/10.1002/icd.434

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, *11*(6), 467-473. https://doi.org/10.1111/1467-9280.00290

Durand, K., Gallay, M., Seigneuric, A., Robichon, F., & Baudouin, J. Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of Experimental Child Psychology*, *97*(1), 14-27. https://doi.org/10.1016/j.jecp.2006.12.001

Ekman, P. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: Times Books.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124. https://doi.org/10.1037/h0030377

Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists.* Mahwah, NJ: Erlbaum.

Evans, C. A., & Porter, C. L. (2009). The emergence of mother–infant co-regulation during the first year: Links to infants' developmental status and attachment. *Infant Behavior and Development*, *32*(2), 147-158. https://doi.org/10.1016/j.infbeh.2008.12.005

Fajen, B. R. (2008). Perceptual learning and the visual control of braking. *Perception & Psychophysics*, *70*(6), 1117-1129. https://doi.org/10.3758/PP.70.6.1117

Fajen, B. R., Riley, M. A., & Turvey, M. T. (2008). Information, affordances, and the control of action in sport. *International Journal of Sport Psychology*, *40*(1), 79-107.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Visual eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 9602-9605. https://doi.org/10.1073/pnas.152159999

Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition, 7*(6), 705-718. https://doi.org/10.1080/13506280050144399

Farroni, T., Mansfield, E. M., Lai, C., & Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology, 85*(3), 199-212. https://doi.org/10.1016/S0022-0965(03)00022-5

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307-314. https://doi.org/10.1016/j.tics.2004.05.002

Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action, thought, and emotion. In W. Damon & R. M. Lerner (Eds.), *Theoretical models of human development. Handbook of child psychology* (6th ed., Vol. 1, pp. 313-399). New York: Wiley.

Flom, R., & Pick, A. D. (2005). Experimenter Affective Expression and Gaze Following in 7-Month-Olds. *Infancy, 7*(2), 207-218. https://doi.org/10.1207/s15327078in0702_5

Flom, R., Lee, K., & Muir, D. (2006). *Gaze-following: Its development and significance*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Fodor, J. A. (1985). Fodor's guide to mental representation: The intelligent auntie's vademecum. *Mind*, *94*(373), 76-100. https://doi.org/10.1093/mind/XCIV.373.76

Fogel, A. (1992). Co-regulation, perception and action. *Human Movement Science*, *11*(1), 505-523. https://doi.org/10.1016/0167-9457(92)90031-6

Frith, U., 2013. Are there innate mechanisms that make us social beings? In A. Battro, S. Dehaene, & W. Singer (Eds.), *Neurosciences and the Human Person: New Perspectives on Human Activities* (pp. 215-236). Vatican City: Pontifical Academy of Sciences.

Fritz, J., Bisenberger, A., & Kotrschal, K. (2000). Stimulus enhancement in greylag geese: socially mediated learning of an operant task. *Animal Behaviour*, *59*(6), 1119-1125. https://doi.org/10.1006/anbe.2000.1424

Galef, B. G. (2013). Imitation and local enhancement: detrimental effects of consensus definitions on analyses of social learning in animals. *Behavioural Processes*, *100*, 123-130. https://doi.org/0.1016/j.beproc.2013.07.026

Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology*, *102*(4), 503-521. https://doi.org/10.1016/j.jecp.2008.11.002

Gao, X., & Maurer, D. (2010). A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities. *Journal of Experimental Child Psychology*, *107*(2), 67-86. https://doi.org/10.1016/j.jecp.2010.05.003

Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, *22*(3), 611-631. https://doi.org/10.1017/s0305000900009971

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, *415*(6873), 755-755. https://doi.org/10.1038/415755a

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, *39*(1), 1-42. https://doi.org/10.1146/annurev.ps.39.020188.000245

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. New York: Oxford University Press.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin. https://doi.org/10.1002/bs.3830260313

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, *62*(1), 32. https://doi.org/10.1037/h0048826

Goldin-Meadow, S. (2007). Pointing sets the stage for learning language—and creating language. *Child development*, *78*(3), 741-745. https://doi.org/0.1111/j.1467-8624.2007.01029.x

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515-523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*(1), 585-612. https://doi.org/10.1146/annurev.psych.49.1.585

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69-78. https://doi.org/10.1002/wcs.26

Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, *130*(1), 116. https://doi.org/10.1037/0096-3445.130.1.116

Goren, C., Sarty, M., & Wu, P. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 4*, 544-549.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*(5), 462-466. https://doi.org/10.1111/1467-9280.02454

Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: a critical test of natural pedagogy. *Child Development*, *89*(6), 2091-2098. https://doi.org/10.1111/cdev.13026

Gredebäck, G., Örnkloo, H., & Von Hofsten, C. (2006). The development of reactive saccade latencies. *Experimental Brain Research, 173*, 159-164. https://doi.org/10.1007/s00221-006-0376-z

Gredebäck, G., Theuring, C., Hauf, P., & Kenward, B. (2008). The microstructure of infants' gaze as they view adult shifts in overt attention. *Infancy, 13*(5), 533-543. https://doi.org/10.1080/15250000802329529

Grossmann, T., Parise, E., & Friederici, A. (2010). The detection of communicative signals directed at the self in infant prefrontal cortex. *Frontiers in Human Neuroscience, 4*, 1-6. https://doi.org/10.3389/fnhum.2010.00201

Guevara, M., Cox, R. F., van Dijk, M., & van Geert, P. (2017). Attractor dynamics of dyadic interaction: A recurrence-based analysis. *Nonlinear Dynamics Psychology and Life Sciences*, *21*(3), 289-317.

Hainline, L. (1998). The development of basic visual abilities. In A. Slater (Ed.), *Perceptual development: Visual, auditory, and speech perception in infancy* (pp. 5-50). Hove, UK: Psychology Press. https://doi.org/10.1002/9781444327564.ch2

Hains, S., & Muir, D. (1996). Infant sensitivity to adult eye direction. *Child Development, 67*, 1940-1951. https://doi.org/10.1111/j.1467-8624.1996.tb01836.x

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, *21*, 167-179. https://doi.org/10.1016/S0163-6383(98)90001-7

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science, 198*(4319), 853-855. https://doi.org/10.1126/science.918670

Hasselman, F. (2015). Classifying acoustic signals into phoneme categories: average and dyslexic readers make use of complex dynamical patterns and multifractal scaling properties of the speech signal. *Peerj, 3*(3), e837. https://doi.org/10.7717/peerj.837

Hasselman, F., Seevinck, M., & Cox, R. (2019, July 17). "So you confirmed, replicated and emptied your file-drawer… now what?" A Structural Realist's Guide to Theory Evaluation in Psychological Science. https://doi.org/10.31234/osf.io/b8csj

Hellendoorn, A. (2014). Understanding social engagement in autism: being different in perceiving and sharing affordances. *Frontiers in Psychology*, *5*, 850. https://doi.org/10.3389/fpsyg.2014.00850

Hellendoorn, A., Langstraat, I., Wijnroks, L., Buitelaar, J. K., van Daalen, E., & Leseman, P. P. (2014). The relationship between atypical visual processing and social skills in young children with autism. *Research in Developmental Disabilities*, *35*(2), 423-428. https://doi.org/10.3389/fpsyg.2014.00850

Hellendoorn, A., Wijnroks, L., & Leseman, P. P. (2015). Unraveling the nature of autism: Finding order amid change. *Frontiers in Psychology*, *6*, 359. https://doi.org/10.3389/fpsyg.2015.00359

Herba, C. M., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry*, *45*(7), 1185-1198. https://doi.org/10.1111/j.1469-7610.2004.00316.x

Herba, C. M., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: Effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry*, 47, 1098-1106. https://doi.org/10.1111/j.1469-7610.2006.01652.x

Heyes, C. (2012). What's social about social learning? *Journal of Comparative Psychology*, *126*(2), 193.

Heyes, C. (2016). Born pupils? Natural pedagogy and cultural pedagogy. *Perspectives on Psychological Science*, *11*(2), 280-295. https://doi.org/10.1177/1745691615621276

Heyes, C. M., Ray, E. D., Mitchell, C. J., & Nokes, T. (2000). Stimulus enhancement: controls for social facilitation and local enhancement. *Learning and Motivation*, *31*(2), 83-98. https://doi.org/10.1006/lmot.1999.1041

Hoehl, S., Reid, V., Parise, E., Handl, A., Palumbo, L., & Striano, T. (2009). Looking at eye gaze processing and its neural correlates in infancy: Implications for social development and autism spectrum disorder. *Child Development, 4*, 968-985. https://doi.org/10.1111/j.1467-8624.2009.01311.x

Hoffman, D. D. (2000). *Visual intelligence: How we create what we see*. WW Norton & Company.

Holt, E. B. (1914). *The Concept of Consciousness*. London: George Allen & Co.

Honey, R. C., Close, J., & Lin, T. E. (2010). Acquired distinctiveness and equivalence: A synthesis. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 159-186). Oxford, UK: Oxford University Press

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science, 9*, 53-56. https://doi.org/10.1111/1467-9280.00024

Hopper, L. M. (2010). 'Ghost' experiments and the dissection of social learning in humans and animals. *Biological Reviews*, *85*(4), 685-701. https://doi.org/10.1111/j.1469-185X.2010.00120.x.

Hoppitt, W., & Laland, K. N. (2013). *Social learning: an introduction to mechanisms, methods, and models*. Princeton University Press.

Horne, P. J., Erjavec, M., & Lovett, V. E. (2009). The effects of modelling, local stimulus enhancement, and affordance demonstration on the production of object-directed actions in 6-month-old infants. *British Journal of Developmental Psychology*, *27*(2), 269-281. https://doi.org/10.1348/026151008X297126

Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson Wadsworth.

Huang, C. T., & Charman, T. (2005). Gradations of emulation learning in infants' imitation of actions on objects. *Journal of Experimental Child Psychology*, *92*, 276-302. https://doi.org/10.1016/j.jecp.2005.06.003

Huang, C. T., Heyes, C., & Charman, T. (2002). Infants' behavioral reenactment of "failed attempts": exploring the roles of emulation learning, stimulus enhancement, and understanding of intentions. *Developmental Psychology*, *38*, 840. https://doi.org/10.1037/0012-1649.38.5.840

Huet, M., Jacobs, D. M., Camachon, C., Missenard, O., Gray, R., & Montagne, G. (2011). The education of attention as explanation of variability of practice effects: Learning the final approach phase in a flight simulator. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1841. https://doi.org/10.1037/a0024386

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130490. https://doi.org/10.1111/desc.12873

Ingold, T. (2001). From the transmission of representation to the education of attention. In H. Whitehouse (Ed.), *The debated mind: Evolutionary psychology versus ethnography* (pp. 113-153). New York, NY, US: Berg.

Ishak, S., Franchak, J. M., & Adolph, K. E. (2014). Perception–action development from infants to adults: Perceiving affordances for reaching through openings. *Journal of Experimental Child Psychology*, *117*, 92-105. https://doi.org/10.1016/j.jecp.2013.09.003

Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, *109*(19), 7241-7244. https://doi.org/10.1073/pnas.1200155109

Jacobs, D. M., & Michaels, C. F. (2007). Direct learning. *Ecological Psychology*, *19*(4), 321-349. https://doi.org/10.1080/10407410701432337

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Henry Holt and Company.

James, W. (1895). The knowing of things together. *Psychology Review*, *2*, 105-124. https://doi.org/10.1037/h0073221

Johnson, S. C., Booth, A., & O'Hearn, K. (2001). Inferring the goals of a nonhuman agent. *Cognitive Development*, *16*, 637-656. https://doi.org/10.1016/S0885-2014(01)00043-0

Jones, B., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry, 65*, 946-954. https://doi.org/10.1001/archpsyc.65.8.946

Kellman, P. J., & Massey, C. M. (2013). Perceptual learning, cognition, and expertise. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 58, pp. 117-165). Academic Press. https://doi.org/10.1016/B978-0-12-407237-4.00004-9.

Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., & Caldara, R. (2011). Developing cultural differences in face processing. *Developmental Science*, *14*(5), 1176-1184. https://doi.org/10.1111/j.1467-7687.2011.01067.x

Kessels, R. P. C., Montagne, B., Hendriks, A. W., Perrett, D. I., & de Haan, E. H. F. (2013). Assessment of perception of morphed facial expressions using the Emotion Recognition Task: Normative data from healthy participants aged 8-75. *Journal of Neuropsychology, 8*(1), 75-93. https://doi.org/10.1111/jnp.12009

Ketelaar, T., & Ellis, B. J. (2000). Are evolutionary explanations unfalsifiable? Evolutionary psychology and the Lakatosian philosophy of science. *Psychological Inquiry*, *11*(1), 1-21. https://doi.org/10.1207/S15327965PLI1101_01

Kibler, J. (2011) Cognitive Schemas. In S. Goldstein S & J. A. Naglieri (Eds.), *Encyclopedia of Child Behavior and Development*. Boston, MA: Springer. https://doi.org/10.1007/978-0-387-79061-9_608

Király, I. (2009). The effect of the model's presence and of negative evidence on infants' selective imitation. *Journal of Experimental Child Psychology*, *102*, 14-25. https://doi.org/10.1016/j.jecp.2008.06.003

Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2017). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, *21*(2), 142-154. https://doi.org/10.1177/1362361316632710

Klin, A., & Jones, W. (2008). Altered face scanning and impaired recognition of biological motion in a 15-month-old infant with autism. *Developmental Science, 11*(1), 40-46. https://doi.org/10.1111/j.1467-7687.2007.00608.x

Klinnert, M., Campos, J., Sorce, J., Emde, R., & Svejda, M. (1983). Emotions as behavior regulators: social referencing in infancy. In R. Plutchnik & H. Kellerman (Eds.), *Emotion in early development* (Vol. 2, pp. 57-86) New York: Academic Press https://doi.org/10.1016/B978-0-12-558702-0.50009-1

Konczak, J., Meeuwsen, H. J., & Cress, M. E. (1992). Changing affordances in stair climbing: The perception of maximum climbability in young and older adults. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 691. https://doi.org/10.1037/0096-1523.18.3.691

Krys, K., Vauclair, C.-M., Capaldi, C. A., Lun, V. M.-C., Bond, M. H., Domínguez-Espinosa, A., … Yu, A. A. (2015). Be Careful Where You Smile: Culture Shapes Judgments of Intelligence and Honesty of Smiling Individuals. *Journal of Nonverbal Behavior, 40*(2), 101-116. https://doi.org/10.1007/s10919-015-0226-4

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174. https://doi.org/10.2307/2529310

Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, *6*, 761. https://doi.org/10.3389/fpsyg.2015.00761

Lee, V., Cheal, J. L., & Rutherford, M. D. (2015). Categorical perception along the happy–angry and happy–sad continua in the first year of life. *Infant Behavior and Development*, *40*, 95-102. https://doi.org/10.1016/j.infbeh.2015.04.006

Lehrer, R., & Schaube, L. (1998). Reasoning about structure and function: Children's conceptions of gears. *Journal of Research in Science Teaching, 35*, 3-25. https://doi.org/10.1002/(SICI)1098-2736(199801)35:1%3c3::AID-TEA2%3e3.0.CO;2-X

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, *10*, 37-47. https://doi.org/10.1038/nrn2554

Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guildford Press.

Lobo, L., Heras-Escribano, M., & Travieso, D. (2018). The history and philosophy of ecological psychology. *Frontiers in Psychology*, *9*, 2228. https://doi.org/10.3389/fpsyg.2018.02228

Lynn, S. K., & Barrett, L. F. (2014). "Utilizing" signal detection theory. *Psychological Science*, *25*(9), 1663-1673. https://doi.org/10.1177/0956797614541991

Macchi Cassia, V., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science, 15,* 379-384. https://doi.org/10.1111/j.0956-7976.2004.00688.x

MacMillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology Press.

Malmberg, K. J., Raaijmakers, J. G., & Shiffrin, R. M. (2019). 50 years of research sparked by Atkinson and Shiffrin (1968). *Memory & Cognition*, *47*(4), 561-574. https://doi.org/10.3758/s13421-019-00896-7

Marchi, F., & Newen, A. (2015). Cognitive penetrability and emotion recognition in human facial expressions. *Frontiers in Psychology*, *6*, 828. https://doi.org/10.3389/fpsyg.2015.00828

Marr, D. (1976). Analyzing natural images: a computational theory of texture vision. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 40, pp. 647-662). Cold Spring Harbor Laboratory Press.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science*, *29*, 587-625. https://doi.org/10.1207/s15516709cog0000_15

Meagher, B. R., & Marsh, K. L. (2014). Testing an ecological account of spaciousness in real and virtual environments. *Environment and Behavior*, *47*(7), 782-815. https://doi.org/10.1177/0013916514525039

Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838. https://doi.org/10.1037/0012-1649.31.5.838

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental Science*, *10*(1), 126-134. https://doi.org/10.1111/j.1467-7687.2007.00574.x

Michaels, C. F., & Carello, C. (1981). *Direct perception*. Englewood Cliffs, NJ: Prentice-Hall.

Moder, K. (2010). Alternatives to F-test in one-way ANOVA in case of heterogeneity of variances (a simulation study). *Psychological Test and Assessment Modelling*, *52*(4), 343-353.

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science, 7*(1), F1-F9. https://doi.org/10.1111/j.1467-7687.2004.00315.x

Monroy, C., Chen, C. H., Houston, D., & Yu, C. (2020). Action prediction during real-time parent-infant interactions. *Developmental Science*, e13042. https://doi.org/10.1111/desc.13042

Montirosso, R., Peverelli, M., Frigerio, E., Crespi, M., & Borgatti, R. (2010). The Development of Dynamic Facial Expression Recognition at Different Intensities in 4- to 18-Year-Olds. *Social Development, 19*(1), 71-92. https://doi.org/10.1111/j.1467-9507.2008.00527.x

Moore, C., & Povinelli, D. J. (2007). Differences in how 12- and 24-month-olds interpret the gaze of adults. *Infancy*, *11*, 215-231.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98,* 164-181. https://doi.org/0.1037/0033-295x.98.2.164

Mostow, A. J., Izard, C. E., Fine, S., & Trentacosta, C. J. (2002). Modeling emotional, cognitive, and behavioral predictors of peer acceptance. *Child Development*, *73*(6), 1775-1787. https://doi.org/10.1111/1467-8624.00505

Mumme, D. L., & Fernald, A. (2003). The infant as onlooker: Learning from emotional reactions observed in a television scenario. *Child Development, 74*(1), 221-237. https://doi.org/10.1111/1467-8624.00532

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science, 16*(5), 269-274. https://doi.org/10.1111/j.1467-8721.2007.00518.x

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide. Eighth Edition.* Los Angeles, CA: Muthén & Muthén.

Neisser, U. (1963). The multiplicity of thought. *British Journal of Psychology*, *54*(1), 1-14.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Netti, S., & Nusantara, T. (2016). The Failure to Construct Proof Based on Assimilation and Accommodation Framework from Piaget. *International Education Studies*, *9*(12), 12-22. https://doi.org/10.1016/10.5539/ies.v9n12p12

Nielsen, M. (2006). Copying actions and copying outcomes: social learning through the second year. *Developmental Psychology*, *42*, 555. https://doi.org/10.1037/0012-1649.42.3.555

Nomikou, I., Leonardi, G., Radkowska, A., Rączaszek-Leonardi, J., & Rohlfing, K. J. (2017). Taking up an active role: emerging participation in early mother–infant interaction during peekaboo routines. *Frontiers in Psychology*, *8*, 1656. https://doi.org/10.3389/fpsyg.2017.01656

Northrup, J. B., & Iverson, J. M. (2020). Multimodal coordination of vocal and gaze behavior in mother–infant dyads across the first year of life. *Infancy*, *25*(6), 952-972. https://doi.org/10.1111/infa.12369

Ostini, R., & Nering, M. L. (2006). *Polytomous Item Response Theory Models.* Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781412985413

Özgen, E., & Davies, I. R. (2002). Acquisition of categorical color perception: a perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General*, *131*(4), 477. https://doi.org/10.1037/0096-3445.131.4.477

Ozturk, O., Shayan, S., Liszkowski, U., & Majid, A. (2013). Language is not necessary for color categories. *Developmental Science*, *16*(1), 111-115. https://doi.org/10.1111/desc.12008

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. MIT press.

Paulus, M. (2011). How infants relate looker and object: Evidence for a perceptual learning account of gaze following in infancy. *Developmental Science*, *14*(6), 1301-1310. https://doi.org/10.1111/j.1467-7687.2011.01076.x

Paulus, M. (2012). Is it rational to assume that infants imitate rationality? A theoretical analysis and critique. *Human Development, 55*(3), 107-121. https://doi.org/10.1159/000339442

Piaget, J., & Cook, M. T. (1952). *The origins of intelligence in children*. New York, NY: International University Press.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*(2), 293-305. https://doi.org/10.1016/j.neuron.2006.11.022

Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., … Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. *The Lancet*, *388*(10059), 2501-2509. http://doi.org/10.1016/ S0140-6736(16)31229-6

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*(428), 1303-1313. https://doi.org/10.1080/01621459.1994.10476870

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences, 99(13)*, 9072-9076. https://doi.org/10.1073/pnas.142165999

Pollak, S. D., & Tolley-Schell, S. A. (2003). Selective attention to facial emotion in physically abused children. *Journal of Abnormal Psychology*, *112*(3), 323. https://doi.org/10.1037/0021-843X.112.3.323

Pollak, S. D., Messner, M., Kistler, D. J., & Cohn, J. F. (2009). Development of perceptual expertise in emotion recognition. *Cognition*, *110*(2), 242-247. https://doi.org/10.1016/j.cognition.2008.10.010

Pouw, W. T., Van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review*, *26*(1), 51-72. https://doi.org/10.1007/s10648-014-9255-5

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110-122. https://doi.org/10.1016/j.lindif.2009.10.005

Rambusch, J., & Ziemke, T. (2005, July). The role of embodiment in situated learning. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1803-1808). Mahwah, NJ: Lawrence Erlbaum.

Read, C., & Szokolszky, A. (2018). An emerging developmental ecological psychology: future directions and potentials. *Ecological Psychology*, *30*(2), 174-194. https://doi.org/10.1080/10407413.2018.1439141

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*(6), 1045-1060. https://doi.org/10.1207/s15516709cog0000_29

Rodger, H., Lao, J., & Caldara, R. (2018). Quantifying facial expression signal and intensity use during development. *Journal of Experimental Child Psychology*, *174*, 41-59. https://doi.org/10.1016/j.jecp.2018.05.005

Rodger, H., Vizioli, L., Ouyang, X., & Caldara, R. (2015). Mapping the development of facial expression recognition. *Developmental Science*, *18*(6), 926-939. https://doi.org/10.1111/desc.12281

Rodriguez, N. M., & Thompson, R. H. (2015). Behavioral variability and autism spectrum disorder. *Journal of Applied Behavior Analysis*, *48*(1), 167-187. https://doi.org/10.1002/jaba.164

Ruba, A. L., & Repacholi, B. M. (2019). Do Preverbal Infants Understand Discrete Facial Expressions of Emotion? *Emotion Review*, 1-16. https://doi.org/10.1177/1754073919871098

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2019). How Do You Feel? Preverbal Infants Match Negative Emotions to Events. *Developmental Psychology*, *55*(6), 1138-1149. http://dx.doi.org/10.1037/dev0000711

Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental Review*, *34*(3), 265-293. https://doi.org/10.1016/j.dr.2014.04.001

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 795-804. https://doi.org/10.1037/h0077714

Russell, J. A., & Bullock, M. (l986). On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology, 22,* 97-l02. https://doi.org/10.1037/0012-1649.22.1.97

Russell, J. A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *Journal of Experimental Psychology: General*, *116*(3), 223. https://doi.org/10.1037/0096-3445.116.3.223

Ruxton, G. D., & Beauchamp, G. (2008). Time for some a priori thinking about post hoc testing. *Behavioral Ecology*, *19*(3), 690-693. https://doi.org/10.1093/beheco/arn020

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*, 181-203. https://doi.org/10.1146/annurev-psych-122216-011805

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*(5489), 265-266. https://doi.org/10.1038/253265a0

Schneider, W., Eschman, A., & Zuccolotto, A. (2002) *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70-76. https://doi.org/10.1016/j.tics.2005.12.009

Segal, S. C., Reyes, B. N., Gobin, K. C., & Moulson, M. C. (2019). Children's recognition of emotion expressed by own-race versus other-race faces. *Journal of Experimental Child Psychology*, *182*, 102-113. https://doi.org/10.1037/a0014113

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*, 668-671. https://doi.org/10.1016/j.cub.2008.03.059

Senju, A., & Hasegawa, T. (2006). Do the upright eyes have it? *Psychonomic Bulletin and Review, 13*, 223-228. https://doi.org/10.3758/bf03193834

Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences, 13*, 127-134. https://doi.org/10.1016/j.tics.2008.11.009

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition, 12*(8), 1474-1496. https://doi.org/10.1080/13506280444000797

Shepherd, S. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience, 4*, 1-13. https://doi.org/10.3389/fnint.2010.00005

Shneidman, L., Todd, R., & Woodward, A. (2014). Why do child-directed interactions support imitative learning in young children? *PloS One*, *9*(10), e110891. https://doi.org/10.1371/journal.pone.0110891

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*(5), 405-410. http://dx.doi.org/10.1111/1467-9280.00076

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8*, 481-520. https://doi.org/10.1016/0010-0285(76)90016-5

Siegler, R. S. (1996). *Emerging minds. The process of change in children's thinking.* New York: Oxford University Press.

Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Eds.), D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2: Cognition, Perception, and Language* (6th ed., pp. 464-510). https://doi.org/10.1002/9780470147658.chpsy0211

Siegler, R. S., & Alibali, M. W. (2005). *Children's thinking*. Upper Saddle River, N.J: Pearson Education/Prentice Hall.

Siegler, R. S., & Chen, Z. (2002). Development of rules and strategies: Balancing the old and the new. *Journal of Experimental Child Psychology*, *81*(4), 446-457. https://doi.org/10.1006/jecp.2002.2666

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *The Journal of Genetic Psychology: Research and Theory on Human Development, 164*(1), 54-71. https://doi.org/10.1080/00221320309597503

Slone, L. K., Smith, L. B., & Yu, C. (2019). Self-generated variability in object images predicts vocabulary growth. *Developmental Science*, *22*(6), e12816. https://doi.org/e12816.10.1111/desc.12816

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Science, 7*, 343-348. https://doi.org/10.1016/s1364-6613(03)00156-6

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*(4), 325-336. https://doi.org/10.1016/j.tics.2018.02.004

Smith, L., & Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. *Artificial life*, *11*(1-2), 13-29. https://doi.org/10.1162/1064546053278973

Smith, P., Fischer, R., Vignoles, V., & Bond, M. (2013). *Understanding social psychology across cultures*. Engaging with others in a changing world. London: Sage.

Southgate, V., Chevallier, C., & Csibra, G. (2009). Sensitivity to communicative relevance tells young children what to imitate. *Developmental Science*, *12*, 1013-1019. https://doi.org/10.1111/j.1467-7687.2009.00861.x

Spelke, E. S. (1998). Nativism, empiricism, and the origins of knowledge. *Infant Behavior and Development*, *21*, 181-200. http://doi.org/10.1016/S0163-6383(98)90002-9.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89-96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Spencer, J. P., Dineva, E., & Smith, L. B. (2009). Comment on "Infants' perseverative search errors are induced by pragmatic misinterpretation". *Science, 325*, 1624-a. https://doi.org/10.1126/science.1172759

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149. https://doi.org/10.3758/bf03207704

Steenbeek, H., & van Geert, P. (2020). Education and Development as Complex Dynamic Agent Systems: How Theory Informs Methodology. In M. F. Mascolo & T. R. Bidell (Eds.), *Handbook of Integrative Developmental Science: Essays in Honor of Kurt W. Fischer* (pp. 162-188). New York: Routledge. https://doi.org/10.4324/9781003018599

Stephen, D. G., Boncoddo, R. A., Magnuson, J. S., & Dixon, J. A. (2009). The dynamics of insight: Mathematical discovery as a phase transition. *Memory & Cognition, 37*(8), 1132-1149. https://doi.org/10.3758/mc.37.8.1132

Stephen, D. G., Dixon, J. A., & Isenhower, R. W. (2009). Dynamics of representational change: entropy, action, and cognition. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1811-1832. https://doi.org/10.1037/a0014510

Stevens, K. A. (2012). The Vision of David Marr. *Perception, 41*(9), 1061-1072. https://doi.org/10.1068/p7297

Stöckler M. (1991). A Short History of Emergence and Reductionism. In E. Agazzi (Ed.), *The Problem of Reductionism in Science. Episteme (A Series in the Foundational, Methodological, Philosophical, Psychological, Sociological, and Political Aspects of the Sciences, Pure and Applied), Vol. 18*. Dordrecht: Springer. https://doi.org/10.1007/978-94-011-3492-7_4

Szokolszky, A., Read, C., Palatinus, Z., & Palatinus, K. (2019). Ecological approaches to perceptual learning: learning to perceive and perceiving as learning. *Adaptive Behavior*, *27*(6), 363-388. https://doi.org/10.1177/1059712319854687

Szufnarowska, J., Rohlfing, K. J., Fawcett, C., & Gredebäck, G. (2014). Is ostension any more than attention? *Scientific Reports*, *4*, 5304. https://doi.org/10.1038/srep05304

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn and Bacon.

The National Research Council. (2014). *STEM Integration in K–12 Education Status, Prospects, and an Agenda for Research*. The National Academic Press.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, *24*, 1-34. https://doi.org/10.1017/S0140525X01003910

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science, 10*, 547-558. https://doi.org/10.1111/j.1467-7687.2007.00614.x

Thompson, D. E., & Russell, J. (2004). The ghost condition: imitation versus emulation in young children's observational learning. *Developmental Psychology*, *40*(5), 882. https://doi.org/10.1037/0012-1649.40.5.882

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Topál, J., Gergely, G., Miklósi, Á., Erdohegyi, Á., & Csibra, G. (2008). Infants' perseverative search errors are induced by pragmatic misinterpretation. *Science, 321*, 1831-1834. https://doi.org/10.1126/science.1161437

Träuble, B., & Pauen, S. (2007). The role of functional information for infant categorization. *Cognition*, *105*(2), 362-379. https://doi.org/10.1016/j.cognition.2006.10.003

Triesch, J., Teuscher, C., Deák, G. O., & Carlson, E. (2006). Gaze following: Why (not) learn it? *Developmental Science, 9*, 125-157. https://doi.org/10.1111/j.1467-7687.2006.00470.x

Trudeau, J. T., & Dixon, J. A. (2007). Embodiment and abstraction: Actions create relational representations. *Psychonomic Bulletin & Review, 14*, 994-1000. https://doi.org/10.3758/BF03194134

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

UNICEF. (2007). An overview of child well-being in rich countries. Retrieved from http://news.bbc.co.uk/nol/shared/bsp/hi/pdfs/13_02_07_nn_unicef.pdf

van Beek, Y., & Dubas, J. (2008). Age and Gender Differences in Decoding Basic and Non-Basic Facial Expressions in Late Childhood and Early Adolescence. *Journal of Nonverbal Behavior, 32(1)*, 37-52. https://doi.org/10.1007/s10919-007-0040-8

Van der Kamp, J., Oudejans, R., & Savelsbergh, G. (2003). The development and learning of the visual control of movement: An ecological perspective. *Infant Behavior and Development*, *26*(4), 495-515. https://doi.org/10.1016/j.infbeh.2003.09.002

Van Der Steen, S., Steenbeek, H., Van Dijk, M., & Van Geert, P. (2014). A process approach to children's understanding of scientific concepts: A longitudinal case study. *Learning and Individual Differences*, *30*, 84-91. https://doi.org/10.1016/j.lindif.2013.12.004

Van der Ven, S. H., Boom, J., Kroesbergen, E. H., & Leseman, P. P. (2012). Microgenetic patterns of children's multiplication learning: Confirming the overlapping waves model by latent growth modelling. *Journal of Experimental Child Psychology, 113*, 1-19. https://doi.org/10.1016/j.jecp.2012.02.001

van der Ven, S. H., van der Maas, H. L., Straatemeier, M., & Jansen, B. R. (2013). Visuospatial working memory and mathematical ability at different ages throughout primary school. *Learning and Individual Differences*, *27*, 182-192. https://doi.org/10.3758/s13421-014-0480-4

van Dijk, L. (2016). Laying down a path in talking. *Philosophical Psychology*, *29*(7), 993-1003. https://doi.org/10.1080/09515089.2016.1213379

van Dijk, L. (2020). Temporalizing ontology: a case for pragmatic emergence. *Synthese*, 1-14. https://doi.org/10.1007/s11229-020-02615-1

van Dijk, L., & Withagen, R. (2014). The horizontal worldview: A Wittgensteinian attitude towards scientific psychology. *Theory & Psychology*, *24*(1), 3-18. https://doi.org/10.1177/0959354313517415 van Geert, P. (2000). The dynamics of general developmental mechanisms: From Piaget and Vygotsky to dynamic systems models. *Current Directions in Psychological Science*, *9*(2), 64-68. https://doi.org/10.1111/1467-8721.00062

van Geert, P. (2019). Dynamic systems, process and development. *Human Development*, *63*(3-4), 153-179. https://doi.org/10.1159/000503825

Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, *25*, 340-374. https://doi.org/10.1016/S0163-6383(02)00140-6

Vecera, S. P., & Johnson, M. H. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition, 2*, 59–87. https://doi.org/10.1080/13506289508401722

Vicari, S., Reilly, J. S., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expressions of emotions in school-age children: the intersection of perceptual and semantic categories. *Acta Paediatrica*, *89*(7), 836-845. https://doi.org/10.1111/j.1651-2227.2000.tb00392.x

Vink, R., Hasselman, F., Cillessen, A. H., Wijnants, M. L., & Bosman, A. M. (2018). Does competence determine who leads in a dyadic cooperative task? A study of children with and without a neurodevelopmental disorder. *Complexity*, *2018*. https://doi.org/10.1155/2018/5379531

von Hofsten, C. (2007). Action in development. *Developmental Science*, *10*(1), 54-60. https://doi.org/10.1111/j.1467-7687.2007.00564.x

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychological Bulletin*, *121*(3), 437. https://doi.org/10.1037//0033-2909.121.3.437

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception, 6*(3), 313-326. https://doi.org/10.1068/p060313

Wang, J., & Feigenson, L. (2019). Infants recognize counting as numerically relevant. *Developmental Science*, *22*(6), e12805. https://doi.org/10.1111/desc.12805

Want, S. C., & Harris, P. L. (2002). How do children ape? Applying concepts from the study of non-human primates to the developmental study of 'imitation' in children. *Developmental Science*, *5*(1), 1-14. https://doi.org/10.1111/1467-7687.00194

Warren, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 683. https://doi.org/10.1037/0096-1523.10.5.683

Weisberg, S. M., & Newcombe, N. S. (2017). Embodied cognition and STEM learning: overview of a topical collection in CR: PI. *Cognitive Research*, *2*(1), 38. https://doi.org/10.1186/s41235-017-0071-6

Welchman, A. E. (2016). The human brain in depth: how we see in 3D. *Annual Review of Vision Science*, *2*, 345-376. https://doi.org/10.1146/annurev-vision-111815-114605

Widen, S. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, *5*, 72-77. https://doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, *39*, 114-128. https://doi.org/10.1037/0012-1649.39.1.114

Widen, S. C., & Russell, J. A. (2004). The relative power of an emotion's facial expression, label, and behavioral consequence to evoke preschoolers' knowledge of its cause. *Cognitive Development*, *19*(1), 111-125. https://doi.org/10.1016/j.cogdev.2003.11.004

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development, 23*, 291-312. https://doi.org/10.1016/j.cogdev.2008.01.002

Widen, S. C., & Russell, J. A. (2013). Children's recognition of disgust in others. *Psychological Bulletin*, *139*(2), 271. https://doi.org/10.1037/a0031640

Winkler, M., Mueller, J. L., Friederici, A. D., & Männel, C. (2018). Infant cognition includes the potentially human-unique ability to encode embedding. *Science Advances*, *4*(11), eaar8334. https://doi.org/10.1126/sciadv.aar8334

Withagen, R., & Michaels, C. F. (2007). Transfer of calibration between length and sweet-spot perception by dynamic touch. *Ecological Psychology*, *19*(1), 1-19. https://doi.org/10.1080/10407410709336948

Woolfolk, A. E., Hoy, A. W., Hughes, M., & Walkup, V. (2007). *Psychology in education*. Pearson Education.

Yan, Z., & Fischer, K. (2002). Always under construction. *Human Development*, *45*(3), 141-160. https://doi.org/10.1159/000057070

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*(3), 271-313. https://doi.org/10.1016/s0010-0277(97)00003-6

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PloS One*, *8*(11), e79659. https://doi.org/10.1371/journal.pone.0079659

Yu, C., & Smith, L. B. (2017). Multiple Sensory-Motor Pathways Lead to Coordinated Visual Attention. *Cognitive Science*, *41*, 5-31. https://doi.org/10.1111/cogs.12366

Zebrowitz, L. A. (2011). Ecological and social approaches to face perception. In A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 31-50). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199559053.001.0001

Zentall, T. R. (2001). Imitation in animals: evidence, function, and mechanisms. *Cybernetics and Systems*, *32*, 53-96. https://doi.org/10.1080/019697201300001812

Zheng, Y., & Fischer, K. (2002). Always under construction: Dynamic variations in adult cognitive microdevelopment. *Human Development*, *45*, 141-160. https://doi.org/10.1159/000057070

Samenvatting

Het algemene doel van dit proefschrift was om vanuit de ecologische psychologie te begrijpen hoe baby's en kinderen verschillende vaardigheden kunnen ontwikkelen met behulp van perceptueel leren. In hoofdstuk 1 werden de cognitivistische visie en de ecologische visie op menselijk gedrag en de ontwikkeling daarvan geïntroduceerd en gecontrasteerd. In essentie gaat de cognitivistische visie ervan uit dat informatie passief wordt ontvangen via de zintuigen, waar deze vervolgens mentaal wordt verwerkt om er betekenis aan te geven en om gedrag aan te sturen. Vanuit deze visie lijkt de ontwikkeling van vaardigheden enerzijds te bestaan uit een accumulatie en (re)organisatie van kennis binnen het brein en anderzijds een verhoogde capaciteit en snelheid van het brein om dit te doen. Vanuit de ecologische visie wordt verondersteld dat waarnemen een actief en exploratief proces is dat gericht is op het ontdekken van mogelijkheden voor actie waarbij waarnemingen betekenis krijgen in termen van adaptieve actiemogelijkheden waarin ze kunnen voorzien. Vanuit deze visie is de ontwikkeling van vaardigheden het resultaat van perceptueel leren waarin baby's en kinderen in toenemende mate visuele informatie kunnen onderscheiden in hun omgeving. Dit kunnen ze op adaptieve wijze koppelen aan gedrag, leidend tot een vergroting van hun gedragsrepertoire. In tegenstelling tot de cognitivistische visie gaat de ecologische visie niet uit van een scheiding tussen de immateriële mentale wereld en de fysieke wereld waarin men zich bevindt. Op basis van deze discussie werd geconcludeerd dat de ecologische visie een elegant en interessant alternatief is voor de meer gangbare cognitivistische visie in het beschrijven van de ontwikkeling van vaardigheden door kinderen. Met de studies die zijn beschreven in dit proefschrift werd vanuit de ecologische visie perceptueel leren onderzocht met het uitgangspunt dat kinderen sociale en niet-sociale vaardigheden ontwikkelen door middel van continue afstemming van perceptie en actie op de relevante informatie in de omgeving.

Het volgen van kijkrichting

In hoofdstuk 2 werd een studie gerapporteerd waarin baby's van 20 maanden werden onderzocht op de vaardigheid om de kijkrichting van een ander te volgen. In een eerder onderzoek werd beweerd dat baby's oogcontact met een ander interpreteren als een communicatieve intentie van die andere persoon om iets te laten zien, en daarom geneigd zijn om de kijkrichting van die persoon te volgen. Vanuit het idee van perceptueel leren werd de hypothese geformuleerd dat baby's een voorkeur hebben om naar de ogen van frontaal gepositioneerde gezichten te kijken. Als die ogen vervolgens een laterale beweging maken, volgen ze deze richting meestal, wat effectief resulteert in het volgen van kijkrichting. Volgens deze hypothese hoeven baby's niet signalen zoals oogcontact te interpreteren als communicatieve intentie om de kijkrichting te kunnen volgen. Om dit idee te onderzoeken is het kijkgedrag van kinderen bestudeerd terwijl ze herhaaldelijk een vrouw keken die naar voren keek waarna ze haar ogen naar links of naar rechts bewoog. Tijdens het naar voren kijken waren er vier variaties toegepast, resulterend in vier verschillende condities die met elkaar vergeleken konden worden. In één conditie keek de vrouw naar voren en maakte oogcontact met de baby, een handeling die baby's zouden kunnen interpreteren als communicatief signaal om iets te laten zien. In een tweede conditie verschenen knipperende en bewegende gekleurde bolletjes over haar ogen die baby's niet zouden kunnen interpreteren als communicatief signaal. In een derde conditie verschenen er knipperende en bewegende gekleurde bolletjes over haar mond. In een vierde conditie verscheen er een animatie van een bloem voor het gezicht van de vrouw. In de eerste twee condities werd op verschillende manieren geprobeerd de visuele aandacht van de baby's naar de ogen van de vrouw te richten terwijl in de laatste twee condities werd geprobeerd de visuele aandacht van de baby's af te leiden van de ogen van de vrouw. Hiermee kon bekeken worden of het richten van de visuele aandacht van baby's naar de ogen het volgen van de kijkrichting kon faciliteren en of intentiebegrip hierin een rol speelt. Uit de resultaten kwam naar voren dat baby's de neiging hadden om de blikrichting van de vrouw te volgen indien hun aandacht was gevestigd op de ogen van de vrouw, vlak voordat ze van kijkrichting veranderde. We toonden verder aan dat het volgen van de kijkrichting gefaciliteerd kan worden door de aandacht van de baby's te vestigen op de ogen, vlak voordat deze van een frontale positie naar een laterale positie bewogen. De baby's volgden de kijkrichting hierbij ongeacht of de persoon voor hen sociale signalen, zoals oogcontact, liet zien, of dat er sprake was van niet-sociale signalen, zoals knipperende en bewegende stippen die over de ogen werden geprojecteerd. Dit suggereert dat baby's niet signalen zoals oogcontact hoeven te interpreteren als communicatieve intentie om de kijkrichting te kunnen volgen, zoals werd gesuggereerd vanuit een cognitivistische visie. Het richten van de visuele aandacht op de frontaal gepositioneerde ogen van een ander, vlak voordat die ogen een laterale beweging maken, lijkt voldoende te zijn voor baby's om de kijkrichting te kunnen volgen.

Imiteren van anderen

In hoofdstuk 3 werd een studie gerapporteerd waarin werd onderzocht of baby's de kijkrichting van een ander kunnen gebruiken om van die persoon te leren hoe bepaalde objecten gecombineerd kunnen worden. In deze studie keken baby's van 20 maanden verschillende malen naar een opname van een volwassene die mislukte pogingen deed om twee objecten op een specifieke manier te combineren, zoals een ketting in een kommetje doen en een elastiekje aan een stokje hangen. Hierna kregen ze zelf de kans om deze objecten te exploreren en te proberen om deze objecten te combineren. Voorafgaand aan de mislukte poging werd de visuele aandacht van baby's getrokken naar de oogregio van de volwassene door daarop knipperende en bewegende gekleurde bolletjes te laten zien (non-sociaal) of door middel van het maken van oogcontact (sociaal). In een derde conditie verscheen er een animatie van een bloem voor het gezicht van de vrouw waardoor de visuele aandacht op een non-specifieke manier naar het gezicht werd getrokken. Wanneer de aandacht van baby's (zowel sociaal als niet-sociaal) werd getrokken naar de oogregio van de volwassene, voorafgaand aan de mislukte poging, volgden de baby's de kijkrichting van de ogen. Dit liet zien dat het volgen van de kijkrichting ondersteund kan worden door de aandacht van de baby's te vestigen op de ogen, waarmee de resultaten van hoofdstuk 2 werden gerepliceerd. Wanneer de aandacht van baby's werd getrokken naar de oogregio van de volwassene, keken de baby's ook meer naar belangrijke onderdelen van de objecten en objectmanipulaties. Tevens waren ze naderhand succesvoller in het combineren van de twee objecten zoals de volwassene poogde te doen dan wanneer hun aandacht niet naar de ogen werd getrokken. Van de groep baby's waarvan de visuele aandacht niet naar de ogen van de volwassene werd getrokken keek echter ook een aanzienlijk deel naar belangrijke onderdelen van de objecten en objectmanipulaties, waarna ze de objecten konden combineren zoals de volwassene dat probeerde te doen. We concludeerden daarom dat wanneer baby's iemand waarnemen die probeert een specifieke handeling te realiseren, het richten van de visuele aandacht op de belangrijkste objectmanipulaties en objectkenmerken belangrijk is om de beoogde handeling te kunnen realiseren. Het volgen van de kijkrichting kan dit proces ondersteunen. Ten slotte werd er geen enkel bewijs gevonden voor de bewering dat baby's de intentie achter de mislukte poging zouden moeten begrijpen om de objecten te combineren zoals de volwassene dat probeerde te doen.

De waarneming van emotionele gezichtsuitdrukkingen

In hoofdstuk 4 werd voorgesteld dat de perceptie van verschillen tussen gezichtsuitdrukkingen en hun betekenis voortkomt uit contextafhankelijke functionele relaties tussen de waargenomen uitdrukkingen en de mogelijke acties die daarop van toepassing zijn. Zo leren kinderen bijvoorbeeld dat de boze blik van een ouder betekent dat ze berouw moeten tonen als ze stout zijn geweest. Op basis hiervan werd voorspeld dat naarmate kinderen ouder worden, ze deze functionele relaties specifieker waarnemen, resulterend in een betere vaardigheid om gezichtsuitdrukkingen te identificeren aan de hand van de context waarin ze gebruikelijk worden waargenomen. Om deze voorspelling te toetsen onderzochten we de vaardigheid van kinderen van vijf tot negen jaar in het identificeren van zes verschillende emotionele gezichtsuitdrukkingen (namelijk angst, verdriet, blijdschap, boosheid, walging, en verbazing) aan de hand van de gebruikelijke context waarin ze normaal gesproken worden waargenomen (bijvoorbeeld een verdrietig gezicht in de context van een kapotgemaakt lievelingsspeeltje). De resultaten lieten zien dat naarmate kinderen ouder zijn, ze beter en sneller zijn in het identificeren van gezichtsuitdrukkingen aan de hand van de context waarin ze gebruikelijk worden waargenomen. Daarnaast werd gevonden dat oudere kinderen minder verschillende uitdrukkingen foutief identificeerden als behorend tot een context waarin ze normaal gesproken niet worden waargenomen. Tezamen suggereren deze resultaten dat kinderen met de leeftijd visuele kenmerken van gezichtsuitdrukkingen en de omgeving op een specifiekere manier kunnen waarnemen en aan elkaar kunnen relateren. Deze perceptuele informatie kunnen ze vervolgens gebruiken om er snel en adaptief op te kunnen reageren.

Inzichten verwerven door te doen

In hoofdstuk 5 is geprobeerd te laten zien hoe inzichtverwerving plaats kan nemen in de waarneming en het handelen, in plaats van in termen van denkprocessen. Hiervoor werd de ontwikkeling in strategiegebruik van kinderen van 11 jaar onderzocht in een leertaak waarin ze series van aaneengeschakelde tandwielen zagen en de rotatierichting van het laatste tandwiel moesten bepalen aan de hand van de rotatierichting van het eerste tandwiel. Vanuit het idee van perceptueel leren werd voorgesteld dat de interactie die het kind heeft met het taakmateriaal perceptuele informatie creëert die gebruikt kan worden voor het ontdekken en toepassen van efficiëntere oplossingsstrategieën. Sommige kinderen begonnen bijvoorbeeld aan de taak met een sensomotorische strategie waarin ze de draairichting van ieder tandwiel simuleerden en volgden met de ogen en vingers tot aan het laatste tandwiel ('force-tracing' strategie). Door dit toe te passen ontdekten de meesten van hen dat de draairichting van ieder volgend tandwiel in tegengestelde richting bewoog ten opzichte van het tandwiel ervoor. Deze pas verworven informatie gebruikten ze vervolgens om een nieuwe, snellere en minder foutgevoelige strategie toe te passen waarin ze ieder opeenvolgend tandwiel afwisselend als 'linksdraaiend' of 'rechtsdraaiend' konden classificeren tot aan het laatste tandwiel ('classification' strategie). Uit de studie kwam naar voren dat het gebruik van de strategie was geordend van simpele en sensomotorische strategieën die foutgevoelig waren en relatief veel tijd vergden om toe te passen, naar geavanceerde en abstracte strategieën die snel toepasbaar waren en nauwkeurige oplossingen voortbrachten. Deze ontwikkelingsvolgorde sloot aan bij de ideeën van perceptueel leren waarin werd voorgesteld dat kinderen nieuwe inzichten of strategieën kunnen verwerven aan de hand van perceptuele informatie die voortkomt uit het werken aan een leertaak. Hiermee heeft dit onderzoek geïllustreerd hoe inzichtverwerving plaats kan vinden in de waarneming en het handelen, in plaats van in termen van denkprocessen.

Algemene conclusies en toepassingen van het proefschrift

In hoofdstuk 6 werden op basis van de gezamenlijke conclusies uit de voorgaande studies suggesties gedaan met betrekking tot toekomstig onderzoek, theorievorming en toepassingen binnen het onderwijs en therapie. Als algemeen discussiepunt werd er gesteld dat perceptueel leren moet worden gezien als een domein-algemene leertheorie die kan worden gebruikt om de volledige breedte van het menselijk gedrag en ontwikkeling te onderzoeken en te beschrijven. Ter illustratie van dit principe werd autisme omschreven met problemen in het perceptuele leren als kenmerkend. Er werd voorgesteld om interventies en onderzoek met betrekking tot autisme te richten op het verbeteren van perceptueel leren in termen van het leren oppikken van visuele informatie die constant blijft over veranderingen en verschillende situaties. Als een algemene vaardigheidstraining kan dit mogelijk bevorderlijk effecten hebben op de ontwikkeling in veel verschillende domeinen, en dan met name in het sociale domein waarin kinderen met autisme doorgaans veel moeite hebben.

Hiernaast werd perceptueel leren besproken in relatie tot een onderwijsvisie en toepassingen binnen het onderwijs. Als onderwijsvisie werd voorgesteld om het onderwijs dusdanig in te richten dat kinderen kunnen leren hoe ze hun visuele aandacht kunnen richten op de relevante informatie in de leeromgeving, wat vervolgens nieuwe inzichten en vaardigheden kan opleveren. Het leren om zelfstandig te kunnen distilleren (of eigenlijk differentiëren) van nuttige informatie om leertaken vervolgens op inzichtelijke wijze uit te kunnen voeren zou kinderen de mogelijkheid kunnen bieden om dit ook buiten een onderwijscontext te kunnen toepassen. Een manier waarop deze vorm van perceptueel leren gestimuleerd kan worden is door variabiliteit te creëren in leertaken. Variabiliteit aanbrengen in de interactie tussen het kind en de leertaak kan bijvoorbeeld bestaan uit het aanbieden van een leertaak in telkens een andere context. Hierdoor worden kinderen gestimuleerd om hun aandacht te richten op de informatie die constant blijft over verschillende contexten heen en dus relevant is voor het ontwikkelen van een inzicht of vaardigheid die in alle contexten van toepassing is. Alhoewel er al enige evidentie bestaat voor dit idee, kan het nog verder uitgewerkt worden teneinde het toe te kunnen passen in de praktijk.

Voorwaarts met de ecologische visie

Dit hoofdstuk werd afgesloten met een epiloog waarin werd aangegeven hoe het onderzoek binnen de ontwikkelingspsychologie vooruit kan gaan door de cognitivistische visie los te laten ten gunste van de ecologische visie. Er werd gesteld dat verklaringen van het menselijk gedrag en hoe het zich ontwikkelt vanuit een cognitivistische visie meestal worden omschreven in termen van denkprocessen en mentale representaties. Deze denkprocessen en mentale representaties kunnen echter nauwelijks empirisch geverifieerd of gefalsifieerd worden, wat werd omschreven als een ernstige tekortkoming. Het gemak waarmee dit soort onfalsifieerbare verklaringen geformuleerd worden zou kunnen liggen aan de menselijke tendens om een empirisch gegeven (bijvoorbeeld: de baby imiteert mijn gedrag) niet los te kunnen zien van de eigen interpretatie van dat empirische gegeven (bijvoorbeeld: als ik imiteer, dan is dat omdat ik de intentie van dat gedrag begrijp, dus als de baby imiteert, dan begrijpt de baby de intentie van mijn gedrag en imiteert deze). Daarnaast werd reificatie besproken als een probleem dat veel voorkomt binnen de cognitivistische visie. Reificatie betekent dat verklaringen van empirische bevindingen (bijvoorbeeld: de baby begrijpt en imiteert mijn intentie) ten onrechte als een onderdeel worden gezien van de empirische werkelijkheid (bijvoorbeeld: de baby imiteert mijn gedrag) terwijl die verklaringen daar eerder een abstractie van zijn (bijvoorbeeld: allicht is intentiebegrip niet nodig om gedrag van iemand te kunnen imiteren). Het begrip incommensurabiliteit werd gebruikt om aan te duiden dat er vanuit de ecologische visie geen logisch dwingende reden kan worden gegeven waarmee de aannames van de cognitivistische visie ontkracht kunnen worden. Daarentegen kan de ecologische visie wel terrein winnen binnen de ontwikkelingspsychologie door hypotheses voor te stellen met een minimum van aannames die de empirische toets kunnen doorstaan en waar men verder kan komen in het onderzoek en de praktijk. Dit zou over tijd kunnen leiden tot een paradigmaverschuiving waarbij we de cognitivistische visie op gedrag en ontwikkeling achter ons kunnen laten ten gunste van een ecologische visie. Het huidige proefschrift biedt daartoe een aantal voorbeelden en voorstellen.

Dankbetuiging

Het was een lange weg naar het ‘Hora Est’. Het zijn van een buitenpromovendus betekent mijns inziens dat je afhankelijk bent van wat andere mensen bereid zijn voor je te doen en van wat ze je gunnen. Ondanks de afwezigheid van directe collega’s waarmee ik kon samenwerken aan het promotieonderzoek, ging aan het afronden van de dissertatie zondermeer een sociaal traject vooraf. Van velen om mij heen heb ik hulp, steun en begeleiding gekregen, waardoor ik op de schouders van reuzen kon staan. Zonder deze mensen had ik dit nooit voor elkaar gekregen, waarvoor deze dankbetuiging.

Allereerst gaat mijn dank uit naar mijn mentoren en begeleiders van mijn bachelorthesis (Jan Boom), masterthesis (Ralf Cox en Fred Hasselman) en dissertatie (Ralf Cox, Fred Hasselman en Paul van Geert). Jan, ik ben al sinds 2004 welkom bij jou om moeilijke vragen te stellen en interessante discussies te voeren. Jij hebt mijn passie voor fundamenteel onderzoek altijd gestimuleerd en ik kon altijd bij jou terecht! Waar soms het onderzoek, de wiskunde en de onderzoekspraktijk lastig werden, heb je mij dikwijls een duw in de rug gegeven of een schop onder mijn kont om door te gaan en verdieping te zoeken. Zonder jou was ik nooit zo ver gekomen, ook ver nadat ik mijn bachelorthesis bij jou had afgerond! Ralf, ik ben je tegengekomen tijdens mijn master op het punt dat ik twijfelde aan psychologie als wetenschap en jij mij introduceerde in de ecologische psychologie. Je enthousiasme en passie voor het onderzoek hebben mij overgehaald om door te zetten in het onderzoek en een uitdagende masterthesis te schrijven. Binnen een jaar hadden we twee interessante onderzoeken afgerond die mij tezamen de beste thesisprijs van het jaar opleverde en wat inmiddels ook nog eens gepubliceerd is in twee aparte artikelen. In de jaren daarop was je altijd bereid met mij in gesprek te gaan en mijn stukken voor publicatie door te nemen tot op het punt dat ik weer student van je was, ditmaal in de vorm van buitenpromovendus. Als begeleider van mijn masterthesis en mijn dissertatie heb jij mij altijd de vrijheid gegeven om op mijn eigen manier precies te onderzoeken wat ik interessant vond. Fred, mijn begeleider van mijn masterthesis en mijn copromotor, jij bent ontzettend belangrijk geweest voor mij wat betreft de wiskunde maar ook het filosofische fundament van het de onderzoeken die ik heb verricht. Ik ken niemand die Dynamische Systeem Theorie, een veelal wiskundige theorie, leeft en ademt zoals jij dat doet. Jouw colleges en de gesprekken met jou hebben mij geweldig geïnspireerd en een nieuwe wereld voor mij laten zien. Paul van Geert, mijn promotor, is zo vrij geweest om op aandringen van Ralf mij aan te nemen als buitenpromovendus slechts 4 dagen voor hij met emeritaat ging. Paul, ik ben je heel erg dankbaar voor deze kans en hoop dat je je kan vinden in de manier waarop ik deze benut heb. Daarnaast ben ik je zeer dankbaar voor je hulp bij het afronden van deze dissertatie.

Ludger van Dijk is de meest rigoureuze, analytische, strenge maar ook beste wetenschapper die ik ken. Sinds onze ontmoeting in 2012 (met zijn befaamde digitus medium) hebben we

regelmatig gesprekken gehad over wetenschap en filosofie onder begeleiding van bier en ballen. Deze gesprekken ervaarde ik als periodiek onderhoud van de geest, waarin hij als een slijptol geduldig maar doelgericht mijn denken scherper maakte. Daarnaast heeft hij mij op een cruciale manier bijgestaan bij het schrijven van verschillende onderdelen van mijn dissertatie. Dankjewel makker! Verder had ik veel steun aan mijn lotgenoten in het promoveren zoals Sylke, Mona, Annika en Tessel. We konden bij elkaar terecht voor een goed potje cynisme, zeuren over de 'lord of death' en sleurden elkaar door het onderzoeksproces heen. Dank voor jullie toegankelijkheid en lichtzinnigheid in soms zware tijden. Daarbij wil ik ook mijn dank betuigen aan Willemijn die het moreel hoog wist te houden in de soms barre omstandigheden waarin wetenschappers moeten overleven.

In 2015 werd ik enigszins uit het niets gebeld door professor Tom ter Bogt met de vraag of ik geïnteresseerd was in het verzorgen van cursussen van de opleiding Algemene Sociale Wetenschappen aan de Universiteit Utrecht. Zonder dat hij het wist, heeft hij me de kans van mijn leven gegeven om aan mijn buitenpromotie te kunnen werken. Vanuit de rol als docent kon ik mijn onderzoeken aansluiten op het onderwijs, zodat ik met studenten data kon verzamelen. Daarnaast heeft hij ervoor gezorgd dat ik naast mijn werk als docent voldoende tijd over had om aan mijn promotie te kunnen werken. Hartelijk dank voor deze onvergetelijke kans!

In de jaren dat ik docent was en aan mijn buitenpromotie werkte, was ik gezegend met een groot aantal collega's bij de Universiteit Utrecht waarmee ik iedere dag lol kon beleven en waar ik heel veel aan heb gehad. Ester, ik heb mij ronduit kapot gelachen om je grapjes en genoten van onze collegialiteit. Je was altijd de welkome afleiding als ik middenin een moeilijke berekening zat en er bijna was! Ik ben heel blij en dankbaar dat je mijn paranimf wil zijn! Evi, slimme meid, ik zal op groene partijen stemmen en ben je veel dank verschuldigd voor de manier waarop je altijd bereid was om mij te helpen als ik weer eens vragen had over ons werk als docent. Robbie, topdocent, het is een eer om aandacht te krijgen van iemand die op speelse wijze de aandacht van duizenden studenten per jaar trekt en ik hoop op een lange voortzetting van gezellige avonden. Wil, dank voor je hulp, meedenken en boekhouding, wat voor mij in vriendschap, tijd, geld en middelen veel meer heeft opgeleverd dan ik voor mogelijk had gehouden. Nelly, rots in de branding waar de afdeling Interdisciplinaire Sociale Wetenschappen (ISW) op draait: ik ben jou net zoals iedere ISW'er simpelweg dank verschuldigd voor het feit dat jij de boel bij elkaar houdt, werkbaar en geloofwaardig maakt. Als jij ooit de universiteit verlaat zal je onherroepelijk een stuk daarvan wegnemen. Ook ben ik dank verschuldigd aan professor Jaap Bos, professor Catrin Finkenauer, dr. Gerrit Bartus Diellissen, dr. Jeroen Hessels en dr. Vincent Duindam voor de wijze raad en ondersteuning waarmee ik me een weg kon banen in het wetenschappelijke bedrijf. Leden

van de Labondersteuning FSW UU (Jurgen, Tom, Son); hartelijk dank voor de technische ondersteuning en het vrij beschikbaar stellen van al het onderzoeksmateriaal waarmee ik mijn data heb kunnen verzamelen. Ook wil ik hier een klein aantal van de vele studenten bedanken die mijde afgelopen jaren hebben geïnspireerd en die mij hielpen met het verzamelen van data waarmee ik mijn dissertatie tot stand kon brengen, namelijk Cas, Jacinta, Julia, Kayla, Lidewij, Michelle, Nina, Patrick, Roosmarijn, Saskia en Susan. Aan het einde van vele lange werkdagen kon ik altijd mijn hart ophalen bij beveiligingsmedewerker Jeffrey Hessing bij de receptie van het Sjoerd Groenmangebouw van de Universiteit Utrecht. Man, wat was het fijn om de werkdag te eindigen met iemand die zo nuchter en positief in het leven staat.

Ten slotte gaat mijn dank uit naar mijn vriendenkring en familie. Bij Tim kon ik altijd terecht met vragen en discussies over complexe wiskunde en programmeren. Met Gereon, Michiel, Remi en Rafa heb ik de wereld van filosofie kunnen ontdekken. Sophie ('Ga jij eens een biertje voor mij halen?') en Marco zijn mijn wetenschappelijke metgezellen sinds de eerste dag van mijn studie. Marco heeft mij als paranimf tot op het laatste punt geholpen met het schrijven van deze dissertatie. Les amis de l'Apero Ghetto (biensûr) en de jongens (Jaake, Oscar, Sam, Julio, Kevin, Tony, James, Giovanni, Bram, Willem en Christian) hebben mijn gezeur en gezwets over m'n monnikenwerk jarenlang moeten aanhoren. Met hun ondersteuning, liefde en vriendschap maakten ze het werk licht voor mij. Mijn moeder is altijd mijn grote voorbeeld geweest. Haar onuitputtelijke energie en sterke doorzettingsvermogen hebben mij geïnspireerd om dit boek af te maken, met name in tijden dat het even tegenzat waarin zij mij dikwijls opdroeg om maar eens een grote biefstuk te kopen. Mijn broers (Willem en Robbie) en zus (Cees) hebben altijd goed voor mij gezorgd en zijn altijd mijn voorbeelden en mijn vangnet geweest. Mijn vader stond aan het einde van de lange dagen altijd voor mij klaar met wijze raad en met een bemoedigend biertje. Mijn lieve vrouw Carole en onze prachtige kinderen Rafaël en Emilie; het werk zit erop! Dank voor jullie zorg, geduld en vertrouwen.

Curriculum vitae

Pieter de Bordes was born in Amersfoort on March 28th in 1985. After finishing high school in Zeist, he moved to Utrecht where he obtained his propaedeutics in teaching in 2004 and started to study psychology at Utrecht University. During his studies, he specialized in both developmental psychology and social psychology, took part in an Erasmus Exchange program at the University of Barcelona, and completed a minor in methods and statistics in social sciences and several courses in philosophy. He obtained his bachelor degree in psychology with honours at Utrecht University in 2008, after which he joined a research master program in behavioural Science at the Behavioral Science Institute of the Radboud University in Nijmegen. In 2010, he obtained his research master degree and was awarded with the price for best master thesis of 2010 of the Behavioral Science Institute. In the years after, he worked as a bartender in Fuerte Ventura in Spain (2011), lecturer and junior researcher at the faculty of Social Sciences of Utrecht University (2011–2014), and study advisor at the faculty of Business and Economics of the University of Amsterdam (2015). In 2016, he started working as a senior lecturer, course coordinator and teacher coordinator at the department of Interdisciplinary Sciences faculty of Social Sciences of Utrecht University. During this appointment, he worked on his PhD at the University of Groningen in his free time. Pieter is currently working as a consultant at the department of Educational Consultancy & Professional Development at University Utrecht.