Development of routine and supervisory processes in sequential action control

https://eprints.bbk.ac.uk/id/eprint/48292/

Version: Full Version

Citation: Carteron, Aude Marie (2022) Development of routine and supervisory processes in sequential action control. [Thesis] (Unpublished)

BIRKBECK, UNIVERSITY OF LONDON

PHD THESIS

Development of routine and supervisory processes in sequential action control

---

Aude CARTERON 2021

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

Department of Psychological Sciences

Declaration of Authorship

I hereby declare that the work presented in this thesis is my own. All work and materials that are drawn from others are always clearly attributed.

Aude CARTERON

Acknowledgements

I am extremely thankful to Rick Cooper and Denis Mareschal for their invaluable guidance while giving me trust and freedom to conduct this PhD.

I am particularly grateful to Rick, my primary supervisor, for his dedication in sharing his knowledge and sharing constructive feedback with so much kindness, as well as for encouraging me to develop my own ideas. I was genuinely lucky to receive his constant support and patience during these three years.

I am also grateful to Denis, my secondary supervisory, for his insightful ideas, for the inspiring discussions that arise out of his passion for developmental science, computational modelling and every possible topic, and relatedly for his contagious motivation.

Thank you Leslie, Berta, Tamsin, Annabel, Isabel and Sarela for making the lab logistics so smooth but also for making the CBCD such a pleasant work environment.

I am also grateful for the enthusiasm with which families who took part in my studies at the Babylab.

I am indebted to my army of proofreaders- Victoria, Joanna, Emma, Chiara, Giulia, Olatz, Penny, Ellie, Cécile, Umay and Meg, who so kindly helped me and made this thesis much clearer.

I was more than lucky to find wonderful friends among my CBCD lab mates, who made this journey much more fun. Most of you are already cited above. Special thanks to Louisa for being like a second Mum.

Being part of the MOTION network made the PhD even more special. Thank you Sabine Hunnius and Kwan Lee for making the MOTION adventure successful. It was a delight to do team work and spend the training schools with all my fellow ESRs — special mention to Lisanne, Penny, Julia, Victoria, Saya, Iara, Joanna and Tommaso. Thank you as well Erwin Schoonderwaldt and Frederik Muller for the great collaboration on Qualisys' Motion capture.

Even though "you don't choose your family"... I would not dream of a better one. Thank you Mum and Dad for your unconditional support. Mum — You are simply the best human on earth; there is not enough space here to list all that I am thankful

for. Alexandre — After from the ones who made and raised me, you are the person who contributed the most in allowing me to get where I am professionally. You not only transmitted some of your geekiness and technical help but more importantly shared your unique perspectives and irreplaceable advice whenever I needed it. You helped me make the best decisions. To my grown-up babies, Flore and Valentine — Your hugs were a significant contribution to helping me complete this journey.

I am grateful to the fabulous friends who cheered me up all along. Special mention to: Les Brévi who have grown up with me and have been a rock for more than 15 years. La Compi crew for the refreshing craziness, and jumping to help on my pilot studies. Soso for your irreplaceable empathy and friendship. Léti for making it easier to disconnect from the thesis. Vivian for supporting me during the less fun period of writing.

Mamie, I wish you were here to read these words. Your endless determination and strength have inspired me and I dedicate this thesis to you.

This PhD was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 765298.

Abstract

The sequential actions that children and adults perform regularly if not daily (e.g., preparing for and going to school/work, preparing meals, and so on) are often under routine control in that they appear not to require overt attention. The study of routine action control in adults has benefited from influential theories, such as the Norman and Shallice's (1986) *dual-systems theory*, supported by comprehensive computational models. Drawing on the latter theory, and comparing it with other existing accounts of sequential action selection, this thesis aims at improving our understanding of the development of routine action control throughout the school-age years. It investigates how children control complex action sequences, at several levels, and with the involvement of various supervisory functions (including inhibitory control and monitoring functions). It furthermore explores the interaction between the two hypothesised action control systems in children under the lens of the dual-systems theory, but also under the lens of the so-called *model-free* and *model-based* types of reinforcement learning. This is done by designing child-friendly tasks, developing a computational model, and proposing novel analysis methods for kinematics data.

The findings in this thesis support the view that children use two modes of control which may follow different developmental trajectories, with a supervisory system following a more protracted development. The results furthermore suggest that the development of inhibitory control throughout the school-age years might reduce children's propensity to interferences from environmental distractors, and might improve their abilities to select the appropriate action in an ambiguous context (e.g., when an action needs to be related more strongly to the overarching goal than to the preceding's action) or under increased cognitive load.

In conclusion, this thesis shows that by 5 or 6 years old, children readily use conjointly two modes of action control and are able to control action sequences in a routinised fashion, yet the supervisory mode of control seems to substantially improve throughout mid-childhood. It furthermore brings evidence for the fact that changes in executive functions underlie improvements in sequential action control with age.

Contents

List of Figures

List of Tables

Chapter 1

The Control of Action Across Development: Theoretical and Methodological Foundations

Abstract

This Chapter starts by introducing the relevance of studying the development of sequential action control and presents the domain of action of concern (the so-called *intermediate domain of action*). It then presents key concepts in action control, breaking it down into small components and laying out the developmental trajectories of such separate aspects of action control. This leads to the introduction of integrative theoretical accounts of sequential action control. The most influential of them involves two systems working together (Norman and Shallice, 1986, Cooper et al., 2014) or in competition (Daw et al., 2005) to select actions. The accounts are then compared, and the gap in our understanding of the interactions of two systems for action control throughout development is highlighted. Finally, important methods to study action control are presented before giving an overview of the thesis Chapters.

1.1 Introduction

Picture a 7-year-old, on a Monday morning, walking into the kitchen for breakfast. She heads to the bottom right cupboard to pick up a bowl, then to the cutlery drawer to fetch a spoon, avoiding the fork that has been placed in the wrong slot. She picks up her favourite cereal box, frees her hand by placing all items on the table, now heads towards the fridge where she pauses an instant facing the temptation of her favourite chocolate bar (knowing she would be told off for eating it at this time rather than as an afternoon snack), and successfully picks up the milk instead. She fails to locate the orange juice thus turns around to find it is already on the kitchen table. Having in front of her a full array of objects for breakfast, she reaches for the cereal box while hearing her mum's reminder to wear warm clothes for the cold day ahead. A few seconds later she realises that she has been pouring the cereal into the glass instead of the cereal bowl- she thus corrects this by transferring the cereal from the glass to the right recipient. She then helps herself to some orange juice, and finally goes on to lift the bottle of milk that she unexpectedly finds empty- a consequence of her sibling's sneaky passage through the kitchen. She decides to eat the cereal dry instead of getting up to get a new bottle of milk.

This scene embodies many aspects of the control of daily-life activities that will be recurrent in this thesis: achieving one's goal, selecting a target among surrounding objects, sequencing actions while collecting external information, maintaining a goal in mind, inhibiting unintended actions, monitoring the environment, updating plans, and so on. This thesis explores the development of action control in the context of routine or routine-like daily action sequences.

1.1.1 Routinisation or modularisation starts early in development

Bruner (1973) introduces the importance of routines and subroutines in the development of goal-directed skilled action, in particular in infancy. He presents modularisation, the process by which an act or a serially ordered sequence of acts are mastered, as an essential phenomenon to construct more and more complex behaviour. In a modularised subsequence, the component acts are well-mastered and thus do not require to be controlled at the individual level. Thus, less attention is needed for controlling such a sequence. The freed-up attention resources consequently allows the incorporation of acts into higher-order sequences and the regulation of such sequences. Bruner (1973) highlights two ways in which such constituent acts or subroutines develop through practice in early childhood. One is through play (especially when the action consists in physically altering states of the world, e.g., Schiller, 1952), the other is through modelling adults action, which can also be a form of play.

Similarly, the skill theory of Fischer (1980) underlies that cognitive development happens through the mastery of simpler skills that are integrated together to build progressively more complex skills. This theory tackles motor skills as well as encompasses other domains such as cognitive, social and language skills in a unified framework. Both early accounts of the development of skilled action highlight the relevance of studying routines, or well-mastered action sequences, in childhood. When fewer resources are needed to perform constituent subroutines, the execution of more complex higher-order acts, or goal-directed sequences, becomes possible. Greater resource availability when executing daily-life routine action sequences also has the advantage of leaving attentional resources. Such resources may be allocated to multitasking or planning subsequent activities, which contributes to building more complex daily behaviours.

1.1.2 School-aged children control numerous action sequences in their daily life

Once acquired, the action sequences must be consistently and efficiently controlled in the face of a sometimes distracting or changing environment. Hence, beyond the mastery of basic motor acts and learning conceptual sequence order, it is essential to understand how action sequences are controlled (initiated and executed until the end of a sequence) with respect to goals and in accordance with the state of the environment. In adults, the control of action has been widely studied. Although school-aged children already partake efficiently in many routine activities, the study of action sequencing in children has traditionally been geared towards learning, and less towards the control of known sequences. The latter is the focus of this thesis.

When reaching school-age, i.e., around 5 or 6 years old, children already need to complete several routine sequences such as preparing breakfast, getting dressed, brushing their teeth, packing their backpack for school, etc. (Weisner, 1996). Furthermore, attending school requires acquiring a concentrated amount of novel knowledge, meaning that children will benefit from freed-up attentional resources when executing action sequences if these are routinised. Additionally, multitasking abilities appear to develop between the age of 7 and 12 (Yang et al., 2019), presumably as a consequence of efficient routine control abilities.

Finally, throughout the years, a set of core processes called executive functions (domain-general control mechanisms responsible for modulating many cognitive subprocesses) are known to improve greatly (e.g., Karr et al., 2018, T. Maldonado et al., 2020). These subprocesses will be described in more detail in Section 1.3.2.2. Such executive functions are also thought to be important for efficient action control. Pooling together the vast knowledge in the executive functions literature about changes throughout childhood together with studying action in children can further our understanding of the developmental changes underlying action control.

For these reasons, this thesis will particularly focus on the traditional primary school ages, i.e. around 5 to 11 years old. Nonetheless, the developmental trajectory observed may expand to preschool years and adolescence.

1.1.3 The *intermediate domain of action* is of particular interest

Action can be studied at several levels. The lowest level of analysis (the execution of action) considers the biomechanics of movements: that is, measuring the movement kinematics as well as the forces behind posture, body configurations, speed of action and fine-grained adaptations to objects' physical properties (Jensen, 2005, Thoroughman and Shadmehr, 2000). This thesis is not directly concerned with this level of description, and a review of the developmental changes in movement biomechanics, including the biomechanics of interaction with objects (reaching, grasping, holding) is given by Adolph and Berger (2011).

At the highest level of analysis, we encounter what Schank and Abelson (1977) referred to as scripts or *Memory organisation packets* (Schank, 1982). For example, visiting a restaurant entails an overall script that includes greeting the staff, taking place at the restaurant table, ordering food, eating, and paying the bill. The subactions for each goal are not specified precisely. For example, the bill may be paid by card or cash (Cooper and Shallice, 2000). Moreover, the duration of the entire script can be highly variable from one execution to the next.

In between the two levels lies a level of its own as defined by Cooper and Shallice (2000): the *intermediate* domain (or level) of action. The intermediate level is the focus of this thesis. Contrary to the higher, *script* level, the *intermediate* level defines sub-actions more precisely and their selection is critical to carrying out the goal as intended (e.g., the sequence 'paying the bill by cash' will include the sub-actions 'fetching the wallet', 'taking and counting notes', 'handing them to waiter', and so on). Furthermore, at the script level the action timing is generally more variable than at the intermediate level. Unlike at the lowest level of action, at the intermediate level the local physics of the environment (e.g., the reachability of objects, the status of objects such as 'the wallet is already open') only need to be defined at a coarse grain. At the lowest level, the physical properties of the objects that are subject to an action are defined at a fine level which encompasses the biomechanics of movement, contrary to the intermediate level. The intermediate domain of action encompasses a wide range of routine activities as presented in the introduction: dressing, preparing one's school bag, preparing coffee, washing one's hand, starting a car, etc.

1.1.4 The development of action can be studied separately from motor skills

The level of action and the age range that are of interest in this thesis allow us to investigate developmental changes in action without considering the fact that motor skills are still changing across school-age years. Although body size and strength are still significantly changing, by school age children are largely proficient in basic motor acts such as reaching, grasping, holding and manipulating objects. In fact, such a motor competence is already present in 12- or 18-month-old toddlers (e.g., opening a bottle of water, using a spoon) and are well-mastered by the third year of life (e.g., Konczak and Dichgans, 1997, Adolph and Berger, 2007, Kahrs et al., 2013). Motor skills continue to evolve beyond school age (e.g., becoming more accurate and consistent Kahrs et al., 2012, Kahrs et al., 2014), but within our level of action of interest (*intermediate*) it is sufficient that the basic motor acts are proficient enough so that cognitive developmental changes may be studied independently from motor developmental changes.

Another important distinction to be made is that studying routine actions differs from studying motor skills, although there is a considerable overlap. Indeed the automatisation process is most likely carried out by the same areas: motor learning has been found to engage parietal areas (Kelly and Garavan, 2005, Jenkins et al., 1994), and more recent work has described the role of the striatum and the associated corticobasal ganglia loops in the production and regulation of movement sequences (Graybiel, 2008, Desmurget and Turner, 2010).

Yet the precise interest of this thesis can be distinguished from activities that rely more heavily on motor skills with a very specific movement pattern, such as typing on a keyboard, or sports such as dancing. Such skill learning is typically studied in the lab with tasks such as the serial reaction time task (Willingham, 1998) or the Discrete sequence production task (Karni, 1995), involving repeating a fixed sequence of button presses, responding to cues or by memory. Skill learning refers to gaining accuracy and/or speed when repeatedly performing such tasks and is reviewed by Diedrichsen and Kornysheva (2015). A simple way to define what types of action sequences fall in the scope of this thesis is to think of routine action sequences that are naturally repeated in everyday life (e.g., preparing meals, grooming, etc.), that can be tied together to an overarching goal and typically involve several objects.

1.1.5 Learning a sequence is different from routinisation

The process of acquisition of a novel goal-directed action sequence can be broken down into two phases (Anderson, 1987, W. Schneider and Detweiler, 1988): a first, *initial learning phase*, that consists in the rapid acquisition of knowledge about the sequencing of individual actions needed to achieve the goal, and a second, *consolidation phase*, that consists in establishing routinisation more gradually through repetition.

This thesis is mostly interested in how control occurs in routinised sequences, i.e. a phase where the sequence has been practised enough such that the involvement of prefrontal resources is substantially less than at the start of learning a sequence- at least for some steps or under easy conditions, as will be detailed later on. The acquisition process also needs to be considered for that purpose, and in particular the second, consolidation phase.

If one is to compare tasks where children and adults have received similar amounts of training, a methodological concern is whether age influences the speed of sequence acquisition. It remains unclear whether there are sensitive periods in acquiring motor skills, due to the methodological challenges posed by such studies (Thomas and Johnson, 2008, White et al., 2013, Solum et al., 2020). When considering learning in non-motor tasks, learning rates do not follow a clear developmental trajectory from childhood to adulthood as they have been found to remain similar (e.g., Javadi et al., 2014), increase (e.g., Master, Eckstein, Gotlieb, Dahl, Wilbrecht, and Collins, 2020) or decrease (e.g., Decker et al., 2015). Thus, the speed of learning is unpredictable and probably highly dependent on the task at hand.

The amount of time and training needed to routinise a sequence will necessarily depend on the sequence complexity and be subject to inter-individual differences. Nonetheless, to give an order of magnitude, Lally et al. (2010) found that participants took between 18 and 254 days to reach a plateau in automaticity when performing an activity incorporated at the beginning of their day (the automaticity was measured by a questionnaire including items such as 'I do it automatically', 'I do it without thinking' and 'I would find hard not to do it').

In this thesis, action sequences are studied in conditions allowing one to assume similar degrees of mastery or routinisation across age groups, or that such differences in routinisation are negligible compared to the fundamental developmental differences on action control systems.

1.2 Key empirical findings and theoretical concepts

In this Section, we review key features of action selection, that are critical when actions are integrated together in a sequence.

1.2.1 Affordances

While it is possible to study action control without considering the level of motor control, it is important to mention in what ways perception can directly affect action execution. Actions are influenced by the visual environment, a consequence of the direct coupling between perception and action (Allport, 1987). A wealth of evidence indicates that information can be transmitted automatically along visuomotor pathways (Meegan and Tipper, 1999a). For example, healthy participants will grasp a target object more quickly if the presentation of the object is preceded by the presentation of a task-irrelevant but grasp-compatible object, a phenomenon called visuomotor priming (Craighero et al., 1996). Additionally, patients with frontal lobe damage can exhibit utilisation behaviour, where they are unable to withhold a motor response such as picking up an object irrelevant to their current intentions (Lhermitte, 1983, Shallice et al., 1989). The anarchic hand syndrome is another example of behaviour triggered by visual stimuli without intention to perform the act (Della Sala et al., 1991, Goldberg et al., 1981).

Gibson (1977) introduced the concept of affordance, that is, that the visual attributes of an object indicate its potential for action. For example, a cup affords grasping. Affordances are determined by the intersection of the skills of the perceiver and the environment's characteristics. Thus, when body dimensions and motor skills change across development, affordances change as well (Adolph, 2008). Such a coupling between perception and action undergoes a long period of development (Plumert et al., 2007). The sensitivity to affordances starts to appear in the first year of life: 4-month-olds will remember better (Mareschal and Johnson, 2003) an object image that is graspable compared to a non-graspable item (image of face or flat cartoon). Before they even have the motor ability to grasp, 5-month-olds will shape their hand correctly as a function of the grasp afforded by an object (Barrett et al., 2008). 6.5-month-olds will adapt the use of one or two hands to the size of an object (Clifton et al., 1991). The imitation behaviour exhibited by 12- to- 18-month-olds grasping of a tool depends on their familiarity with the tool.

By school-years, children's perception of action patterns is remarkably similar than that of adults: actions are affected by affordances the same way in adults and 6-year-olds, and similar grasp-related areas of the brain are activated when 6-year-olds or adults look passively at tools (Dekker et al., 2010).

1.2.2 Goal-directedness

Together with the bottom-up (environmental) influences on action just presented, top-down processes must be at play to select actions that achieve the intended goal. A goal-directed action can be defined as an action that is organised or directed towards a specific observable end-state, such end-state achieving an intended goal (Lutkenhaus et al., 1987, Hofer et al., 2005, Von Hofsten, 2004, Von Hofsten, 2004).

The goals in children's goal-directed sequences may follow predefined task goals, either instructed (e.g., being told "prepare your school-bag for tomorrow") or belonging to daily-life routines (e.g., preparing breakfast after waking-up). The goals may also be related to play. Play can be goal-directed in a visible manner (e.g., completing a jigsaw puzzle) or look like random exploration but may satisfy curiosity, which can be considered a goal in itself.

From 18 months of age, toddlers are able to modify their actions with respect to an overarching goal by complying to the commands of adults (e.g., Kagan, 1981, Kopp, 1991) and are capable of choosing the most efficient means to their goal (Paulus and Sodian, 2015). From this age, they rapidly become better at selecting an action amongst competing actions and at correcting an action to meet a goal (Bullock and Lutkenhaus, 1988, Vaughn et al., 1984). Regarding behaviour sensitive to valuation of outcomes, or reward-driven behaviour, the sensitivity of children's goals to objects' value seem to appear between 19 and 24 months of age, as evidenced by a button-pressing frequency dependent on the value of an outcome in 24-month-olds but not in 19-month-olds. (Kenward et al., 2009). Similarly, from the age of two, children spontaneously select actions that allow them to obtain the most desirable outcome (Klossek and Dickinson, 2012). By the age of one and a half, toddlers engage with proximal goals, but they only engage with more distal goals from the age of two and a half (Jennings, 2004).

The development of action processing (the creation of action representations centred around a higher-level goal) is thought to be important for the later goal-directed action representations used in action production (Sommerville and Woodward, 2005). As early as 6 months of age, infants can interpret the individual actions of an actor as goal-directed (Hofer et al., 2007). At 7 months of age, infants can effectively link actions to their effects (see Hommel and Elsner (2009) for a review on action-effect learning). At 12 months of age, infants appear capable of interpreting different individual actions within a sequence as directed towards an overarching goal (Woodward and Sommerville, 2000).

1.2.3 Action planning

To efficiently achieve one's ends, it is useful to not just improvise the means (that is, goal-directed actions) at the moment when a situation is faced, but to plan actions ahead in time (whether it is planning for the next seconds, the next half hour, the next day or next year).

When reaching for an object, infants under 8 months already exhibit prospective motor control (Hofsten, 1993). Indeed, they adjust their speed prior to contacting the object (Clifton et al., 1994), adjust their hand orientation to the object grasp (Lockman et al., 1984, McCarty et al., 2001) and by 13 months of age they adjust the size of their grip (Hofsten and Ronnqvist, 1988) to object size. Furthermore, from 10-month-old infants will modulate their speed of reaching as a function of their intended subsequent action after picking up the object (Claxton et al., 2003). By 12 months of age infants have the cognitive substrate for controlling action sequences (Verschoor et al., 2015).

At 14 months of age, toddlers are able to plan their movement as a function of future difficulties (Gottwald et al., 2017). 21-month-olds show early planning of a subsequent action (precise or imprecise manipulation) in their earlier movement (Chen et al., 2010). Between 2 and 3 years of age, children start adjusting their dominant/non-dominant hand choice to goal-related information when using a tool (Cox and Smitsman, 2006). Finally, during preschool years, children are able to plan actions on a larger time scale and according to more complex rules such as the ones required in pretend play (Fenson and Ramsay, 1981, Fein, 1981, Hudson and Fivush, 1991) or in daily routines (e.g., Hudson et al., 1995).

1.2.4 Multiple levels and hierarchy

We now turn to longer action sequences such as the daily sequences that school-age children take part in.

Action sequences are hierarchical: they can be broken down into subgoals, and the subgoals themselves can be broken down into sub-actions. This breakdown may be conceived at a progressively finer grain to span a large number of levels (for example, preparing breakfast may contain the subgoal making coffee, which itself containing the sub-action adding sugar, itself containing the sub-action taking a spoonful of sugar).

The simplest hierarchical organisation is to consider the individual actions in a sequence as tied to an overarching goal. There is evidence that infants readily represent action with respect to a higher-level goal from the first year of life, as seen in Section 1.2.2. During preschool years, children are able to represent and attend to goals at any of a higher- or lower-level level (Freier et al., 2017,Yanaoka and Saito, 2017).

The hierarchy of everyday life action sequences often spans more than two levels: at the highest level sits the overarching goal, and at the lowest level, the details of the actions to be executed (e.g., picking up an item, unscrewing a bottle cap). In between, sub-actions can often be grouped together as achieving subgoals. For example, the sequence with the higher-level goal *preparing a lunch box* could be broken down into the subgoals *making a sandwich* and *making a fruit salad*. The subgoal *making a sandwich* can itself be broken down into the subgoals *preparing the bread*, *adding* *butter*, *adding tomato* and *adding cheese*. The *adding butter* subgoal can be further broken down into the lower-level actions *fetching a butter-knife*, *getting butter with the knife*, *spreading butter onto the bread* and *discarding the knife*.

The view of the hierarchical control of action dates back from the study of neurologically-damaged patients (Duncan, 1986, Schwartz et al., 1991) and is now widely accepted. Evidence from cognitive neuroscience strongly suggests the existence of a mapping between the hierarchical organisation of neural systems and the hierarchical structure of action sequences (Dehaene and Changeux, 1997). Different task structure levels are thought to be represented at different levels of hierarchically-organised cortical areas (Fuster, 1990, Koechlin et al., 2003, Courtney, 2004), with the frontal cortex acting upon the highest level of the hierarchy (Luria, 1966, Botvinick, 2008). However, studies based on computational work argue that an explicit neural hierarchical structure is not needed to perform hierarchically-structured action sequences (Botvinick and Plaut, 2004), and the debate is still ongoing (Botvinick, 2007, Botvinick, 2008).

Action can be controlled at diverse levels of this hierarchy (Heckhausen and Beckmann, 1990, Cooper and Shallice, 2000). Adults are able to execute a very well-learnt routine sequence by attending only to the higher level of the hierarchy. It is always possible to control at lower levels, up to the level of each individual action.

1.2.5 Branch points

A consequence of the hierarchical organisation of action is that transitions among individual actions that together achieve a subgoal (or *within-sequence* actions) are more strongly linked since actions within a subgoal are more frequently repeated together, compared to transitions between actions from different subgoals. Experimental data support this view, where transitions from one action to the next are more vulnerable to errors and/or take longer between-subgoals than within-subgoals (Ruh et al., 2008, Ruh et al., 2010). Similarly, Arnold et al. (2017) have shown more difficulties in adults' action selection in building a brick wall at critical *decision boundaries*. Further evidence for the hierarchical structuring, reviewed in Rhodes et al., 2004, comes from motor tasks such as typing (Logan and Crump, 2011), hand writing (Cheng and van Genuchten, 2018) or 'discrete sequence production' tasks (Karni, 1995).

1.2.6 Goal maintenance

To execute the appropriate actions at the lowest level of the hierarchy, a representation of the overarching goal must be maintained throughout the sequence. Thus, maintaining the higher-level goal contributes to eliminating inappropriate actions. When compatible but inappropriate actions compete (e.g., when distractors visually similar to a target are present), top-down selection processes come into play. The ability to maintain a goal improves significantly during pre-school and school years and is strongly related to other supervisory functions (Towse et al., 2007, Marcovitch, Boseovski, et al., 2007, Rueda et al., 2004, Jones et al., 2003). This will be detailed in Sections 1.3.2.1 and 1.3.2.2.

1.3 Theories of action control

1.3.1 Dual-systems theory

1.3.1.1 The theory

The *dual-systems* theory (Shallice, 1972, Norman, 1981, Norman and Shallice, 1986), originally termed *theory of attention to action* was developed to account for sequential action selection in the *intermediate domain of action*. Comparable to the commonly accepted distinction between automatic and controlled processes (Shiffrin and Schneider, 1977), and related distinction between implicit and explicit knowledge in development (Karmiloff-Smith, 1986), the theory of Norman and Shallice (1986) posits the existence of two systems involved in action control: the Contention scheduling and the Supervisory system, in original terms. This thesis will more generally call them *routine/non-routine system* or *routine/supervisory systems*.

The contention scheduling or routine system is held to be responsible for and able to carry out autonomously very well-practised action sequences (or routines). It can nonetheless be modulated by the supervisory system when required, i.e., when a sequence is novel or not routine enough, when the environment causes a challenge (e.g., unexpected consequence of an action), in case of error, distraction, or when a strong habit must be inhibited.

The dual-systems framework has been specifically developed to account for a wide range of phenomena in the control of action. A first source of data at the origin of the theory is the lapses in action performed by healthy adults. Diary studies (Reason, 1979, Reason, 1984, Reason, 1990, Norman, 1981) have allowed a systematic categorisation of action slips, including: capture errors (when instead of the intended sequence, an unintended yet appropriate to the environment sequence is executed, such as taking one's socks off when intending to only take one's shoes off), omission errors (omitting an intended action or subsequence, e.g., omitting to add coffee grounds to the machine when making a cup of coffee), anticipation errors (carrying out an action earlier than intended, e.g., pouring water from the tap before opening the lid of a kettle), perseveration error (unintentionally repeating an action or subsequence, e.g., adding too many spoonsful of sugar to coffee) or object substitution (executing an intended action but not with the intended object, e.g., using shaving cream instead of toothpaste when brushing one's teeth). Such action slips happen only occasionally in healthy patients and are found frequently in patients with the *action disorganisation syndrome* (Humphreys and Forde, 1998, Schwartz et al., 1991, Schwartz, Montgomery, et al., 1998).

Other neurological impairments mentioned previously, such as the *anarchic hand syndrome* (Della Sala et al., 1991, Goldberg et al., 1981) and *utilisation behaviour* (Lhermitte, 1983, Shallice et al., 1989) are relevant in how they affect the control of actions. Indeed, this type of syndrome reveals that action can be elicited by the environment without the participants' intention to act. For example, patients exhibiting utilisation behaviour will pick up objects in a way that is appropriate to the object features (e.g., grasp an apple with the appropriate opening of the hand) but subsequently report they did not want to perform the action.

1.3.1.2 The Interactive Activation Network (IAN) model of Cooper and Shallice (2000)

The dual-systems account is supported by the computational model of Cooper and Shallice (2000) which successfully explained critical aspects of healthy and disordered action control. The original theory of Norman and Shallice (Norman, 1981, Norman and Shallice, 1986) proposed that routine action sequences are controlled by schemas, representation units which may perform operations towards achieving a goal or subgoal. Importantly, schemas are organised according to an explicit hierarchical structure, that mirrors the hierarchical structure of behaviour.

The schema hierarchy is implemented in the Cooper and Shallice (2000) model (referred to later as IAN model) with an Interactive Activation Network (a type of model used by McClelland and Rumelhart, 1981, Rumelhart and Norman, 1982, McClelland, 1992). As shown in Figure 1.1, the model comprises of a central schema network, that interacts with a resource network and an object network, and that carries out action selection. The schemas may receive multiple sources of activations and activations are passed on through the hierarchy of schema components. Nodes in the schema network represent actions at different levels: at the highest level, complex goals (in the example simulated in the original study, preparing instant coffee), at intermediate levels, subgoals (e.g., add milk into coffee) and at the lowest level, direct action (e.g., pick up implement).

In essence, the nodes are linked together such that the superordinate schemas in the hierarchy may send excitation to the lower-level schema that achieves the superordinate goal. Additionally, schemas have arguments that connect them with the object representation network. The links between the schema and object representation networks are such that nodes that have overlapping requirements are mutually inhibitory; on the contrary schema and objects that are routinely used together are mutually excitatory. Together with the effector network and the completion of preconditions and postconditions, activations flow within this network to select actions sequentially. An action is held to be performed when a schema activation reaches the selection threshold. Consequently, the model represents important sources of activation of behaviour mentioned previously: influences from the environment (bottom-up influences), and from intentions (top-down).

The model successfully simulates a routine sequence of coffee making. Crucially, it is able to simulate a range of error patterns that are characteristic of the slips of action in healthy participants or patients with the disorders mentioned previously. It does so by modifying the weights of different sources of activations. For example, omission or anticipation errors arise when top-down or environmental influences do not have enough weight for the omitted schema to be activated. Perseveration errors are simulated when self-activation in a schema is too large or if lateral inhibition between mutually incompatible schema is too low, consequently failing to resolve the competition in the network thus not deselecting a schema at the right time. Furthermore, utilisation behaviour (Lhermitte, 1983, Shallice et al., 1989) is reproduced when the relative influence of external (environmental) influences is too high compared to internal influences. A final example is action disorganisation syndrome, that is simulated by increased noise in the schema network.

A limitation of the model is that it does not provide a mechanism for the learning and routinisation of the sequences.

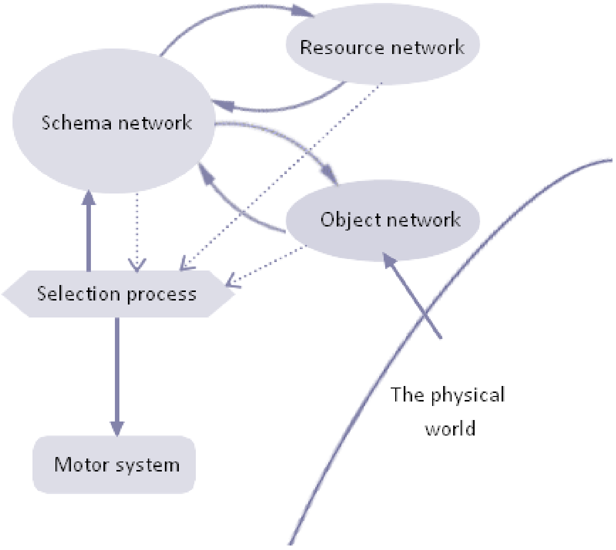


Fig. 1.1: **Main components of the contention scheduling implemented in Cooper and Shallice (2000). The schema network represents goals and their hierarchy. The resolution of competition within this network leads to the selection of action. The object network contains activation-based internal representation of objects. The resource network is analogous to the object network and takes charge of solving resource allocation (for example, if the right hand is taken, one should either wait or use the left hand). Figure adapted from Cooper and Shallice (2000).**

1.3.1.3 The Simple Recurrent Network (SRN) model of Botvinick and Plaut (2004)

Botvinick and Plaut (2004) proposed an alternative model to the IAN model (Cooper & Shallice, 2000), which consists in a Simple Recurrent Network and will be referred to as the *SRN*. The SRN model challenged the need for a localist approach with explicit hierarchical schema organisation and proposed instead a recurrent connectionist network (Jordan, 1986, Elman, 1990, Elman, 1991) with distributed representations.

Within the model (shown in Figure 1.2), perceptual input units encode environmental features (object representations) and output units encode actions upon the object. In between, internal units with recurrent connectivity allow forming distributed representations of the task context and of temporal information. The model was able to reproduce sequences of the same complexity as coffee making in Cooper and Shallice (2000), and to simulate error patterns characteristic of action disorganisation syndrome by increasing noise in the network. However, the SRN model can not account for as many phenomena as the Cooper and Shallice (2000) model (see the comparison in Cooper and Shallice, 2006). Since the dual-systems remains the most plausible account, it is the framework used throughout this thesis.

Nonetheless, the SRN model generated behavioural predictions that have led to interesting experimental work, including with children, that we detail hereafter. Action selection within the SRN relies on the model's internal representation of task context. For sequences with overlapping then diverging sequences (e.g., *making coffee* always involves coffee and cream, but can be with or without sugar), the diverging point (or branch point; here the moment before *adding sugar*) necessitates a strong representation of the context or goal (e.g., *make a coffee with sugar*) at the end of the preceding subsequence. If that representation is not strong enough at that time (e.g., at the end of *adding cream*), the subsequence *adding sugar* may fail to be added. However, within a subsequence, the task context is less relevant since the subsequence (that is, the actions required for adding cream) is consistently performed the same way. Thus, the context is represented more robustly towards the end than the middle of a subsequence. Consequently, the context may be more easily disrupted in the middle, where it is less strongly represented. A prediction that follows is that, when subject to a distraction, errors are more likely to occur if the distraction occurred towards the middle of a subsequence than if it occurred at the end of a subsequence.

This model prediction was tested experimentally in adults performing real coffee-making sequences (Botvinick and Bylsma, 2005). As predicted, participants made more errors after being interrupted midway through a subsequence than after being interrupted at the end of a subsequence. Such results were replicated in preschoolers (Yanaoka and Saito, 2019), indicating that task representations in children are qualitatively identical to that of adults.

1.3.1.4 The Goal Circuit Model (GCM) of Cooper et al. (2014)

A shortcoming of the IAN (Cooper and Shallice, 2000) was that it does not model a routine sequence acquisition process nor explicitly implement the supervisory functions. In the original dual-systems theory, the supervisory system is held to intervene to modulate the operations of the routine system when needed (e.g., resisting a habit or dealing with a non-routine situation).

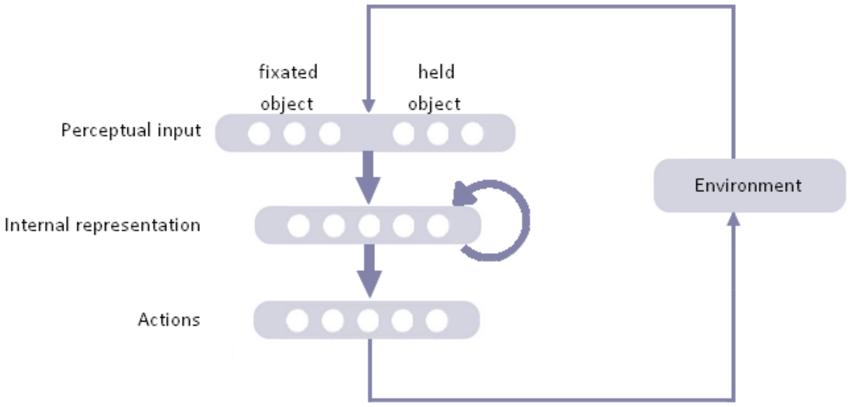


Fig. 1.2: **Architecture of the model of Botvinick and Plaut (2004). Each line of white disks represent units grouped in a layer. Arrows denote connections between group of units.**

Cooper et al. (2014) address this problem by proposing a novel model, the Goal Circuit Model (GCM) which implements an interface of the supervisory system with the contention scheduling system. In that model, the contention scheduling system is a variation to the SRN of Botvinick and Plaut (2004). The SRN is interfaced with a bank of goal units, representing the supervisory system, that exerts control over the contention scheduling system.

The GCM architecture is presented in Figure 1.3. A bank of inputs units represent the objects (fixated and held), and another bank of input units represent the current goal. In order to generate an action and a representation of the following goal (denoted as *predicted goal*), all input units feed through a hidden layer with recurrent connections. When the model progresses through the task, the hidden layer forms an implicit representation of the context. Thus, the GCM model simulates the learning process of action sequences, whereby the repetition of sequences leads to a distributed representation within the SRN, which ultimately renders the biasing from the supervisory system optional.

Through this architecture, multiple pathways may direct or contribute to action selection. The first, *direct pathway* (panel A of Figure 1.3), is the direct influence of the representation of the external environment. It reflects the affordances of Gibson as presented previously (Gibson, 1977). The second, *context pathway* (panel B), represents the influence of well-learned action sequences, which have been represented in the context layer. The third and last, *goal circuit route*, represents the influence of intention on action in the supervisory system.

The input from goal units (supervisory system) is particularly needed at subtask boundaries, or branch points, where the environment is insufficient to predict the successor action. It occurs by providing additional activations to the appropriate schemas.

The GCM model thus addresses how a routine (contention scheduling) system is able to produce complex action sequences autonomously in sufficiently well-learned situations, while being supported by a supervisory system when needed.

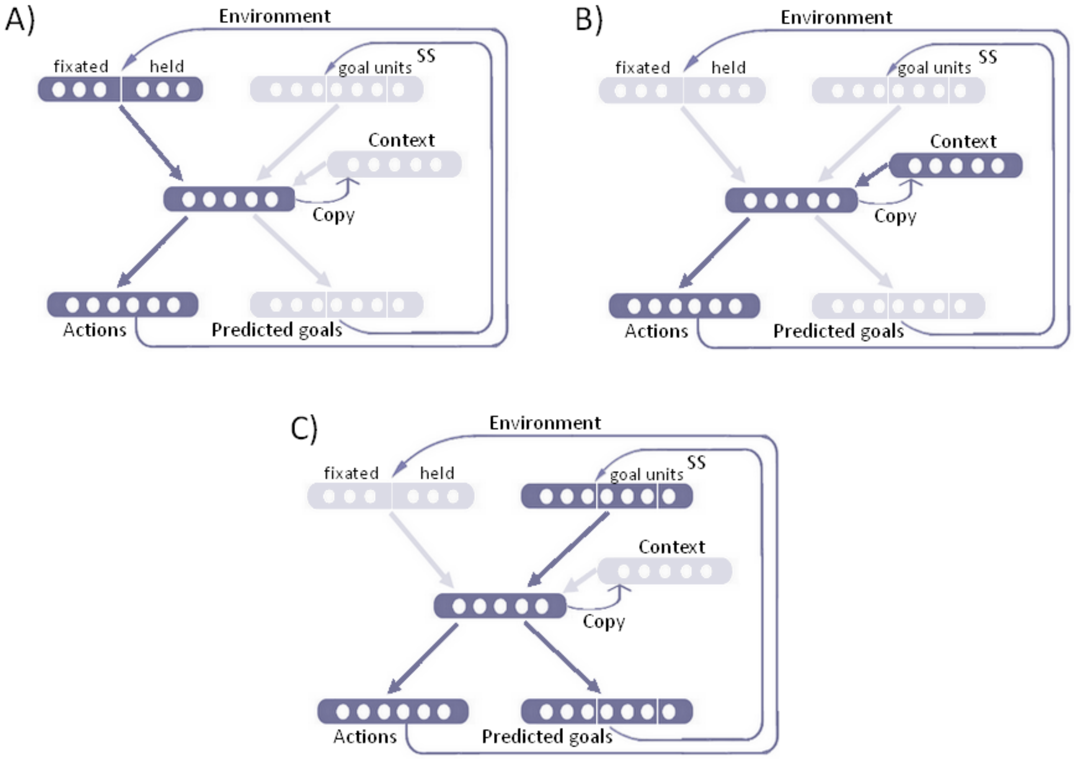


Fig. 1.3: **Architecture and different pathways of action control within the Goal circuit model. Panel (A) illustrates the direct pathway, (B) the goal pathway and (C) the goal circuit. Figure adapted from Cooper et al. (2014).**

1.3.2 Development of action control and development of Executive Functions (EF)

1.3.2.1 The role of supervisory functions in action control

The supervisory functions involved in the Norman and Shallice (1986) account are held to be located in the prefrontal cortex. Indeed, disorders of action selection mentioned previously such as the utilisation behaviour (Lhermitte, 1983) and anarchic hand syndrome (Goldberg et al., 1981) have their roots in prefrontal damage. The prefrontal location of supervisory functions is a view widely accepted (see Burgess and Stuss (2017) for a recent review). A range of supervisory processes are involved in the control of action, we summarise the ones that are relevant given their developmental trajectories.

A key role attributed to the prefrontal cortex, and similarly a central supervisory function, is the inhibition of prepotent responses (e.g., Perret, 1974). Following this view, behaviours triggered by the environment or over-learnt routines must be inhibited to allow the selection of appropriate actions.

Working memory has been postulated to be crucial in carrying out routine sequences, because it allows one to maintain information related to the goal in mind throughout the operation of an action sequence (e.g., Goldman-Rakic, 1987). Nonetheless, the goal maintenance may not only rely on working memory per se, as suggested by the phenomenon of *goal neglect* where one may have knowledge about the rule and the appropriate response but fail to execute it (Duncan et al., 2008). The development of inhibitory control and working memory will be tackled in the next Section.

Attention is a critical component of supervisory functions, albeit its definition varies according to context. It plays an essential role in monitoring for potential errors and thus in error correction mechanisms. In the GCM model (Cooper et al., 2014), the attentional system and supervisory system are considered interchangeably. Attention can also be taken in its broader sense, i.e., any task-direction information pick-up (Schwartz, 1995). To investigate the development of attentional networks, Fan et al. (2002) have divided attention into networks in charge of alerting, orienting and executive control. Such networks appear to have independent developmental trajectories, with the alerting network changing between the tested range of 6 to 9 years old, and further until adulthood, while the orienting and executive networks did not appear to change in that range (Rueda et al., 2004).

An influential partition of the supervisory system (e.g., Shallice et al., 1996, Shallice et al., 2008) identifies the following separable processes: goal generation, strategy generation, and monitoring and checking. Supervisory functions are also grouped and referred to as Executive Functions (EF). The development of such processes has been widely studied under the lens of executive functions, which will be used in the rest of the thesis.

1.3.2.2 Executive functions and development

A partition of EF into switching, working memory and inhibition components, stemming from the work of Miyake et al. (2000b), is widely used in the literature, including the developmental literature. Miyake et al. (2000b) picked the most commonly used executive functions among tasks that seemed to tap into independent components as much as possible: switching (shifting of mental sets), working memory (monitoring and updating of working memory representations) and inhibitory control (inhibition of prepotent responses) to verify how separable the components actually were.

Miyake et al. (2000b) selected multiple tasks that have ben postulated to involve the functions of interest and applied two complementary analyses. The first, a confirmatory factor analysis, aimed at extracting the common latent variables behind tasks, held to represent the executing functions and further measure the degree of commonalities among executive functions. The second, a structural equation modelling analysis, aimed at testing whether the tasks were actually measuring the executive functions they were intended to measure. The fourteen tasks used were the original tasks, or adaptations of: the antisaccade (Roberts et al., 1994), number-letter (Rogers and Monsell, 1995b), keep track (Yntema, 1963), stop-signal (Logan, 1994), local-global (Navon, 1977), Stroop (Stroop, 1935), and letter memory (N. Morris and Jones, 1990), plusminus (Jersild, 1927), tone monitoring (Larson et al., 1988), operation span (Turner and Engle, 1989), Random Number Generation (RNG), (Tower of Hanoi, TOH; Humes et al., 1997), dual-task (the Maze tracing speed test, Ekstrom et al., 1976 and a word generation task), and Wisconsin card sorting task (WCST; Kimberg et al., 1997). The results have shown that the switching, WM and inhibition components could indeed be separated, but not perfectly. They have also shown that executive functions are not uniquely composed of these three components.

EF components and development

A line of research stemming from Miyake et al.'s (2000b) confirmatory analyses has observed that early on, EF components are indistinguishable, and that through development the EF components gradually differentiate from each other (e.g., Brydges et al., 2014). However, research is not clear yet about the precise age when this transition operates. Using 9 different tasks adapted from the original set of tasks of Miyake et al. (2000b), Brydges et al. (2014) observed that the separability, from one unitary EF to a two-factors (inhibition/shifting)EF model, increased between 7 and 9 years old. Shing et al. (2010) used 5 different tasks and have found inhibitory control and memory maintenance to separate later in 9.5- to 14-year-olds compared to 4- to 7-year-olds. Consistent with this finding, Lehto et al. (2003) found 8-to 13-year-olds to present distinguishable (yet still correlated) inhibition, working memory and shifting components, using 9 tasks. However, Klauer et al. (2010) found inhibition and working-memory to be hard to separate.

Although it is difficult to get a definitive answer about the separability, and nature of the different developmental trajectories for each component with the methodology of Miyake et al. (2000b), it is interesting to look at the trends of the developmental trajectories. Such trends can be examined by pulling together numerous single task studies. A meta-analysis of 401 studies (T. Maldonado et al., 2020) found that all tested subcomponents of EF (inhibition, shifting, and updating, following the University-Diversity model of Friedman and Miyake, 2017) varied equally with age except from updating.

When looking at the school-age range more specifically, Carver et al. (2001) measured inhibition with a stop-signal task (where one should respond to a stimulus, or withhold responding to the same stimulus when it changes into a *stop* stimulus after a more or less long delay, sometimes after response initiation). They found improvements in inhibitory control before 5 years old as well as between 5 and 7 years (but have not tested beyond this age range). Mehnert, Akhrif, Telkemeyer, Rossi, Schmitz, et al. (2013) assessed inhibitory control changes throughout the school years and beyond, in relationship with protracted frontal cortex development, with a *Go-NoGo* task. In this task, a prepotent response is built by successive trials with stimulus requiring a response, typically a key press (*Go trials*), followed by a mixture of *Go trials* and *NoGo trials*, the latter requiring to withhold the response. The 4 to

6 years old range was the most sensitive period for the development of inhibitory skills. Such improvements continued gradually until 7 years old, and to a lesser extent throughout adolescence. Macdonald et al. (2014) found improvements in inhibition with one of three versions of a Stroop task, but not the others, between 5 and 7 years of age, with performance plateauing at 8 years old.

Nevo and Breznitz (2013) found improvements in working memory between a 5-6 and a 6-7 years old groups, with 11 tasks (9 from the Automated Working Memory Assessment test suite of Alloway et al. (2008), that includes phonological and visuo-spatial working memory tests, and 2 episodic buffer tests). Michalczyk et al. (2013) also found gains in working memory from 5 to 12 years old, with three different tasks (phonological, visuo-spatial and central executive). Let us note that if one considers memory span broadly, steep increases are seen between 4 and 8 years old, slower increase until 12 years old after which they are thought to plateau at adults' level (Gathercole, 1999).

In a *task and response set*-switching task, 10- to 12-year-olds had less switch cost than 7- to 8-year-olds (Crone et al., 2006). Cragg and Nation (2009) found improvements in switching between decisions rules when comparing 9- to 11-year-olds to 5- to 8-year-olds. Finally, Huizinga and van der Molen (2011) found 7-year-olds to have less good task-switching abilities than 11-year-olds.

What is clear is that executive functioning as a whole significantly improves throughout childhood until early adulthood (Karr et al., 2018, T. Maldonado et al., 2020). This is very plausibly linked to protracted prefrontal cortex development (Casey et al., 2005a, Lenroot and Giedd, 2006, Mehnert, Akhrif, Telkemeyer, Rossi, Schmitz, et al., 2013).

EF tested in relationship with action sequences

A series of studies have related performance on routine actions to standard EF measures such as inhibition, working memory and switching components to better understand which specific top-down processes are critical in children's action control.

In a study where 4- to 6-year-olds repeated sequential actions such as making toasts, children who had higher switching ability as measured by the advanced Dimension Card Change Sort (DCCS; Chevalier and Blaye, 2009) were more able to recover from an end-of-task interruption (Yanaoka and Saito, 2019). Such a link between action control and switching as measured by the advanced DCCS were partially replicated in Yanaoka and Saito (2020). This study had similar action sequences as Yanaoka and Saito (2019), performed by 4- to 6-year-olds, and in addition a condition providing visual reminders to test how the reminders would impact the interruption recovery. Interestingly, they found that reminders erased individual differences in goal maintenance capacity before a branch point.

In a *Lego* house-building task, Schröer et al. (2021) found that 3- to 5-year-olds with higher updating skill (one of the scores of the auditory reverse digit span task; Carlson et al., 2002) were more able to follow the subgoal order of the sequence.

Schröer et al. (2021) propose that updating skill may help maintain the ongoing subgoal in mind as well as maintain the temporal information about the previously executed steps and access to the subsequent action. In the same study, children with higher inhibitory control, as measured by a variant of the Go-NoGo task (the BAT tasks, Kaller, Rahm, Spreer, Mader, and Unterrainer, 2008), were less vulnerable to distractors in their sequence execution. Working memory, as measured by the main score of the auditory reverse digit span task (Carlson et al., 2002), and set-shifting were not found to correlate with the action planning measurements of the study.

In a naturalistic everyday routine tasks, Perone et al. (2020) measured individual differences in EF with the Minnesota Executive Function Scale (Carlson and Zelazo, 2014) which is built around the DCCS task but with an adaptive level of complexity. Individual differences in EF in 6- to 10-year-olds predicted some accuracy features regarding actions order, as well as the correctness of the incorporated items, above and beyond age (Perone et al., 2020). The authors interpret the importance of EF to maintain in mind instructions and subgoal information, as well as to ignore distractors when executing routine tasks.

Taken together, the developmental findings indicate important changes in executive functions (in particular, inhibitory and switching components) throughout the school years that may explain changes in sequential action skill.

As discussed previously, it is to some extent possible to measure distinct developmental trajectories for distinct components of supervisory or executive functions (Miyake et al., 2000b). The separation can be compromised by the strong interdependence between components, which can be overcome with a large number of tasks administered together. We will occasionally consider the particular role of a given component (e.g., inhibition) with a 'one task, one component' approach, given the widespread investigation of the three components via unique tasks in the developmental literature, and because looking at separate processes gives a better understanding of the developmental mechanisms at play. Nonetheless, we acknowledge the limitations of separating components by using single standard tasks, and our primary interest and conclusions will be the development of executive functions as a whole.

1.3.3 Reinforcement Learning (RL)

1.3.3.1 General reinforcement learning framework

A last line of research contributing to the understanding of sequential action control is the one of reinforcement learning (RL; Sutton and Barto, 1998). Reinforcement learning has been extensively used to study reward-based learning. It has been successful in generating new computational theories of value-based learning and choice and has identified neural correlates of the underlying components (see Daw and Doya (2006) and Dayan and Niv (2008) for reviews). Although RL is traditionally framed in terms of learning and decision-making, it is equally useful to study action selection. First, even if our focus is not on learning itself nor the generation of novel sequences, the framework remains relevant because it also tackles the control of sequences that have been previously learnt. Second, the same set of systems are responsible for making a decision or choosing an option (in decision-making terms) and for selecting an action. In fact, what studies refer to as *making a decision* (or choosing an option) is analogous to *selecting an action* because choosing involve making the choice apparent (whether by pressing the corresponding key and thus selecting the action 'press key x', or by vocalising the response, or by performing any action that enacts the choice).

RL describes how an agent explores an environment and learns by trial-and-error to select the actions that will lead to rewards and avoid punishments. The RL accounts used in decision-making traditionally involve three steps: predicting the values of candidates' actions, selecting the action that maximises the predicted values, and learning from experience to improve future predictions. A RL environment is formally defined as follows: an agent can find itself within a set of states S; at each state, the set of actions A is available, actions lead from one state to another as defined by the transition function T, and after each transition an outcome or reward R is obtained. The agent has the goal of finding a policy π that maximises its cumulative long-term reward. The policy is a function that describes the probability of taking each possible action in every possible state. To do so, the agent predicts an action value, that isn the sum of expected future rewards that would follow each action taken in the state it finds itself. The value of action a in state s is denoted $Q(s, a)$.

Although the agent may select the action that maximises $Q(s, a)$, animal and human choices are commonly modelled by a *softmax* decision rule. This rule implements a choice guided by a mixture of the maximal $Q(s, a)$ and some randomness. The mixture is such that the highest valued action is given the highest selection probability and the other actions are weighted according to their valuation. Accordingly, action a is chosen in state s with the probability:

$$P(a \mid s) = \frac{\exp \beta \cdot Q(s, a)}{\sum_{a'} \exp \beta \cdot Q(s, a')}$$

where β is the inverse temperature parameter. Decreasing the value of β leads to more equiprobability of choice among actions.

A common implementation of RL is the actor-critic approach. In such architecture, an agent is divided into an actor and a critic part, as shown in Figure 1.4, panel A. The actor follows the policy, denoted $\pi(s)$, while the critic stores a value function denoted $V(s)$ that estimate the cumulative long-term reward that can be expected after visiting the state s. Both the policy and the value functions are learned from experience. Notably, after collecting a reward, the critic computes the prediction error δ. At the neural level, a central finding is that midbrain dopaminergic neurons seem to encode the reward prediction error δ (Schultz et al., 1997, Berns et al., 2001, Haruno et al., 2004).

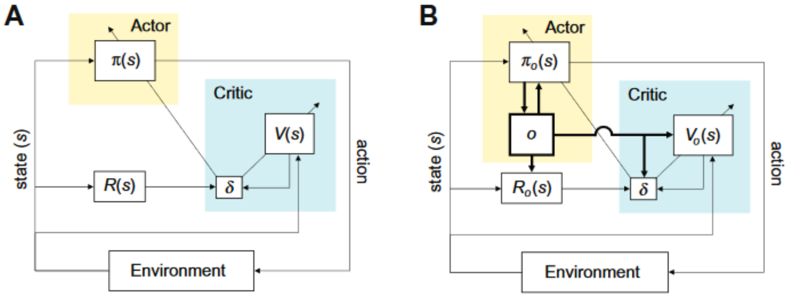


Fig. 1.4: **Actor-critic implementation of standard RL (panel A) and HRL (panel B). PI: policy, R: reward function, V: value function, δ: reward prediction error. O: option being executed. The o subscript indicates an option-specific function. Image reproduced from Botvinick et al. (2009) with permission of the rights holder, Elsevier.**

1.3.3.2 Model-free and model-based classes of reinforcement learning

We will consider two classes of reinforcement learning, model-free and model-based RL, that operate the value optimisation in different ways. With model-free RL, an agent estimates the action values from experience and stores them in memory. The agent does not form a representation of the transition function T. In the model-free Q-learning algorithm (Watkins and Dayan, 1992), this is formalised as follows. The value of action a in state s estimated by a model-free agent $Q_{MF}(s,a)$ is updated at each stage proportionally to a reward prediction error δ via a certain learning rate α:

$Q_{MF}(s,a) = Q_{MF}(s,a) + \alpha * \delta$ with: $\delta = r + \max_{a' \in A} Q_{MF}(s',a') - Q_{MF}(s,a)$

where r is the outcome (reward) obtained and s' the subsequent state.

In contrast, a model-based agent both learns from experience and constructs an internal model of the transitions between states and actions (referred to as internal model). The agent selects action by using this internal model of the environment. In general terms, this is formalised by considering the reward function R (i.e., continuously updated by trial-and-error) and the transition function T:

$Q_{MB}(s,a) = \sum_{s'} T'\,(s,a,s') * (R'\,(s,a,s') + \max_{a' \in A} Q_{MB}\,(s',a')$

1.3.3.3 Model-free / model-based reinforcement learning approaches mapped onto habitual / goal-directed behaviour

An influential body of work has considered two forms of control working together, habitual and goal-directed control, and related them to the model-free and model-based RL algorithm (Daw et al., 2005). The notion of *goal-directed* control used within the RL framework is used to oppose to habitual, and differs from the general definition of goal-directed action introduced in Section 1.2.2.

The precise separation and definition of habitual and goal-directed behaviour varies and is reviewed in Dolan and Dayan (2013). This thesis will consider the following definition. Habitual control (modelled by model-free RL) consists in trial-by-error learning where a reward expectation guide choices, and is updated following the discrepancy between an obtained outcome and an expected reward (that is, a reward prediction error), which in turn guides subsequent choices. Goal-directed control (modelled by model-based RL) uses a more sophisticated evaluation process to guide choices. It assumes the use of an internal model of the environment (e.g., how actions prior to the final 'outcome action' are connected to states which have a higher or lower probability of reward) and computes prospective evaluation of all the possible paths within that model of the environment to select the action that will later lead to the best outcome. The search throughout the path of the internal model has been compared to mental simulation (e.g., Doya, 1999, Schacter et al., 2012).

Following the distinction between habitual and goal-directed control, studies have looked into the interaction or competition between the forms of control. Indeed, it has been found that even simple tasks incorporate both types of control to some extent (Collins and Frank, 2012). To evaluate the relative involvement of model-free and model-based types of control, a "two-stage stage" was introduced by Daw et al. (2011). The task will be described in Chapter 3. The task has been extensively used to characterise behavioural implications and influences over such modes of control in healthy adults (e.g., Eppinger et al., 2013, Otto et al., 2013, Otto et al., 2014, Kool et al., 2017a, Keramati et al., 2011b) and clinical populations (e.g., Sharp et al., 2015, Voon et al., 2014, Voon et al., 2015, Gillan et al., 2016).

There is evidence that the two forms of action evaluation used for model-free and model-based types of control are subserved by distinct neural systems. The prefrontal cortex is held to be used to evaluate the potential consequences of actions (Owen, 1997), as used in model-based RL. The dorsolateral striatum is held to support stimulus-triggered responses (Yin et al., 2004) as used in model-free RL. Additionally, brain imaging studies found correlations between dorsal and ventral striatum regions and the reward prediction error in model-free learning (Berns et al., 2001, Haruno et al., 2004).

1.3.3.4 Hierarchical reinforcement learning

A challenge posed by traditional RL methods is that they do not scale well. That is, when the task domain (i.e. the set of states and/or actions) becomes too large, the time required to reach a stable policy can become excessively large to be feasible. Hierarchical reinforcement learning (HRL) is an approach that tackles this issue. It has been proposed that the brain might deal with the scaling issue in an analogous way (Botvinick et al., 2009, Botvinick, 2012).

In HRL, the RL framework is augmented with *options* (Sutton et al., 1999). Options are a set of interrelated actions that are grouped together under a higher-level action (for example, the higher-level action *buttering toast* groups together the actions *grasping a butterknife*, *collecting butter* and *spreading it onto the toast*). Options thus carry out temporal abstractions over a sequence of lower-level actions.

An actor-critic implementation of HRL is shown in Figure 1.4, panel B. It extends the standard RL actor-critic architecture from panel A. First, it adds the representation of the option o currently maintained by the actor. The actor also maintains a separate policy for each option π_o. The prediction errors are no longer computed after each action, but at the end of an option. Therefore, the actor must send the critic the information of when an option termination occurs. Finally, to compute the prediction error for an option, the critic must also keep track of the rewards obtained throughout the option execution, as well as the state in which the option was initiated.

The selection of options, and the execution of actions within an option may be linked to the dual-systems framework by considering that the start of an option is like a branch point, thus requires supervisory influence, while the execution of individual actions within an option can be carried out by the routine system, without extra supervisory influence. Relatedly, a computational problem to solve is the discovery and acquisition of useful options (i.e., what actions binds together to form an option), as highlighted by Botvinick et al. (2009). Solving the problem in a general case is beyond the scope of the thesis. Nonetheless, in a simple task with a hierarchical structure, we can hypothesise that what constitutes an option may be discovered by using transition probabilities between actions: actions that follow each other with certainty or very high probability may be bound together to form an option, while low transition probabilities would constitute the end of an option.

Although to our knowledge HRL has not been used to study development, the options as defined in HRL resembles the subroutines and the modularisation process proposed by Bruner (1973).

1.3.3.5 Reinforcement learning and development

Applications of the reinforcement learning framework across development have focused greatly on the exploration/exploitation balance (e.g., Schulz et al., 2019) or on the learning side. For example, it has investigated the sensitivity to negative or positive reward prediction error (e.g., Bos et al., 2012), or social learning (Bolenz et al., 2017). Additionally, a reinforcement learning analysis of the Tower Of London (TOL; Shallice, 1982) task suggested that the worse performance of 3- to 4-year-olds compared to 5- to 6-year-olds may be due to a greater tendency of younger children to be influenced by the perceptual similarity of an ongoing state compared to the goal state, at the expense of efficient planning (Mitsopoulos et al., 2015).

Despite the success of the approach of the model-free/model-based distinction, there have been few studies tackling the development of the two types of control and their interaction in children. One of them, which will be discussed extensively in Chapter 3, found overall that model-based use increases with age from 8 years old to adulthood. Eppinger et al. (2013) have found that in adults, model-based choice decreased with age, even after controlling for working memory which also decreased with age. Together with neural evidence (e.g., Eppinger et al., 2012, Frank and Kong, 2008), the authors suggest that age-related changes are linked to under-recruitment of the PFC for integrating the expected rewards into model-based decisions.

Janacsek et al. (2012) studied the development of implicit learning across the lifespan, between 4 and 85 years of age, and related it to the model-free / model-based RL framework. The authors used an adjusted Serial response time task, with probabilistic transitions of both high and low frequencies. The rationale is that implicit statistical learning based on raw probabilities (such as measured by the task) may be important in developing new abilities and skills (Hikosaka et al., 2002, Keele et al., 2003). Janacsek et al. (2012) found that the sensitivity to raw probabilities decreases after 12 years old. The authors connected the finding to the development around the same age of cortical areas underlying the capacity to build internal models, which are the basis of model-based learning (Giedd et al., 1999, Blakemore and Choudhury, 2006).

Based on this, they propose that a developmental shift occurs around 12 years of age (see Figure 1.5). Before the shift, skill learning relies mostly on the detection of simple raw probabilities while after the shift, it relies more on complex internal models. Accordingly, this shift is driven by the experience-driven maturation of the model-based system, which makes possible the efficient use of internal models, and simultaneously accompanied by a neglect of model-free computations. The view of Janacsek et al. (2012) matches closely the one of Decker et al. (2016), whereby a shift between a model-free-dominant strategy to model-based-dominant strategy occurs between a group of 8 to 12 years old children and adolescents.

1.3.4 Two diverging views of the interaction between two systems for the control of action

In the picture depicted by Janacsek et al. (2012) and Decker et al. (2016), and summarised on diagram 1.6 (A), model-free and model-based types control are competing, and development occurs not only via maturation of the model-based system, where more complex internal model building abilities drives increase in model-based use, but also via the decrease of the model-free involvement (e.g., via a decrease in the ability to use raw probabilities as suggested by Janacsek et al. (2012)).

The dual-systems account, proposed by Norman and Shallice (1986) and implemented in the GCM (Cooper et al., 2014), uses a similar dichotomy between a more implicit, procedural type of system (routine) and a system with explicit representations of goal (supervisory system). Nonetheless, the functioning of the two systems as proposed by Norman and Shallice (1986) is not a competition but rather an experience- and action-dependent combination of involvement of each system. The routine system is held to be able to carry highly-practised sequences almost autonomously, and benefit from biasing of the supervisory system when needed

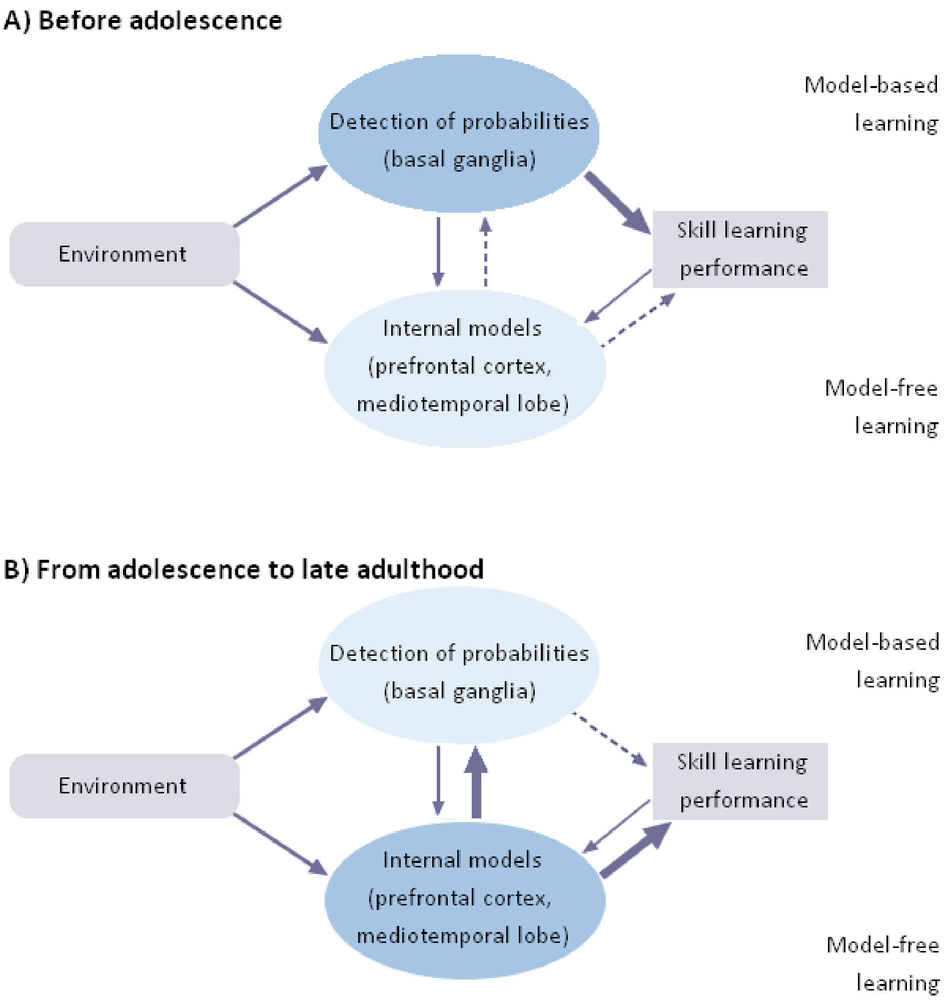


***Fig. 1.5:* Developmental shift in the use of model-free and model-based types of skill learning, proposed by Janacsek et al. (2012). (a) Before adolescence, skill learning is mostly driven by detection of raw probabilities, defined as model-free learning. (b) Later in adolescence, internal models have developed, making possible the use of more complex relations from the environment, resulting in model-based type of learning. Adapted from Janacsek et al. (2012).**

(e.g., at branch points where many compatible actions compete for selection, when the transition probability in the sequence is low). The view is summarised on diagram 1.6 (C).

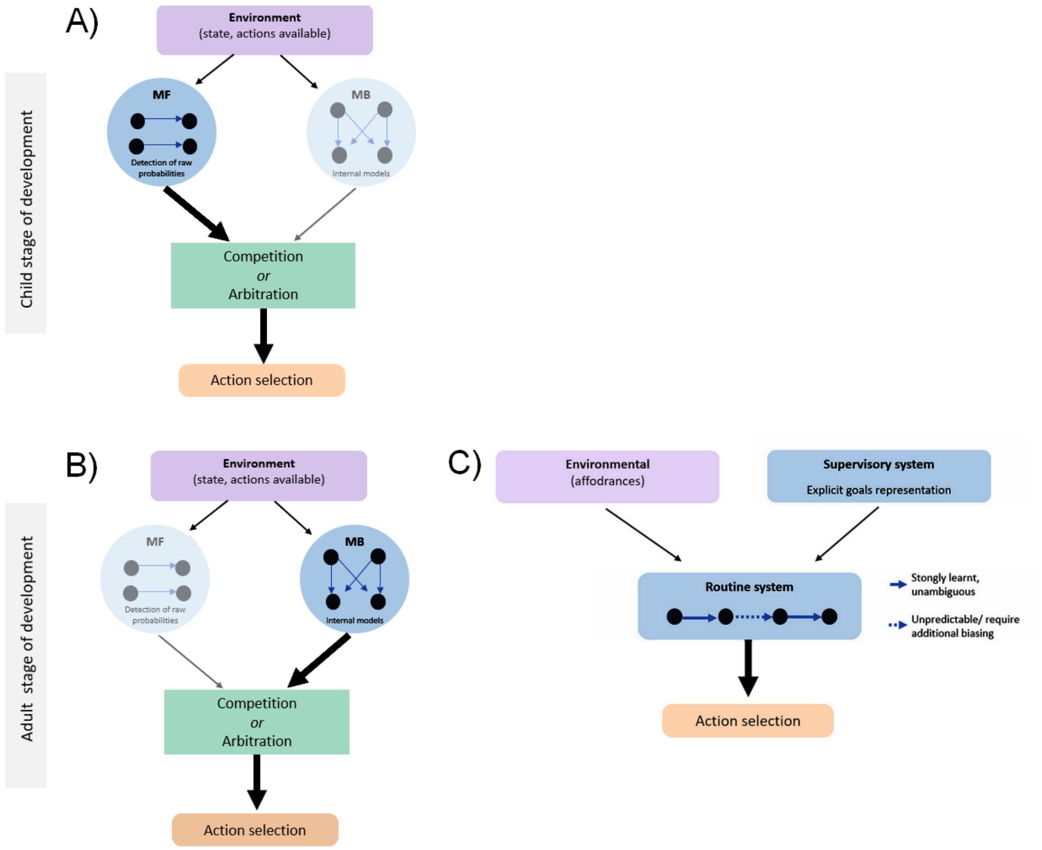


Fig. 1.6: **Summarised view of two systems interaction in the model-free/-model-based reinforcement learning framework (panel A and B) and in the dual-systems framework (Norman and Shallice, 1986, Cooper et al., 2014). The bottom panels (B and C) represent the adult stage of development, while the top panel (A) represents the child (before 12 years old) stage of development proposed by Janacsek et al. (2012). In the view proposed by Janacsek et al. (2012), there is a developmental progression where children rely mostly on the model-free system, and adults presumably rely more on model-based system after the necessary component for model-based computations have developed, while model-free type of computations may regress. The dual-systems framework only specifies the adult stage, where the two systems work in cooperation, with the supervisory system biasing the routine system at the points where it is most needed.**

The HRL view is most similar to the latter collaborative view. Indeed, in that HRL view the selection of options may be related to the selection and maintenance of an overarching goal at punctual points (branch points), like in the supervisory system of Norman and Shallice (1986), while the sub-actions within an option can be carried out by the other system.

A last distinction between the existing accounts is that the GCM account includes a direct pathway of environmental influences on action, but this is missing from the model-free/model-based reinforcement learning account. In the latter, the environment only indicates what options are present in a given state but it does not influence the action selection mechanism.

1.4 Methods

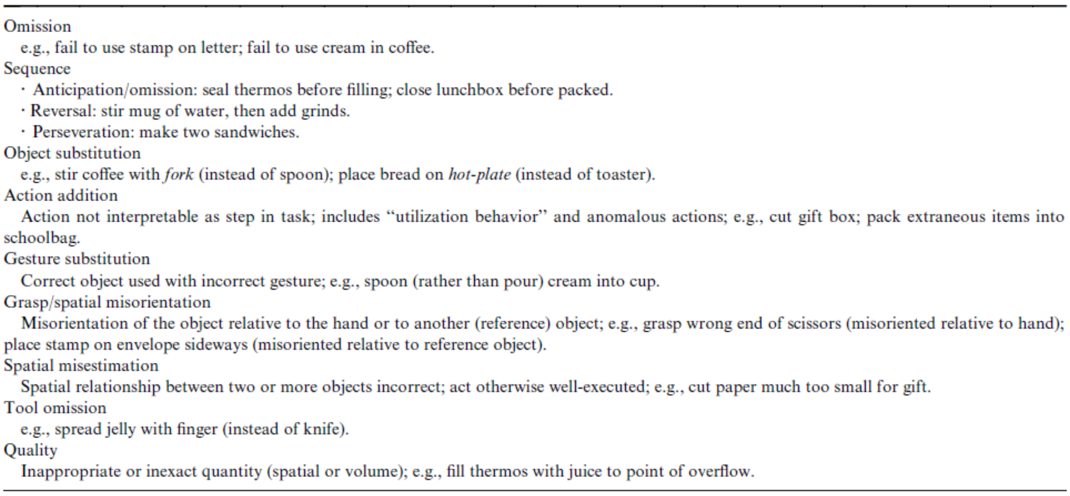
1.4.1 Paradigms to study action control in adults and children

The investigation of routine action control in children has focused on different levels than the level of action we are concerned with, or on the relevant level of action but with very little practice (whether in real-life or in the lab). For example, Yanaoka and Saito (2019) had children make toy toasts (for a mouse or a cat character) for 4 practice trials and 12 test trials. Although the study led to interesting insights (presented in Section 1.3.1.3), 4 practice trials are not sufficient to study routinised action control. Perone et al. (2020) had children perform in the lab some routines typical of daily life (making a sandwich, packing lunch, packing homework, and packing snow gear). However, children performed each task only once and the study focused on information gathering and prospective memory (from the instruction phase) as well as the link with executive functions, rather than the level of action control we are interested in.

In adults, reports of real-life errors from daily routines of healthy participants, or observed in neurological patients, have been among the firsts methods of investigation. Errors in daily-life executions of routine action systems provide insights into the operation of the system. Common routines such as preparing coffee are often taken as a subject of study. For example, objects errors (such as adding butter into coffee), as observed by brain-damaged patients (Schwartz et al., 1991) or reported in healthy patients diaries (Reason and Mycielska, 1982), have contributed to the view that multiple action plans may be activated in parallel. An object error thus results from a blend of the arguments of an ongoing action plan. Relatedly, the Multi-level action test (MLAT) standardises the study of such routine action errors in an experimental setting (Schwartz et al., 1999). It consists in the primary tasks of making a slice of toast with butter, wrapping a present, and packing a lunch box; in the presence or absence of distractor objects. It classifies the errors according to taxonomy presented in table 1.7.

Giovannetti et al. (2007) studied routine performance by healthy participants of a real-life coffee-making task containing 18 steps, over 12 experimental trials. The rate of errors was manipulated by adding time pressure and distractors objects. The task appeared promising to study naturalistic action, nonetheless it was limited to 10 practice trials which may not be enough to tap into routine performance. Arnold et al. (2017) approached the question of hierarchical action control by a Lego blocks building task. Participants were asked to learn to build and to build two block walls that had similar and dissimilar chunks. The start of a dissimilar chunk constituted a decision point thus allowed measuring differences in action selection latency at decision compared to no-decision points.

The study of Ruh et al. (2010) was designed to provide direct empirical support to the dual-systems account of routine action, by analysing errors and response times on a very large number of trials of a well-routinised task performed by real adult



Omission
e.g., fail to use stamp on letter; fail to use cream in coffee.
Sequence
- Anticipation/omission: seal thermos before filling; close lunchbox before packed.
- Reversal: stir mug of water, then add grinds.
- Perseveration: make two sandwiches.

Object substitution
e.g., stir coffee with *fork* (instead of spoon); place bread on *hot-plate* (instead of toaster).
Action addition
Action not interpretable as step in task; includes "utilization behavior" and anomalous actions; e.g., cut gift box; pack extraneous items into schoolbag.
Gesture substitution
Correct object used with incorrect gesture; e.g., spoon (rather than pour) cream into cup.
Grasp/spatial misorientation
Misorientation of the object relative to the hand or to another (reference) object; e.g., grasp wrong end of scissors (misoriented relative to hand); place stamp on envelope sideways (misoriented relative to reference object).
Spatial misestimation
Spatial relationship between two or more objects incorrect; act otherwise well-executed; e.g., cut paper much too small for gift.
Tool omission
e.g., spread jelly with finger (instead of knife).
Quality
Inappropriate or inexact quantity (spatial or volume); e.g., fill thermos with juice to point of overflow.

Fig. 1.7: **Summary of the error taxonomy used in the Multi-level action test in Schwartz et al. (1999). Reproduced with permission from Elsevier.**

participants. A first version of the task consisted of a computerised version of a coffee-making task, like the initial simulations of the model of Cooper and Shallice (2000). Nonetheless, the coffee-making task suffered some limitations that impaired interpretability, so we will focus on the second version of the task, virtual gardening. The task implied drag-and-dropping icons on a screen with the computer mouse to achieve actions (e.g., moving a tool to stir a mixture). The task had a variant of six overarching goals that were achieved by an ordered sequence of 6 subgoals, each subgoal constituted of several steps. Each resulting sequences had a total of 48 to 69 steps. The set of sequences reproduced key features of real-life actions sequences, such as a set-up subsequence at the start and a clean-up subsequence at the end of each full sequence. Additionally, some actions constituted branch points: the task involved variants of sequences (with overlapping subsequences followed by diverging actions) such that, at a diverging point, the transition to the next action was ambiguous. These were compared to non-branch points, comparable action after which there was only one legitimate option. Together with the use of a task taxing supervisory resources, analyses of response time and errors brought support to the dual-systems framework, for example with longer latencies at branch points, as well as further slowing down by a secondary task, indicating local modulation of action control by the supervisory system.

The design of the Ruh et al. (2010) inspired the last experimental study of this thesis. The advantage of a computerised task such as Ruh et al. (2010) is the elicitation of a substantial amount of data on highly-practised action performance, which is hard to obtain in other situations. The main limitation is the artificial nature of the task. Nonetheless, results obtained via this computerised task are consistent with the types of errors obtained in laboratory studies with real-objects (e.g., Giovannetti et al., 2007) or in real-life diary reports (e.g., Reason, 1979).

Given the scarcity of studies that may or may not support the dual-systems theory in child action control (as opposed to adult action control), there is a need to test the appropriateness of the dual-systems approach to study development. This must be done by striking the right balance between the ecological nature and the controllability of a paradigm to measure and test precisely predictions from the dual-systems theory with methods that have been validated in adults.

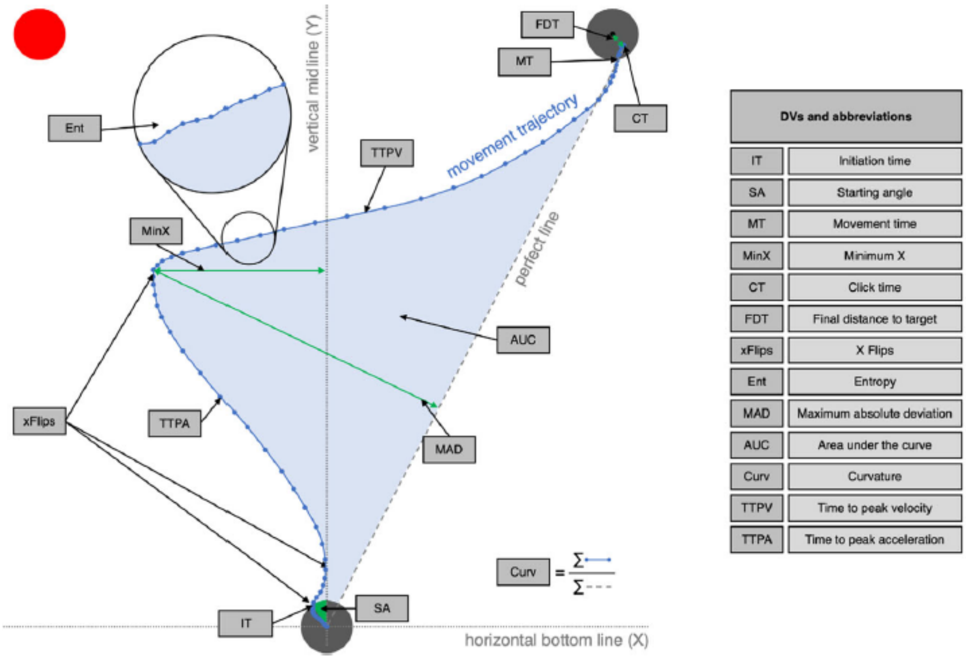
1.4.2 Kinematics

Hand-movement tracking methods (mouse-tracking, touchscreen finger-tracking, 3D Motion capture etc.) have been of increasing interest in the psychological literature in the past decades, possibly facilitated by the rise of open-source software that democratised experiment development and analyses, in particular online with mouse tracking (for example *MouseTracker* Freeman and Ambady, 2010; or *mousetrap* Kieslich and Henninger, 2017). Numerous methodological studies (Hehman et al., 2015, Kieslich et al., 2020, Schoemann et al., 2019, just to name a few) presented important considerations and recommendations about design such as starting procedures (self-paced or not), target sizes, mouse hardware setup (sensitivity), response type (hover vs. click), stimulus positions (centred vs. edged), and so on. However, to our knowledge there is no consensus on what approach is the best across all experiments, nor there is a standard that most researchers follow.

While some kinematic analyses are constrained or emerge naturally from a specific paradigm (e.g., Dotan et al., 2018), others have a wide margin of manoeuvrer. Part of the problem is idiosyncratic to studying movement in 3D space. Indeed, in each experimental paradigm the specific start points, end point, trajectories etc. will differ (we will call this problem *spatial specificity*). The diversity of approaches can be explained by two challenges of such data: its spatial specificity and its high-dimensionality. Kinematics is highly-dimensional data because it includes x, y (and sometimes z) coordinates at each sampled point of a trial and several derivatives (commonly speed and acceleration). Another difficulty is that that kinematic analyses must be adapted to each task, since the spatial arrangement of the stimulus and responses will affect how the trajectory can vary (e.g., how much space there is, whether there is there a single optimal path, whether there are there obstacles, etc.) and what is the volume in which actions are to be performed, and thus the kinematics data itself.

A widespread approach, found applicable across various studies, is what we call *single feature approach* where particular measures such as ’maximal deviation from a straight path’ are computed on trajectory, speed or acceleration data (for example, see Figure 1.8 reproduced from Wirth et al., 2020).

With this approach, a broad type of measures is commonly used for similar constructs but there is notable heterogeneity in the choice of precise measure definition for similar constructs. For example, for constructs related to co-activation of multiple targets, measures of curvature are often used, but the features can be: the area under



DVs and abbreviations	
IT	Initiation time
SA	Starting angle
MT	Movement time
MinX	Minimum X
CT	Click time
FDT	Final distance to target
xFlips	X Flips
Ent	Entropy
MAD	Maximum absolute deviation
AUC	Area under the curve
Curv	Curvature
TTPV	Time to peak velocity
TTPA	Time to peak acceleration

Fig. 1.8: **Multiple kinematics features commonly used in the psychology literature analysing movement data. The grey circle at the bottom represents the start area, the red and grey circles at the top represent the target areas. Figure reproduced with permission from Wirth et al. (2020), licensed under CC BY 4.0.**

the curve (Wifall et al., 2017, Miles and Proctor, 2011), maximal deviation (Desmurget et al., 1997, Song and Nakayama, 2006 Erb et al., 2018), distance travelled in pixels (Dale et al., 2007), and proximity of a target relative to a distractor (Spivey et al., 2005). To get better cumulative evidence in the mapping between cognitive process and kinematic signatures, it would be desirable to use common features across tasks when investigating overlapping or similar processes. On the other hand, by limiting oneself to features defined in past lines of research, one finds oneself at risk of missing out on crucial information in their own paradigm (that would not be captured by features used in other paradigms). To our knowledge, within this *single feature approach*, there is no consensus on the approach for the features selection. This means that there is a heterogeneity of choice of feature for a cognitive construct, i.e., even if a feature is well-suited to a given single study, the choice of that feature may not allow one to get cumulative evidence for which cognitive signatures a feature is characteristic of. This also means that there might not be a set of best practices and caveats associated with the implementation of a given feature (although there exist general methodological guidelines such as Schoemann et al., 2019) and that one may miss factors that could influence a feature in one design but not another. For example, the value of a feature such as maximal angle may be dependent on the instructed speed of reply. There is also a problem of multiple comparisons if one explores too many features without a-priori selection. Given the high-dimensionality of kinematics data and the multiple ways in which one can imagine how kinematics may be affected by cognitive processes, it might be tempting to explore a large number of features but then eventually be ensure on the probability that a significant result is a false positive. Thus, without the existence of commonly accepted best practices, the *single feature approach* might suffer from features being affected by subtle design change or from the multiple comparisons issue.

Another type of approach was used by M. Maldonado et al. (2019) who used a *machine learning* classifier on mouse-tracking data. Specifically, they created a first 'artificial' experiment in which they induced a 'change of decision' by changing the cue that indicated a previous answer (that required for example clicking a box on the right-hand side of the screen) to a cue that indicated another answer (for example click on the left-hand side of the screen). They used kinematics data points (with dimensionality reduced to 13 principal components with a Principal Component Analysis) to train a supervised classifier (Linear Discriminant Analysis, LDA) that would indicate how likely data were to come from an artificial 'change of mind' or 'no change of mind' condition. They then collected data on their main study (looking at linguistic processing of double-negation) and applied the classifier to infer whether the experimental conditions looked rather like a 'change of mind' or 'no change of mind' condition. This approach was fruitful for the study given that the question related to whether processing of double negation included a 'change of mind'. However, the LDA does not describe the kinematics beyond their similarity to one or the other artificial experimental condition. The LDA classifier training on this dataset is also not usable for any experiment that does not have a similar design. This *machine learning approach* can thus be powerful but not generalisable.

We propose that the *machine learning approach* just mentioned and the *single feature approach* lie on two opposite sides of a spectrum. One extremity of that spectrum uses a maximal amount of information from the kinematics at the expense of generalisability and interpretability. The other extremity has generalisability and interpretability at the risk of not exploiting all the information captured by kinematics data. In this thesis, we will propose a novel approach hybrid between the *machine learning approach* and the *single feature approach* that attempts to maximize generalisability while taking advantage of a maximum of information in the kinematics data. In brief, this new approach consists in a first step in using a dataset such as a pilot dataset (or any dataset collected with the paradigm of interest, but on different participants than the target study) and implement a very large number of *interpretable* features (including features used in the literature). In a second step, it consists in deriving a dimension reduction technique to select the features that capture the most variance, and finally applying the pre-selected features on the target study.

1.5 Research questions and thesis overview

1.5.1 Summary of existing findings and open research questions

There is a significant amount of evidence for the early emergence of goal-directed actions as soon as the first year of life. Infant and child actions are highly influenced by the environment affordances – like adults, if not even more strongly. We also know that toddlers can anticipate a few steps ahead, and that by school age children are capable of planning and executing daily-life routines. Some work has addressed the learning of routines, or the interaction between habitual and goal-directed learning in school-aged children. We also know that the development of EF throughout childhood has a considerable influence on the development of many social and cognitive faculties.

Much less is known about the nature and mechanisms underlying sequential action control such as in daily-life routine sequences. Sequential action control has been extensively tested and theorised in adults, notably with the dual-systems framework of Norman and Shallice (1986); but little or even no previous work has formally investigated the existence and use of two systems of action control in children as defined by the dual-systems framework. Critically, there is evidence suggesting that the development of executive functions is involved in improvement in action control throughout childhood. However, the role of executive functions as measured by standard tasks and its relationship with the action modes from the dual-system framework has never been tested in children and is a central aim of this thesis. To address such theoretical questions, there is a need for new experimental paradigms. It would be desirable that such paradigms are able to link children's action control to theories of adult action control, in order to benefit from the extensive computational and empirical support about adults' action control theories. Developing such research methods is a parallel goal of this thesis.

Finally, this thesis aims to unify developmental findings obtained from the dual-systems perspective of Norman and Shallice (1986) and from the Reinforcement Learning (RL) perspective. Specifically, there has been developmental work on the trajectories of habitual and goal-directed control modes as defined in the RL tradition, but not from the dual-systems of Norman and Shallice (1986). This thesis also attempts to bridge the gap between dual-systems mode of action control and RL modelling and the related empirical data, by developing a mechanistic model that makes closer contact to child data.

1.5.2 Thesis overview

Chapter 2 investigates the role of environmental, bottom-up influences of objects when reaching for a target among distractors. It reveals that the process of resolving the competition from distractors when selecting a target is still developing throughout childhood, as evidenced by interference effects from distractor objects in 6 - 12 years old but not in adults, and possibly decreasing interference effects within this age range. The Chapter provides evidence that the development of inhibitory control abilities is crucial in the development of action control performance throughout school years.

Chapter 3 looks at action selection at a higher level. It proposes a novel computational model to account for the interaction between two systems of action control in children's and adults' repeated two-actions sequential selection. The two systems are modelled, following the reinforcement learning literature, as model-free and model-based types. The novel model is tested against pre-existing experimental data, and appears more informative than the original study's reinforcement learning model in that the novel model is able to make contact with response time data and consequently distinguish between various mechanistic accounts of the balance between model-free and model-based modes of control. Specifically, it suggests that the increasing use of model-based control from childhood (8-12 years old group) to adulthood may either result from a change in an externally-arbitrated recruitment balance, or result from a reduction in noise in the model-based system with age.

To further understand the development of the action control systems within the span of school years, the subsequent two Chapters rely on newly designed studies and newly acquired data.

Chapter 4 tackles the acquisition, performance and adaptability to unpredictable changes of short action sequences between 5 and 9 years of age. It uses a touchscreen-based task requiring the performance of action sequences and incorporating unpredictable environment changes (or, probabilistic transitions) requiring one to update an ongoing action sequence. The use of probabilistic transitions is inspired by the pre-existing reinforcement learning task modelled in the previous Chapter 3 (that used probabilistic transitions to dissociate between model-free and model-based control strategies). However, the task in Chapter 4 has been substantially adapted to measure the recruitment of routine and supervisory modes in a task tapping more concretely into life-like action control and suited to study development within the school-age years. Chapter 4 reveals that action sequence performance following certain types of transitions can be updated equally well across the age range, but that other types of transition cause a decrease in performance of the youngest age group. The findings also indicate that the monitoring and/or updating processes, which belongs to supervisory functions, undergo developmental changes between 5 and 9 years of age.

Chapter 5 pursues the investigation of routine-like action sequence performance, by 7- to 12-year-olds and adults, this time with longer action sequences that capture essential features of real-life sequences. It consists of two studies using a computer-based mouse-tracking task, with two sequence variants, containing actions that have particular irregularities (called branch points), and that are compared to more regular actions. It also uses a secondary task to tax supervisory resources and thus clarifies the role of supervisory control functions, which have been highlighted in previous Chapters, in longer action sequences. Chapter 5 finds that children have a higher cost than adults at the irregular sequence points, and that 7- to 8-year-olds might have a higher cost than 9- to 12-year-olds. Together with the result that the secondary task impairs performance more specifically at branch points, the chapter reveals that when children execute (partially) routinised action sequences, supervisory control is most needed for resolving the selection of actions at branch points. Furthermore, it finds that this ability to successfully engage supervisory control at the most crucial parts of a sequence improves during and beyond school years.

Chapter 6 discusses all the experimental findings and models together to form a comprehensive view of the complex interactions between various influences and modes of action control and their change across development. In this chapter, we also challenge the view that the routine (model-free) system becomes less important with development, or that the system loses ground on the competition because supervisory (model-based) abilities become more mature and prominent. We propose instead that, throughout development, the supervisory system does mature and consequently gains more importance in controlling skilled behaviour, but only at the levels where supervisory influence is needed; and furthermore that the routine system keeps playing an important role throughout life.

In summary, the thesis first lays out the key action selection challenges that children face when executing sequences – from the lower-level inhibition of task-irrelevant distractors to the appropriate selection of actions that are partly compatible with the context. Secondly, it provides an account of how two action control systems resolve action selection challenges and what their specificity is, notably in routinised action sequences. Thirdly, it describes changes in action control performance throughout the school-age years, and provides converging evidence towards the development of inhibitory control being an important driver of change in action skill. Finally, it proposes a broader view of the developmental trajectory of the interaction of the two systems supporting action control.

Chapter 2

Reaching in the Face of Distractors: A 3D Motion Capture Study

Abstract

This chapter investigates the susceptibility to distractor interferences when selecting an action, and changes in the extent of distractor interferences across development. It presents new analyses on previously collected kinematics data about the interference effects caused by distractor objects when reaching to a target in an ecological 3D environment.

It reveals that distractors, regardless of their semantic or perceptual properties, affect reaching duration and kinematics parameters in children between 6 and 12 years old, and affect the youngest children to a larger extent. The chapter suggests that mechanisms of action-based inhibition are recruited to resolve the competing activations generated in selecting the appropriate action, and that such mechanisms improve with age, supporting more accurate and efficient (less susceptible to interference) reaching actions throughout development. Therefore, it suggests that development of inhibitory control plays a central role in the improvement of action control skills throughout the school-age years.

2.1 Introduction

2.1.1 The importance of objects in action selection

Action can be studied from different perspectives. On the one hand, it is possible to study action by abstracting away the relationship between motor commands and the physical properties of objects, and focusing more on the hierarchical relationship between goals and subgoals. For that reason, action control can be investigated using computerised tasks (as it will be in chapters 3, 4 and 5). On the other hand, it is not always possible to isolate an action from the physical possibility relative to the action (e.g., if one wants to pick up a kettle, one needs to have a hand free, but one may be able to hold both a cup and a cube of sugar in the same hand). Before studying action at a level of abstraction that ignores objects' physical properties (in chapters 3, 4 and 5), this chapter looks into what extent the physical, perceptual and semantic properties of a target and its surroundings affect action selection processes. As mentioned in the Chapter 1, subtle kinematic features between the start and end of a movement can reveal precious information about the processes underlying action control as action unfolds in real-time, and this study takes advantage of it by measuring the kinematics of reaching and grasping.

The experimental data in this chapter was collected by Livia Freier who designed the study, presented in Freier (2016). In the present chapter, we re-used the data from Freier (2016), on which we performed new analyses from scratch (only re-using the time stamps previously determined manually, as described in the Section 2.2). In addition to the analyses, the current literature review, methods and discussion in this chapter are all novel outputs of this thesis.

We first review the theoretical background of selection-for-action. We then expand on experimental studies revealing what features of reach-to-grasp can be subject to interference, and review what object properties may lead to such interference on reaching-to-grasp a target among distractors. We also specify the developmental findings related to these factors.

2.1.2 Theories of selection-for-action

Tipper et al. (1992) observed that reaching properties differed as a function of the surroundings of a target, in particular distractor objects. A number of theories have proposed explanations of the causes of such interferences on reaching to a target. The *premotor theory* (Rizzolatti et al., 1987) postulates that when attending to a location, goal-directed spatially coded movements are prepared through spatial selective attention processes. Relatedly, the *visual attention model* (W. X. Schneider, 1995) advances that target selection is performed by common mechanisms for *selection-for-action* (located in the dorsal pathway) and *selection-for-perception* (ventral pathway). Accordingly, motor programs are prepared towards attended objects, whether they are the intended target or not.

In both the premotor theory (Rizzolatti et al., 1987) and the visual attention model (W. X. Schneider, 1995), the processing of objects and preparation of actions happen together and motor activations occur for action regardless of whether the action will be executed. An alternative hypothesis held in the late 1990s, the *visuospatial hypothesis* claimed that interference was solely dependent on the spatial location of distractors. Meegan and Tipper (1999a) refuted this. Instead, they found that interference depended on the nature of actions afforded by distractors. Supported by the identification of vision-for-action neurons (e.g., Ottes et al., 1987), they proposed the *visuomotor processing hypothesis* which states that distractor objects elicit action representations even in absence of intention to act upon the objects.

Multiple evoked action activations

In all three theories there is the common view that when acting on a target surrounded by other objects, the visual inputs from the target and non-target objects evoke motor activations or action representations. Multiple possible motor action representations can be elicited in parallel, and such activations exist even for objects that are task-irrelevant, independently of the participant's awareness or intention to act (Castiello, 1996).

Before performing an action, a process of selection must therefore occur. Thus, prior to and during the executing of an action, interference may be measured in neural and behavioural data which reflects the competition of multiple activations competing for action.

Inhibition of competing actions

Following the encoding of multiple actions, Tipper (1985) (see also Howard and Tipper, 1997) proposed that action-based inhibitory selection mechanisms are required to facilitate selecting a target. Such inhibitory processes, which suppress the competing activations of (some) objects in the visual field, are reflected in the kinematics of action initiation or execution. For example, Tipper et al. (1998) further proposed that the amount of inhibition determines if a trajectory will deviate away from or towards a distractor.

Interference in response time and kinematics

Before focusing on the parameters of distractors that interfere with reaching actions, we specify how such interference can manifest itself and be detected in behavioural measures. Reaction time studies have found the presence of distractors (even task-irrelevant) to slow down responses (e.g., Bradley, 1969, Pratt and Abrams, 1994, Jackson et al., 1995, Simone and Baylis, 1997, Gnanaseelan et al., 2014). Other studies have found distractors to affect trajectory paths when reaching for a target: Welsh et al. (1999) or Welsh and Elliott (2004) found that the reaching path deviated away from distractors. On the contrary, Chang and Abrams (2004) found distractors to attract the path towards them.

According to Tipper et al. (1998), deviations towards or away from distractors depend on the amount of inhibition required. Deviations towards a distractor are caused by residual activations that have not been sufficiently inhibited. However, if the distractor requires greater inhibition, the path may deviate away from the inhibited distractor. A last possibility is that the level of inhibition is intermediate so that no deviation occurs.

2.1.3 Distractor properties affecting reach-to-grasp actions

We saw that the competition between target and distractor representations can generate interference in reaching actions. We now turn to the factors (properties of the stimulus, appearance, relevance, and their relative location to the target) that influence the extent of interference.

Spatial parameters

A multitude of factors have been found to modulate the presence and the extent of distractor interference on reaching: the object's (target or distractor) size (Jervis et al., 1999, Castiello, 1996), the object's shape (Castiello, 1996), the distance between target and distractors (Keulen et al., 2004), the hand starting position (Tipper et al., 1992), the object's being in reachable space or not (Gallivan et al., 2009). The distances, hand start position and reachability are held constant in the study reported here in order to concentrate on the affordance, semantic and set size factors, whose roles are still unclear and can be expected to have different influences on reach across development.

Perceptual

According to the *affordance hypothesis* or *perceptual hypothesis*, interference occurs due to conflicting activations by incompatible affordance representations between distractors and the target (e.g., Gazzaniga et al., 1962). Consistent with the affordance theory of Gibson (1979), the affordance or perceptual hypothesis predicts that a distractor will compete with, and thus affect the kinematics of reaching to a target, only if the target and distractor afford different actions.

Semantic

Jervis et al. (1999) investigated the role of semantic relatedness between a target and a distractor in reaching interference. Kinematics were measured in adults reaching for an apple placed alone (control condition), placed with a compatible distractor object (other apple; semantically compatible condition) or placed with an incompatible distractor object (box the same size as the apple; semantically incompatible condition). The semantic compatibility modulated the grasp but not the reach kinematics. The reaching duration did not vary with compatibility but the peak acceleration of finger opening in grasping was found earlier with a semantically incompatible distractor than with a semantically compatible distractor (and earlier with any distractor than when no distractor was presented). However, the apple and box distractors differed not only in their semantic compatibility but also in their grip *affordance (perceptual compatibility)*. In a complementary study, Jervis et al. (1999) controlled for the role of perceptual compatibility on interference and found that only semantic compatibility affected the interference.

Set size

A key contrast between the studies of Jervis et al. (1999) and those of, for example, Tipper et al. (1992) is that in the former the target object was always located at the same position, while in the latter there was an element of visual search. To our knowledge, no study has explored whether the number of distractor objects affects reach-to-grasp actions when the target object is in a fixed location. In visual search studies (e.g., Beck et al., 2010, and Treisman and Gelade, 1980), and studies of multiple object tracking (e.g., Bettencourt and Somers, 2009), increasing the number of distractors (ranging from 1 to 16) typically increases response times with each added distractor item. This increase is not (solely) due to serial processing of distractors. For example, Baldauf et al. (2006) provided behavioural evidence that more than two reaching actions can be prepared in parallel rather than serially, and neurophysiological data supports the simultaneous representation of target and multiple distractors' representations (e.g., Basso and Wurtz, 1997, Cisek and Kalaska, 2005, Cisek, 2006). Given these considerations, it is unclear whether one should expect the presence of multiple distractors to lead to more interference than one distractor in the absence of visual search.

2.1.4 Development of action-based selection mechanisms

If the distractors elicit activations that must be suppressed via inhibitory mechanisms, then different age groups are expected to face different degrees of interference. Indeed, inhibitory control is known to improve throughout childhood until late adolescence (e.g., Luna, 2009, Davidson et al., 2006, Blaye and Chevalier, 2011). Along with other executive functions, changes in inhibitory control have been linked to the protracted maturation of the prefrontal cortex (e.g., Lenroot and Giedd, 2006, Diamond, 2002). More precisely, inhibitory control appears to change very intensely around 4-6 years old, relatively less until 7 years old (Mehnert, Akhrif, Telkemeyer, Rossi, Schmitz, et al., 2013) and, albeit possibly at a slower rate, throughout adolescence (Casey et al., 2005a, Klenberg et al., 2010). At the earliest stage of development such low inhibitory control results in the incapacity to resist the most activated response at a given time (Diamond, 1990a). Throughout development, gains in inhibitory control may lead to quicker conflict resolution.

A variety of tasks are used to measure the inhibition of prepotent responses. These tasks commonly face the task impurity issue (Burgess, 1997, Phillips, 1997): they often measure a variety of processes rather than a pure inhibition component. We may distinguish two broad types of tasks held to measure response inhibition. A first type includes tasks like the Stop signal (Logan, 1994) or the Go-NoGo (Shapiro et al., 2013), where, depending on the stimuli, a response must be given or withheld. Importantly, there is only one possibility of response (e.g., the task requires to press a specific key, or to withhold responding). The second type of inhibition task involves more than one possible responses (e.g., both left and right keys), where one response may be prepotent but need to be inhibited to give the other, appropriate response. This is the case in the Flanker task and Stroop tasks (Eriksen and Eriksen, 1974, Stroop, 1935). We will focus here on the second type of task, which is relevant to the current chapter's. Indeed, the reaching paradigm used in this chapter task investigates the development of inhibitory control applied to conflict resolution among more than one possible responses.

Erb et al. (2018) used a variation of the Eriksen flanker task (Eriksen and Eriksen, 1974). This classical version of the Flanker task requires participants to respond to a central stimulus 'flanked' by distractors stimuli on the sides. On congruent trials, the target and distractor stimuli cue the same response (e.g., all arrows pointing leftwards, and the leftward key must be pressed) while on incongruent trials they cue competing responses (e.g., a central arrow pointing rightwards and the others leftwards, and the rightward key must be pressed). Inhibitory control is held to be recruited to override the response evoked by the distractor cues. Thanks to a reach-to-touch adaptation of the Flanker task, Erb et al. (2018) were able to measure within-trial changes in reach curvature hypothesised to index the co-activation between response alternatives during the movement, whereby greater curvature indicates a less efficient conflict resolution process (Erb et al., 2016). The results of Erb et al. (2018) suggest that there might be improvements in the conflict resolution process from 8 to 10 years of age and until adulthood.

The Stroop task (Stroop, 1935) has been proposed by some as an indicator of inhibitory control, either because it indexes the ability to inhibit prepotent representations (Cepeda et al., 2001, Diamond et al., 2005) or the ability to inhibit task sets (Mayr and Keele, 2000). A child-friendly version of the Stroop task, the Day/Night task (Gerstadt et al., 1994) has been used widely. In the Day/Night task, participants must say 'day' when presented with a black card with a moon and stars, and say 'night' when presented with a brightly-coloured card with a sun. Performance on the Day/Night task has been found to improve across preschool and school years (Gerstadt et al., 1994, Hughes et al., 2009, Willoughby et al., 2012).

Thus, diverse tasks that presumably recruit inhibitory control (among other processes) have found that improvements in inhibitory control with age led to more accurate or faster selection among competing options. Drawing on these findings, we predict that such developmental changes in inhibitory control will result in reduced effects of distractors on reaching with development.

2.1.5 The current study

The present study was designed to better understand the roles of distractor properties (number, semantic and perceptual compatibility) and the maturation of inhibitory control on reach-to-grasp interference. It was built on the reaching task used by Jervis et al. (1999) and presented above. Here, children at 6, 9 and 12 years of age, as well as adults reached for a target object placed in the centre of the table and flanked either by one or two distractors that had various perceptual and semantic relationships to the target object, or no distractor. This constituted a total of seven conditions described below and illustrated in Figure 2.1:

- (A) no distractor object condition,
- (B) perceptually incompatible / semantically incompatible condition,
- (C) perceptually incompatible / semantically compatible condition,
- (D) perceptually compatible / semantically incompatible condition,
- (E) perceptually compatible / semantically compatible condition,
- (F) perceptually compatible and semantically incompatible condition,
- (G) perceptually incompatible and semantically compatible condition.

Thus, conditions differed in the combination of the semantic and perceptual compatibility between the target and distractor and in the number of distractors presented simultaneously (set size). For example, in condition *(D) perceptually compatible / semantically incompatible condition*, the distractor was a tennis ball which afforded the same grip as the target apple but belonged to a different semantic category; and in condition (E) the distractor was an orange which offered the same grip and belonged to the same semantic category (all objects are shown in Figure 2.1).

As interference in reaching actions can be revealed either with temporal or spatial parameters, we measured the reach and grasp movement durations as well as the spatial kinematic features of the reach movement.

2.2 Methods

2.2.1 Participants

Sixty-five participants (9 left handed / 56 right handed) were included in the final sample. Participants constituted four age groups as follows: seventeen 6-year-olds (M = 6.4 year, SD = 0.3 years; 7 females), twenty-one 9-year-olds (M = 9.3 years, SD = 0.3; 10 females), eighteen 12-year-olds (M = 12.7 years, SD = 0.4 years; 6 females) and nine adults (M = 29.2 years, SD = 8.3 years; 7 females). Only participants for whom there was sufficient data was included in the final sample, as described hereafter. Six 6-year-olds, four 9-year-olds and two adults who were tested but had excessive marker occlusion were excluded from the final sample (not counted in the above). Three 6-year-olds and one 12-year-old were excluded as well because they failed to consistently bring, as instructed, their reaching hand back to the start location at the onset of trials. Participants were tested individually. They were naïve to the purpose of the study and had normal-to-corrected vision. Written consent was obtained from the parents or caregiver, and verbal consent to participate was obtained from the children. The study had ethical approval from Birkbeck's institutional Ethics Committee and the study sessions were conducted in accordance with the declaration of Helsinki. All the participants were offered travel reimbursement and children were given small thank-you gifts.

2.2.2 Materials

The stimuli were a combination of one to six different objects: three round objects of equal diameters (6 cm), each part of a different semantic category (an apple, a tennis ball and a tea-ball infuser); and three common objects that varied in shape, size as well as semantic category (a banana, grape and scissors). Before the testing session children and adults were able to identify all objects and were given an opportunity to touch them.

2.2.3 Procedure

The experimental sessions took place indoors under normal lighting conditions. Participants were sat with their dominant hand in front of them, place on a 50cm x 50cm table. The start location was marked by a line on the table. Participants were sat with the midline of their body aligned to the start location. They were instructed to place their dominant (reaching) hand at this start location with their middle digits touching the line without crossing it. This hand placement was practised before recording the trials. To control for the visual processing time prior to reach onset, a screen was placed on the table such that it was hiding the placement of stimuli at the start of each trial.

Behind the screen, the target apple was located in the centre of the table, 30cm away from and aligned with the hand start location. One or two distractor objects (as indicated in Figure 2.1) were located also at 30 cm and at a 20°angle from the line between hand and apple. Consequently, distractors appeared closer to the start location relative to the target, and were well within reach of participants, without being obstacles to reaching the target. The location of distractors was counterbalanced to the left and right hand side of the target.

Participants were instructed to reach for the central apple in a way that felt natural, to then pick up the apple and finally place it back on the table. They were also instructed to ignore all other objects. Each participant carried out three practice trials before the test trials.

Each condition was presented 10 times, except the No distractor condition which was presented 30 times. The resulting 90 trials were randomly ordered. At the start of a trial, participants were reminded to place their reaching hand at the start location, and a new trial began only when this was confirmed. The onset of each trial was the moment when the screen was lifted. The reach-to-grasp actions were recorded throughout with an optical motion-capture camera system as described below.

2.2.4 Motion-capture setup

Spherical reflective markers (diameter: 4 mm) were attached to the participants' reaching hand with skin-safe double-sided adhesive tape, at four points. The first three points followed the placements used in Jervis et al. (1999), that is: (1) on the wrist - radial styloid process, (2) on the index finger dorsal side of the tip of the

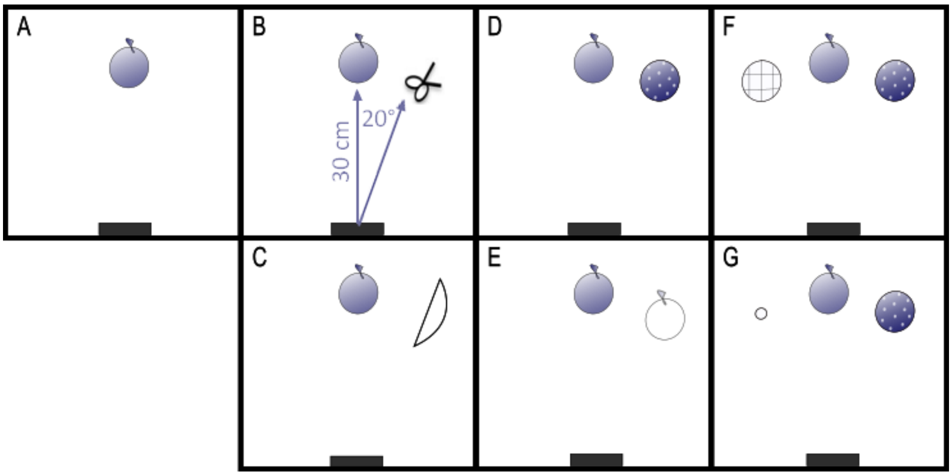


Fig. 2.1: **Experimental setup: stimuli across conditions. The target object (apple) always appeared in the centre, 30cm away from the hand starting position (black rectangle). (A) No distractor condition (alone), and with (B) one perpetually incompatible/semantically incompatible distractor condition (pair of scissors); (C) one perceptually incompatible/semantically compatible distractor (banana); (D) one perceptually compatible/semantically incompatible distractor (tennis ball); (E) one perceptually compatible/semantically compatible distractor (apple); (F) two distractors, one of which is perceptually compatible and the other semantically incompatible (tennis ball and tea infuser); (G) two distractors, one of which is perceptually incompatible and the other semantically compatible (banana and grape).**

second finger and (3) on the thumb - dorsal side of the tip of the index finger. The last (4) marker was attached to the first knuckle of the index finger. The latter was chosen following previous motion-capture studies with children (e.g., Coluccini et al., 2007, Jaspers et al., 2011). A marker was also attached to each object to record their location.

The movements of the markers were tracked by a five-camera optical motion capture system (from Qualisys Inc., Sweden: 4 x *Oqus* 1 cameras to track the markers and 1 x *Oqus* 210c camera to record and synchronize colour video). The marker tracking worked through the emission by each camera of near-infrared light spectrum (800nm wavelength) by strobe-emitting LEDs. Data was sampled at 100 Hz. The marker-tracking cameras were placed on tripods around the table, arranged in a quarter-circular fashion from 1 to 1.5 m to the closest edge of the table, and 0.3 to 1 m apart from one another. Calibration was made so that cameras would record within a volume 1.5 m long x 1 m wide x 1 m high, via a calibration wand with two markers 299.3 mm apart. The coordinate system origin was given by a static, L-shaped frame with four coplanar markers (one at the junction; one on the short side at 200 mm distance from the junction marker; two markers on the long side at 90 mm and 300 mm distances from the junction marker) placed on the table during calibration.

Data was acquired via the Qualisys Track Manager (QTM) software package (version 2.9). Each session was recorded continuously. The three-dimensional marker positions were reconstructed with the QTM software.

2.2.5 Coding

The following coding was performed by the original author of the study (Livia Freier), and the coded data files were re-used in this thesis. Successful trials were assessed visually to identify the frame at which one of the following events occurred: (1) reaching start, (2) grasping start and (3) end of the movement (both reaching and grasping). The reaching start component was coded as the moment at which the hand was lifted off the table, as reflected by an acceleration in the velocity curve. The grasping start component was coded as the point at which the arm had its maximal extension. The end of the movement was coded as the moment when the fingers closed around the object. The interval from reaching start, (1), to the end, (3), formed the reach movement, and the interval from grasping start, (2), to the end (3) formed the grasp movement. Three independent coders carried out the visual coding. They first checked the video for correct placements of the hand at the start of trials. Coders also screened the velocity curves for signal failures, shown by discontinuous signal. All coders were trained similarly for the manual coding procedure, and the final Krippendorff's Alpha coefficient of inter-coder reliability was 0.90 or higher. Trials with signal failure (caused by marker occlusion), as well as trials with wrong hand placement at the start were excluded.

2.2.6 Processing of kinematics data

From this Section and on, the analyses were done (or replicated) anew for this thesis. The kinematics data was smoothed with a second-order Butterworth filter (lowpass, 10 Hz cuttoff frequency). We analysed separately the two-dimensional motion parallel to the table workspace and the movement perpendicular to the table workspace. Trials were filtered as follows. Individual trials were retained if the knuckle marker data was available, or alternatively the wrist marker (analyses on data where both markers were available indicated that wrist and knuckle markers yielded similar path deviation features, with Spearman correlation coefficient of 0.62, $p < 0.001$). Subsequently, a condition was deemed sufficiently represented when more than 4 trials remained (more than 8 for condition (A) *No distractor*). Indeed, except for the *No distractor* condition, at least two conditions contribute to each level of the predictors (semantic, perceptual, set size) which means at least 8 trials per level per participant were available. Thus, the average number of trials per participant included in the final analyses were: 60 trials for the 6-year-olds, 63 for the 9-year-olds, 72 for the 12-year-olds, and 62 for the adults.

In the retained trials, we computed the following features: (a) reach movement duration (from the reach start event to the end of the movement), (b) grasp movement duration (from the time of the maximal extension of the limb to the time that the target is moved from its initial location), (c) the maximal lateral deviation from a straight path in the horizontal plane, referred to as ’maximal x-deviation’, and (d) maximal deviation in the vertical plane (i.e., maximal altitude above the table), referred to as ’maximal z-deviation’. The kinematic parameters were derived from the position data of the knuckle or wrist markers. The temporal parameters were individually computed for reach and grasp components, and the deviation parameters only for the reach component. The maximal x-deviation and maximal z-deviation trajectory features were log-transformed prior to averaging to meet normality assumptions. Preliminary analyses did not reveal any significant effects of gender handedness.

Further kinematics other than response time were not derived for the grasp component because contrary to the knuckle or wrist marker, the index and thumb markers suffered from excessive occlusion thus we did not have enough data for that purpose.

2.2.7 Mixed models and Bayes Factors

To investigate the difference between conditions across age groups, we conducted a series of linear mixed-effects regressions with the *lme* function of the *nlme* package (Pinheiro et al., 2020) for the R software environment (Version 3.6.0; R Core Team, 2019), fitting the models by maximizing the restricted loglikelihood (REML method). We defined a random intercept per participant to account for the repeated measures (Steele, 2014), and defined the following fixed effects: semantic compatibility and perceptual compatibility as categorical predictors (coded as 0 if incompatible, 1 if compatible), set size as a continuous predictor (1 in condition A, 2 in B, C, E and G, 3 in D and F), and age as an ordered categorial predictor. In addition to the main effects, we entered all combinations of 2-ways interactions because this model was more parsimonious than a model which included both all 2-way and all 3-way interactions. For the ordered categorical predictor (age), we tested for linear trends.

Additionally, we computed Bayes Factors (BF) using the R package *BayesFactor* (Morey and Rouder, 2018). For each age group and each predictor P (semantic compatibility, perceptual compatibility and setsize), we generated a BF_{10} comparing the Bayes Factor of model H_1 over the Bayes Factor of model H_0, where H_1 was a mixed model (function *lmBF*) which had all three predictors and a random subject intercept, and H_0 the same without the predictor P. We used the recompute function with 1,000,000 Monte Carlo samples, so that the proportional error on the Bayes Factor all fell under 1%. We use the classification scheme proposed by van Doorn et al. (2020) to interpret the Bayes Factors strength, where a BF_{10} from 1/30 to 1/10 is taken as Strong evidence for H_0, from 1/10 to 1/3 as Moderate evidence for H_0, from 1/3 to 1 as Weak evidence for H_0, from 1 to 3 as Weak evidence for H_1, from 3 to 10 as Moderate evidence for H_1 and from 10 to 30 as Strong evidence for H_1.

2.3 Results

2.3.1 Reach movement duration

Unsurprisingly, response times decreased with age and this held across all conditions (main effect of age, linear trend: β = -0.117, SE = 0.042, t(61) = -2.818, p = 0.007, 95% CI [-0.201, -0.034]). However, young children's movement durations were sensitive to condition, unlike that of the adults.

As is apparent in Figure 2.2 (right panel), across the one-distractor conditions there was no significant effect of either semantic (β = 0.014, SE = 0.008, t(293) = 1.701, p = 0.09, 95% CI [-0.002, 0.03]) or perceptual compatibility (β = 0.008, SE = 0.008, t(293) = 1.003, p = 0.317, 95% CI [-0.008, 0.024]) on reach movement duration. Moreover, the Bayes Factors provide moderate to weak evidence for the absence of an effect of semantic (BF_{10} = 0.27, 0.25, 0.46, and 0.31 for the 6-, 9-, 12-years old and adult respectively) or perceptual compatibility (0.23, 0.23, 0.35, 0.52 for the 6-, 9-, 12-years old and adult respectively) on reach movement duration.

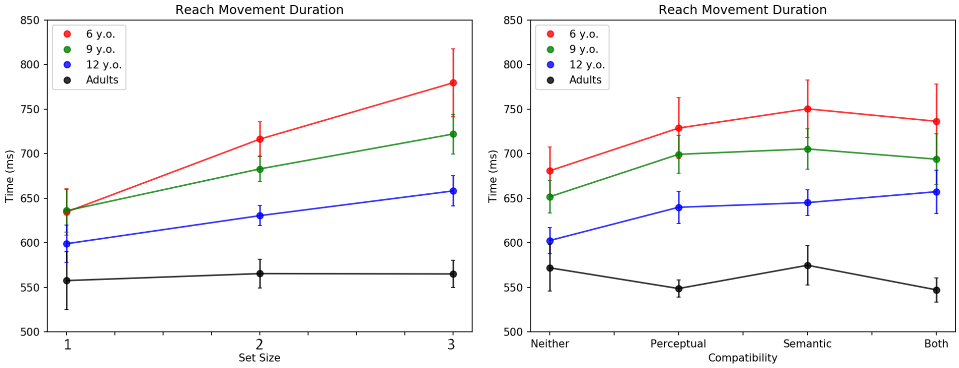


***Fig. 2.2:* Reach movement duration as a function of age and set size (left panel), or age and compatibility (right panel). Setsize 1 corresponds to the No distractor condition, Setsize 2 to conditions with one distractor and Setsize 3 to conditions with 2 distractors). Error bars represent one standard error from the mean.**

In contrast, the effect of set size was significant (β = 0.039, SE = 0.006, t(293) = 6.187, p = <0.001, 95% CI [0.027, 0.052]), as was the interaction between age and set size (linear trend, β = -0.048, SE = 0.014, t(293) = -3.462, p = 0.001, 95% CI [-0.076, -0.021]). As shown in Figure 2.2 (left panel), the younger the children, the higher the effect of set size, with adults not showing this effect. The entire regression results are also presented in table 2.1.

The Bayes Factors provide further insight into the developmental changes in the set size effects: there is moderate evidence for a set size effect in the 6-year-olds (BF_{10} = 3.62), weak evidence for this effect in the 9-year-olds (BF_{10} =1.50), and (in contrast) weak evidence for the absence of such an effect in 12-year-olds (BF_{10} = 0.60), and adults (BF_{10} = 0.34).

Table 2.1: **Results of the Linear Mixed Model on Reach movement duration with the predictors *Age*, *CatSem* (Semantic compatibility, baseline: Incompatible), *CatPer* (Perceptual compatibility, baseline: Incompatible), *Ndistractors* (number of distractors) and their interactions. β': Estimate (regression coefficient), 'CI (l)': Confidence Interval (lower bound), 'CI (u)': Confidence Interval (upper bound), 'SE': standard error, 'DF': degrees of freedom, 't': t-value, 'p': p-value, 'sig.': significance levels: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.)**

Predictor	β	CI (lower)	CI (upper)	SE	DF	t	p-value	sig.
(Intercept)	0.750	0.712	0.787	0.019	293.000	39.323	<0.001	***
Age	-0.117	-0.201	-0.034	0.042	61.000	-2.818	0.007	**
CatSem	0.014	-0.002	0.030	0.008	293.000	1.701	0.09	
CatPer	0.008	-0.008	0.024	0.008	293.000	1.003	0.317	
Ndistractors	0.039	0.027	0.052	0.006	293.000	6.187	<0.001	***
Age:Ndistractors	-0.048	-0.076	-0.021	0.014	293.000	-3.462	0.001	***
Age:CatPer	0.010	-0.026	0.046	0.018	293.000	0.539	0.591	
Age:CatSem	-0.018	-0.053	0.017	0.018	293.000	-1.021	0.308	

2.3.2 Grasp movement duration

For grasp movement duration (Figure 2.3 and table 2.2), the effects were different than for reach movement duration. The effect of age was significant (linear trend, β = -0.037, SE = 0.012, t(61) = -3.094, p = 0.003, 95% CI [-0.061, -0.013]); however, the other effects and interactions were not significant (Semantic compatibility: β = -0.009, SE = 0.008, t(290) = -1.198, p = 0.232, 95% CI [-0.024, 0.006], Perceptual compatibility: β = -0.006, SE = 0.008, t(290) = -0.762, p = 0.447, 95% CI [-0.022, 0.01], set size: β = 0.004, SE = 0.004, t(290) = 0.869, p = 0.385, 95% CI [-0.005, 0.012]), interaction between age and set size: linear trend, β = -0.007, SE = 0.005, t(290) = -1.44, p = 0.151, 95% CI [-0.016, 0.002]).

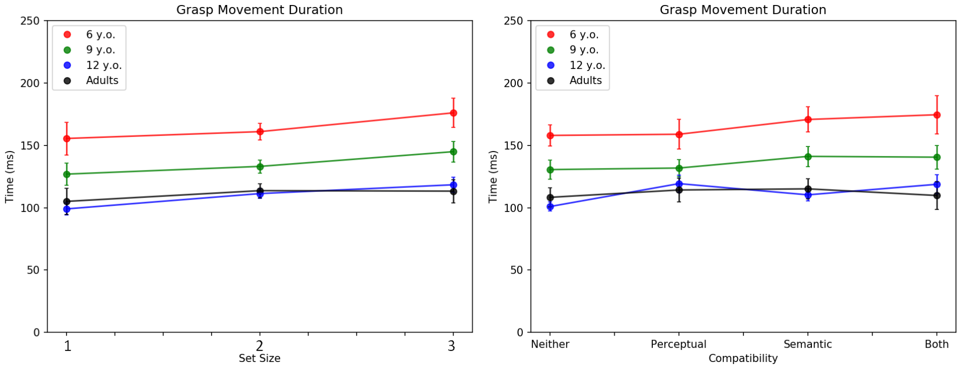


Fig. 2.3: **Grasp movement duration as a function of age and set size (left panel), or age and compatibility (right panel). Setsize 1 corresponds to the No distractor condition, Setsize 2 to conditions with one distractor and Setsize 3 to conditions with 2 distractors). Error bars represent one standard error from the mean.**

Table 2.2: **Results of the Linear Mixed Model on Grasp movement duration with the predictors *Age*, *CatSem* (Semantic compatibility, baseline: Incompatible), *CatPer* (Perceptual compatibility, baseline: Incompatible), *Ndistractors* (number of distractors) and their interactions. β’: Estimate (regression coefficient), ’CI (l)’: Confidence Interval (lower bound), ’CI (u)’: Confidence Interval (upper bound), ’SE’: standard error, ’DF’: degrees of freedom, ’t’: t-value, ’p’: p-value, ’sig.’: significance levels: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.)**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	sig.
(Intercept)	0.124	0.112	0.135	0.006	290.000	21.737	<0.001	***
Age	-0.037	-0.061	-0.013	0.012	61.000	-3.094	0.003	**
CatSem	-0.009	-0.024	0.006	0.008	290.000	-1.198	0.232	
CatPer	-0.006	-0.022	0.010	0.008	290.000	-0.762	0.447	
Ndistractors	0.004	-0.005	0.012	0.004	290.000	0.869	0.385	
Age:Ndistractors	-0.007	-0.016	0.002	0.005	290.000	-1.440	0.151	
Age:CatPer	0.007	-0.005	0.019	0.006	290.000	1.145	0.253	
Age:CatSem	-0.005	-0.017	0.007	0.006	290.000	-0.839	0.402	
CatPer:Ndistractors	0.007	-0.005	0.019	0.006	290.000	1.110	0.268	
CatSem:Ndistractors	0.009	-0.003	0.021	0.006	290.000	1.436	0.152	
CatSem:CatPer	0.007	-0.006	0.019	0.006	290.000	1.080	0.281	

2.3.3 Summary of movement duration results

In all conditions (i.e., regardless of distractor presence or type), the reaching component of younger children was slower than older children, who were in turn slower than adults. These main effects of age are not surprising. The more interesting effects relate to compatibility and set size. The compatibility of distractors (semantic or perceptual) did not affect the movement time of the reach and grasp components. That is, the reach and grasp time were similarly unaffected by the perceptual or semantic features of distractors. This null effect is supported by the Bayes Factors cited above. When analysed in terms of the number (rather than type) of distractors, however, a different pattern was apparent. Reach time (but not grasp time) was significantly affected by the number of distractors in the younger groups, while older participants were less affected by set size.

2.3.4 Kinematics variables

Analyses of the movement trajectories (i.e., the maximal deviation in the horizontal and vertical planes) revealed that the maximal lateral deviation (in the horizontal plane) was only affected by age. However, the maximal deviation in the vertical plan mirrored the duration results, with a larger deviation as a function of set size for children but not for adults.

Linear mixed-effects regressions on the maximal lateral deviation (Figure 2.5) showed a main effect of age (linear trend, β = -0.521, SE = 0.126, t(61) = -4.153, p = <0.001, 95% CI [-0.772, -0.27]), but no other effect nor interaction (see table 2.3). More interestingly, the maximal vertical deviation (Figure 2.4 and table 2.4) showed a significant effect of set size (β = 0.051, SE = 0.022, t(307) = 2.34, p = 0.02, 95% CI [0.008, 0.093]) and a significant interaction between age and set size (linear trend, β

= -0.039, SE = 0.019, t(307) = -2.096, p = 0.037, 95% CI [-0.076, -0.002]). This mirrors the pattern of set size effects and interactions found for reach movement duration.

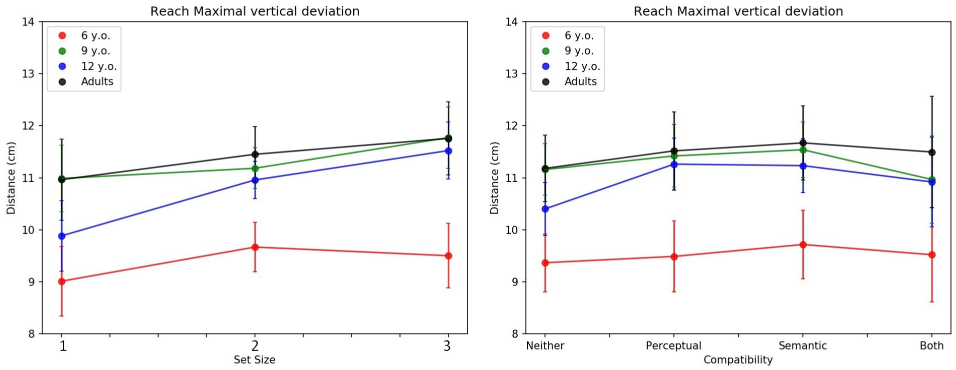


Fig. 2.4: **Maximal vertical deviation: maximal deviation in the vertical plane (i.e. altitude) as a function of age and set size (left panel), or for conditions with exactly 1 distractor as a function of age and compatibility (right panel). Setsize 1 corresponds to the No distractor condition, Setsize 2 to conditions with one distractor and Setsize 3 to conditions with 2 distractors). Error bars represent one standard error from the mean.**

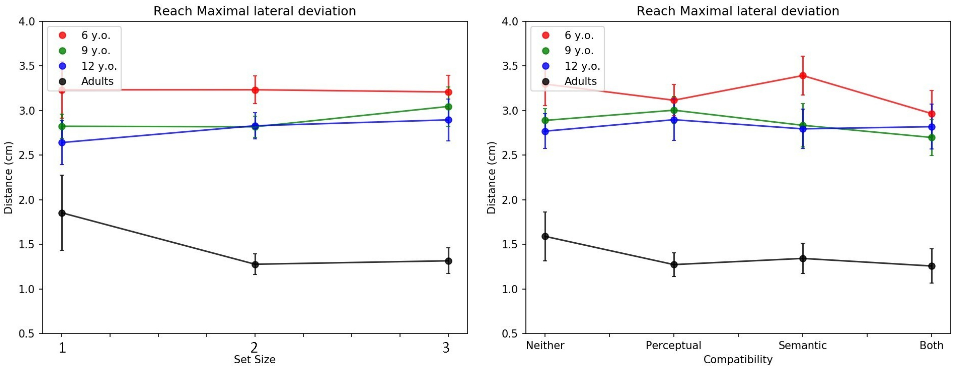


Fig. 2.5: **Maximal lateral deviation: maximal distance in the horizontal plane from a straight path (that is, direct path from start position to target location) as a function of age and set size (left panel), or for conditions with exactly 1 distractor as a function of age and compatibility (right panel). Setsize 1 corresponds to the No distractor condition, Setsize 2 to conditions with one distractor and Setsize 3 to conditions with 2 distractors). Error bars represent one standard error from the mean.**

To test whether reaching time might derive directly from longer trajectories (in the horizontal or vertical plane), we performed correlational analyses across subjects. The maximal x-deviation had a small but significant correlation with reach movement duration (Spearman correlation coefficient: r = 0.308, p = 0.012), but the maximal z-deviation was not correlated with reach movement duration (r = 0.187, p = 0.135). This suggests that the effects revealed in reach duration and the vertical deviation measures are capturing partly independent processes.

Table 2.3: **Results of the Linear Mixed Model on Maximal horizontal deviation with the predictors *Age*, *CatSem* (Semantic compatibility, baseline: Incompatible), *CatPer* (Perceptual compatibility, baseline: Incompatible), *Ndistractors* (number of distractors) and their interactions. β': Estimate (regression coefficient), 'CI (l)': Confidence Interval (lower bound), 'CI (u)': Confidence Interval (upper bound), 'SE': standard error, 'DF': degrees of freedom, 't': t-value, 'p': p-value, 'sig.': significance levels: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.)**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	
(Intercept)	3.027	2.910	3.143	0.059	307.000	51.204	<0.001	***
Age	-0.521	-0.772	-0.270	0.126	61.000	-4.153	<0.001	***
CatSem	-0.111	-0.251	0.029	0.071	307.000	-1.562	0.119	
CatPer	-0.002	-0.144	0.139	0.072	307.000	-0.031	0.975	
Ndistractors	0.018	-0.063	0.098	0.041	307.000	0.434	0.664	
Age:Ndistractors	-0.036	-0.122	0.049	0.044	307.000	-0.835	0.404	
Age:CatPer	0.037	-0.073	0.146	0.056	307.000	0.653	0.514	
Age:CatSem	0.030	-0.079	0.138	0.055	307.000	0.535	0.593	
CatPer:Ndistractors	0.001	-0.113	0.115	0.058	307.000	0.017	0.986	
CatSem:Ndistractors	0.069	-0.044	0.182	0.058	307.000	1.200	0.231	
CatSem:CatPer	0.012	-0.103	0.126	0.058	307.000	0.199	0.842	

2.4 Discussion

2.4.1 Distractor interference effects

We found that the presence of task-irrelevant distractors changes the kinematics of reach-to-grasp action to a target. This is consistent with a large body of evidence (e.g., Tipper et al., 1992, Castiello, 1996, Baldauf and Deubel, 2009) that underlies theorising that processing of a scene with multiple objects generates action representations in parallel, prior to a selection process. The selection process then consists in suppressing unwanted responses by inhibiting their representation and thus select the action to the target object.

Specifically, we found that reaching duration increased with the number of distractors, for the children's group. There were age differences in the extent of that duration interference within the children group. This was the case specifically for the reach component duration but not the grasp component. The Bayesian analyses provide evidence for a slowing in the younger age group reaching when more distractors are present. Additionally, they suggest an absence of an effect of distractors on reaching duration in the older age groups, although the low Bayes Factors call for further studies to test this absence of effect more fully. The duration interference was corroborated by interference on the vertical path.

Another important finding was that the distractor compatibility (semantic, perceptual, incompatible) had no effect on reaching and grasping durations, nor on the vertical and horizontal deviations measures. The Bayes factors provide moderate to weak evidence against the effects, pointing towards the complete absence of an effect of distractor compatibility.

Table 2.4: **Results of the Linear Mixed Model on Maximal vertical deviation with the predictors *Age*, *CatSem* (Semantic compatibility, baseline: Incompatible), *CatPer* (Perceptual compatibility, baseline: Incompatible), *Ndistractors* (number of distractors) and their interactions. β': Estimate (regression coefficient), 'CI (l)': Confidence Interval (lower bound), 'CI (u)': Confidence Interval (upper bound), 'SE': standard error, 'DF': degrees of freedom, 't': t-value, 'p': p-value, 'sig.': significance levels: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.)**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	
(Intercept)	4.530	4.439	4.620	0.046	307.000	98.962	<0.001	***
Age	0.162	-0.035	0.360	0.099	61.000	1.642	0.106	
CatSem	0.067	-0.007	0.141	0.038	307.000	1.772	0.077	
CatPer	-0.022	-0.097	0.053	0.038	307.000	-0.582	0.561	
Ndistractors	0.051	0.008	0.093	0.022	307.000	2.340	0.02	*
Age:Ndistractors	-0.039	-0.076	-0.002	0.019	307.000	-2.096	0.037	*
Age:CatPer	0.005	-0.053	0.064	0.030	307.000	0.186	0.853	
Age:CatSem	-0.006	-0.063	0.052	0.029	307.000	-0.201	0.84	
CatPer:Ndistractors	0.003	-0.058	0.063	0.031	307.000	0.087	0.931	
CatSem:Ndistractors	-0.051	-0.111	0.009	0.030	307.000	-1.686	0.093	
CatSem:CatPer	-0.017	-0.078	0.043	0.031	307.000	-0.567	0.571	

When looking at interference in terms of the number of distractors, the condition with two distractors was also characterised by having opposite perceptual and semantic compatibly. From the results on the one-distractor conditions, it is unlikely that compatibility was more relevant when having two distractors. Consequently, interference effects plausibly arise from the number of distractors in itself.

While it has been found by many that objects present around a target generate automatic processing even of task-irrelevant objects and task-irrelevant properties of an object (Diamond, 1990b, Castiello, 1996, etc.), it remains unclear which specific properties of distractor objects (affordance, semantic relatedness to target) are responsible for interference in reaching actions. By manipulating the perceptual and semantic compatibility of distractors when children and adults reached for a target object, we observed that the interference was solely dependent on the presence and number of distractors, independently of their perceptual or semantic properties.

The response time interference was present in the reaching but not in the grasping component. In previous studies, distractors also had a different effect on reach compared to grasp. For example, Jervis et al. (1999) measured changes in the grasp kinematics as a function of distractor semantic compatibility, while reaching measures remained unaffected. However, contrary to Jervis et al. (1999) who had kinematics measures such as the opening rate of the fingers or the time of maximum grip aperture, our analyses were limited to grasp duration due to excessive occlusion of the thumb and finger markers. It is possible that we would have measured changes in grasp kinematics as a function of setsize or perhaps compatibility if we had accessed to finer-grained measures of the grasp movement. As a consequence of the occluded markers, unlike Jervis et al. (1999) our grasp component was not defined based on the movement of the thumb or of the index finger, but was based on the extension of the arm. This further limits the direct comparison between this study and the one of Jervis et al. (1999). We also note that the visual control of grip aperture evolves between early childhood and 12 years of age (Kuhtz Buschbeck et al., 1998), a change relevant to consider for future studies not suffering from marker occlusion problems.

The distractors affected not only response times but also measures of the unfolding kinematics (maximal deviation). This is consistent with prior findings where interference from distractors appeared in response time (e.g., Pratt and Abrams, 1994) and/or trajectory features (e.g., Chang and Abrams, 2004). In terms of trajectory, studies have observed deviation of the reaching path in the plane formed by the hand start position, the target and distractors locations (e.g., Tipper et al., 1997, Welsh et al., 1999, Chang and Abrams, 2004). A novelty of our analyses is that we found traces of interference in the vertical plane of reaching, even if the distractors were placed in a way that did not physically require any change of trajectory from the condition without distractors (as would have been the case if obstacles had been placed on the no-distractor trajectory). To our knowledge, this vertical dimension of trajectory has never been investigated in studies on interference to reaching and is an interesting phenomenon to further explore to understand more precisely its cause. Within the context of action-based inhibitory mechanisms, we suggest that the vertical deviation may result from residual inhibition that creates avoidance, that is, keeping an ’excessively cautious’ distance away from any object until reaching the target.

As argued by others (e.g., Tipper, 1985, Howard and Tipper, 1997), in the face of multiple objects, the selection of a target requires action-based inhibitory mechanisms to suppress non-target representations. Although inhibition will typically result in slowing down or deviating away from a distractor (e.g., Tipper et al., 1997), very strong inhibition, such as required by salient distractors, may lead to reduced interference compared to less salient distractors (Mohler et al., 2015). This counterintuitive finding is not unlike the fact that enhanced conflict on a task may result in the recruitment of higher levels of cognitive control, resulting in better performance (Botvinick et al., 2001). From this perspective, reaching speed and path straightness may positively benefit from highly conflicting visual input surrounding a target. However, our findings do not support this view: a distractor both perceptually and semantically incompatible did not produce more interference than a distractor both perceptually and semantically compatible. Thus, the movement duration and path deviation variables do not suggest a processing advantage for fully incompatible distractors compared to more compatible conditions. At the same time, the efficient suppression of salient distractors may become relevant only when actions precede other forthcoming actions, as suggested by Mohler et al. (2015). In other words, it is possible that the evidence or interference in single reach actions would not fully apply to the action part of a bigger sequence, and thus should be studied in the context of action sequences.

2.4.2 Methodological differences to previous studies

A key difference with many previous studies is that the present task did not require target identification per se. The target location was known in advance and constant throughout the experiment, and the distractors were never potential targets. Given the low difficulty level of the task, it is possible that participants' encoding of the task requirements was loosely defined by its action and location ('reach for the central object'), ignoring the target identity (e.g., the semantic features of the apple) or the specific affordance (e.g., reach and hold with a large grip). Indeed, the grasping action may initially be encoded at a gross-grained resolution in which the affordances across objects would be similar. In other words, at each level of set size, all types of distractors evoked reaching plans that differed only by their location, thus requiring the same degree of inhibition regardless of their relationship (semantic or perceptual) to the target apple. This is consistent with Tipper et al. (2000) who advance that target and distractors may evoke reaching actions to a location without other properties being represented. Other studies with a single, fixed-location target and in which distractors were not potential targets (e.g., Tipper et al., 2000), also found that the distractor interfered with reaching as revealed by path deviations. This brings additional support for the view that even non-target objects may evoke action representations.

We should note that across studies distractor interference manifests itself in different ways: deviation away from distractors as in Chapman and Goodale (2008), deviation towards distractors as in Chang and Abrams (2004), or prolonged reaching time, as in our study. These differences have been attributed to differences in designs, that is the distance between target and distractors or their relative position, and consequently whether the distractors were possibly encoded as obstacles (as in Tipper et al., 1997). Crucially, all types of interference have been taken to reflect the potentiation of action-related representations.

The novel finding, that interference was a linear function of set size (comparing not only 0 distractor to 1 distractor, but also 2 distractors to 1), is consistent with neurophysiological findings. Empirical data indicate there is some overlap in the neural substrates for reaching and saccadic control (e.g., Meegan and Tipper, 1999b). In the latter domain, Basso and Wurtz (1997) found that the activity of neurons responsible for saccadic responses is a function of the number of visual targets. P. Smith and Ratcliff (2004) review evidence indicating that the neural firing rate of sensory cells encoding stimulus and target representations build up more slowly for more difficult decisions, resulting in increased response time. Hence, it is possible that each added distractor adds neural activity that needs to be suppressed for selection of the target response, resulting in longer RT with each additional distractor. An alternative but related possibility is that the greater spatial angle occupied by the extra distractors causes the observed interference effects (Bock and Eversheim, 2000;

Ellis et al., 2007). According to this view, the increased spatial angle with increased set size would be the driver of the interference observed in younger children. Either way, improvements in inhibitory control and selective attention processes with age remain plausible explanations of the observed developmental pattern of set size interference.

2.4.3 Response inhibition and the reduction of distractor interference

Inhibiting the distractors appear key in performing the reaching action without interference, and the youngest children (6 years old) struggled most to do so. This suggests a central role of the development of inhibitory control on efficiently carrying out low-level actions such as reaching for a target object.

Inhibitory-based mechanisms have been identified in a variety of paradigms where the worst performance of the younger age group appears at least partially due to low inhibition. For example, a reduction in interference in a Day/Night variation of the Stroop task has been observed with increasing age during preschool and school years (e.g., Gerstadt et al., 1994). Furthermore, the reach-to-touch task of Erb et al. (2018), which unlike the present chapter's task used visual cues and not physical objects, found gains in the conflict resolution process between 8 and 10 years of age.

It can be argued that in this reaching task, as well as in the Day/Night Stroop task and reach-to-touch Flanker task, school-aged children's changing performance with age is mainly linked to changes in inhibitory control because the demands of such tasks on other processes (e.g., rule understanding, goal maintenance) are not too elevated for children above 6 years old as is our target sample. We propose that the process that places different demands within this age range is the inhibition of the inappropriate cues (distractor objects in the present chapter's task, incongruent arrows in the Flanker task or description of the image in the Day-night task) to allow one to execute the appropriate response. Therefore, the development of inhibitory control is a very likely explanation for the reduction of interference effects.

This finding is important given that most daily-life goal-directed sequences involve the selection and manipulation of targets among a number of objects irrelevant for the task (e.g., picking-up a toothbrush on a bathroom sink or cupboard, or picking up one's jacket on a coat rack). To develop action competence, children must be able to inhibit distractors to efficiently pick up the objects relevant to their goals.

2.5 Conclusion

This chapter found that distractor interference in a reaching action, as measured by response time or kinematic variables, depended on the number of distractors regardless of the type of compatibility between the target and distractors (perceptual or semantic). The strength of the interference effect depended on age in the range tested (6, 9, 12 years of age and adults), where the younger children showed the most marked increase in response time with the increase in the set size.

Our findings are consistent with theoretical views (premotor theory, visual attention model, visuomotor theories) advancing that when facing a target and distractor objects, multiple motor actions are represented and competing representations must be suppressed to select the target action. Our results further add that the competition is not the sole product of semantic, nor the sole product of perceptual object compatibility. Such properties appear to be considered simultaneously. The fact that the trajectory path is also affected reflects that action planning is not a sequential process where the preparation would occur prior to the start of the movement (measurable in response time only)- instead this is consistent with the view that a non-negligible part of motor planning is carried out on-line and that unfolding cognitive processes flow into motor outputs (e.g., Song and Nakayama, 2009, Freeman et al., 2011, Erb et al., 2018).

Importantly, we only found an interference (by increasing set size) in the child participants and not the adults, and the intensity of the interference on reaching duration decreased with age in children. This is consistent with the trends of improvements with age in resolving conflicting activations (e.g., reduction of the Stroop interference with age: Roy et al., 2018, Leon-Carrion et al., 2004). It highlights the crucial role of the maturation of executive functions, in particular inhibitory control, in supporting better action control with development.

Chapter 3

A Mechanistic Account of the Model-Free / Model-Based Trade-off: A Computational Study

Abstract

This chapter proposes a new model of value-based action selection via the interaction of two systems within the reinforcement learning framework, known as model-free and model-based systems. The chapter models pre-existing experimental data, via an alternative account of a proposed model of this experimental data. Specifically, the original study (Decker et al., 2016) argued that young children select actions using a mixture of habitual (model-free) and goal-directed (model-based) systems. It furthermore argues that the development of adult competence involves an increasing influence of the model-based system. After exploring the data and replicating the original study's model, the chapter proposes an interactive activation model that operationalises the combination of model-free / model-based influences on action selection and simulates response time in addition to choice data. The novel model also considers four competing developmental hypotheses that are modelled and pitted against each other. The model reproduces key behavioural patterns, like the reinforcement learning model of the original study of Decker et al. (2016), but also allows one to discriminate more precisely between several possible developmental mechanisms by leveraging response time data.

The findings point towards the idea that the difference in performance between the child and adult groups may be due to either greater noise in children's than adults' model-based system, or to a differing trade-off in children's and adults' recruitment of the systems. Nonetheless, further evidence is needed to separate both accounts and to exclude other accounts.

3.1 Introduction

Action selection by humans can be informed by past actions. The outcome (positive or negative) of an action taken in a given environment modulates the probability of repeating the action in a similar situation. The influence of an action's outcome on subsequent action selection to maximise expected future rewards has been successfully modelled using reinforcement learning (RL; Sutton and Barto, 1998). At the same time, many accounts of action control draw upon the distinction between two systems, whether they be routine and non-routine (e.g., Cooper and Shallice, 2000) or goal-directed and habitual (e.g., Daw et al., 2011).

Within the RL tradition, a dominant approach is to consider two systems that together direct action selection, with each system modelled by a particular reinforcement learning algorithm (Daw et al., 2011). The habitual system, implemented as model-free (MF) RL, attributes values to an action in a given environmental state following the action's final outcome. The goal-directed system, implemented as model-based (MB) RL, uses both outcomes and knowledge about how actions affect the environment to inform action selection. While most researchers seem to agree on the existence of two distinct systems, there remains some debate about how these systems combine to guide action selection. For example, the two systems have been proposed to compete based on speed-accuracy trade-offs (Keramati et al., 2011a, Kool et al., 2017b) or to be arbitrated by combined probabilities based on uncertainty (Daw et al., 2005). Yet these accounts are purely mathematical in nature. To our knowledge, there is no mechanistic account that satisfactorily explains how the values computed by the systems are effectively integrated.

3.1.1 The two-stage decision task

A standard task introduced by Daw et al. (2011) has emerged to quantify the relative contribution of each system to decision-making (the so-called two-stage Markov decision task, or *two-stage task* as will be referred to later). Each trial of the task consists of two successive stages, in which one must select one of two options (see Figure 3.1). The first stage of the task always starts in the same state (i.e., the same two options are presented). One of the first-stage actions (action A) leads to one second-stage state (state A) with a probability of 0.7 (*common transition*), and to another second-stage state (state B) with a probability of 0.3 (*rare transition*). The probabilities are swapped for the first-stage action (action B), as indicated in Figure 3.1. In the second state, there are two further possible actions. Each of the four second-stage actions (two actions per state) has a certain probability of reward. The reward is given for each second-stage action according to a slowly-drifting probability (between 0.25 and 0.75).

The task is designed to explore the trade-off between habitual / model-free and goal-directed / model-based systems. The key dependent variable is the probability of staying with the same first-stage choice on consecutive trials, as a function of state outcome and reward on the preceding trial. The two systems may be modelled using standard RL techniques as in Daw et al. (2011). The trade-off between the systems can then be quantified in terms of a weighted average between the systems' outputs, with the weight potentially varying by participant. We are specifically interested in

Decker et al.'s (2016) developmental study which uses the standard task to explore potential changes in the trade-off from childhood to adulthood. First, this chapter presents exploratory and replication work based on the original study (Section 3.2). Second, this chapter explores an alternative model of Decker et al.'s (2016) data (Section 3.3). The novel model provides a mechanistic account of the trade-off based on an interactive activation model in which the two systems activate, to varying degrees, the potential responses at each stage. The primary advantage of the model is that it addresses the time-course of the assumed trade-off between model-free and model-based systems. We fit our model to Decker et al.'s (2016) data and explore three mechanistic hypotheses underlying developmental changes in the trade-off between habitual / model-free and goal-directed / model-based systems. It is shown that, at a mechanistic level (i.e., Marr's algorithmic and representational level; Marr, 1982), all three hypotheses are consistent with the behaviour observed by Decker et al. (2016). It is furthermore argued, however, that additional empirical work might differentiate the hypotheses.

3.1.2 The study and model of Decker et al. (2016)

3.1.2.1 The task and model

The study of Decker et al. (2016) used a child-friendly version of the standard two-stage task to study developmental changes in action selection from childhood to adulthood. The narrative involved choosing between two spaceships (the first-stage options) leading to two possible planets (the second-stage states), where the choice was then between two aliens (the second-stage options) that had their own propensity to offer coins (the probability of reward) — see Figure 3.1. Three age groups (n=20 children between 8 and 12 years old, n=20 adolescents between 13 and 17 years old, n=20 and adults between 18 and 25 years old) performed 200 trials of the task.

Decker et al. (2016) were concerned with a variation of the standard 'hybrid reinforcement learning model' introduced by Daw et al. (2011). The critical aspect of the model (see Decker et al. (2016) for details) is that action selection at stage 1 involves computing both model-free and model-based values (Q_{MF} and Q_{MB}) for each action. Actions are then selected based on a weighted average of Q_{MF} and Q_{MB}, with the contribution of each system determined by inverse temperature parameters, β_{MF} and β_{MB}.

The model-free values are updated at each step via to the temporal difference algorithm according to the formulas:

$Q_{MF}\left(s_{i,t+1}, a_{i,t+1}\right) = Q_{MF}\left(s_{i,t}, a_{i,t}\right) + \alpha\delta_{i,t}$ where α is the learning rate, $s_{i,t}$ is the state i at the start of trial t, and $a_{i,t}$ is the action i taken at trial t;

and with δ the reward prediction error calculated as:

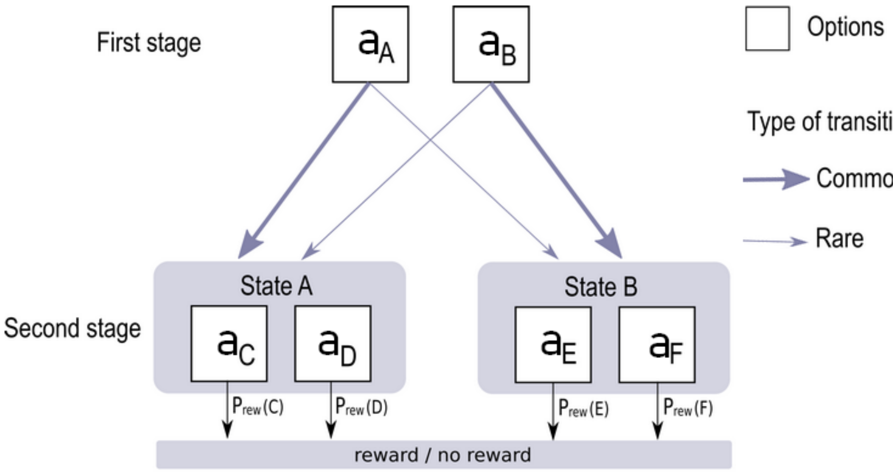


Fig. 3.1: **Structure of a trial of the standard task introduced by Daw et al. (2011) and used in Decker et al. (2016): options and consequences of options' selection at each stage (transitioning and receiving outcome). The actions a_x are denoted with different letters because they bring about different outcomes, but they all consisted of a left or right key press. By selecting the action a_x, there was a 0.7 probability to reach the state s_x (*common* transition) and a 0.3 probability to reach $s_{x'}$ (*rare* transition). By selecting the action a'_x, there was a 0.7 probability to reach the state s'_x (*common* transition), and a 0.3 probability to reach s_x (*rare* transition). $P_{rew}(a_x)$ is the probability of reward of the action 'selecting option a_x'.**

$\delta_{i,t} = [r_{i,t} + Q_{MF}(s_{i+1,t}, a_{i+1,t})] / \alpha - Q_{MF}(s_{i,t}, a_{i,t})$ where r is the reward (0 at stage 1, 0 or 1 at stage 2);
and for the first stage only:

$Q_{MF}(s_{1,t+1}, a_{1,t+1}) = Q_{MF}(s_{1,t}, a_{1,t}) + \lambda \delta_{2,t}$ where λ is the eligibility trace which is carried over stages for one trial.

The model-based values at stage 2 are the same as for model-free. At stage 1, they are updated according to the formula:

$$\begin{aligned} Q_{MB}(s_A, a_j) =& P(s_B \mid s_A, a_j) \max_{a \in \{a_A, a_B\}} Q_{MF}(s_B, a) \\ &+ P(s_C \mid s_A, a_j) \max_{a \in \{a_A, a_B\}} Q_{MF}(s_C, a) \end{aligned}$$

where s_A is the unique state at stage 1, s_B and s_C are the two possible states at stage 2, a_A and a_B are the possible actions at stage 1, and the probability of state according to the action taken follows the 70/30-transition structure illustrated by the arrows in Figure 3.1.

The probability of choosing action a at trial t, being in the first-stage state $s_{1,t}$, given $Q_{MF}(s_{1,t}, a)$ and $Q_{MB}(s_{1,t}, a)$ is calculated as:

$$P(a_{1,t} = a|s_{1,t}) = \frac{exp[\beta_{MF} \cdot Q_{MF}(s_{1,t}, a) + \beta_{MB} \cdot Q_{MB}(s_{1,t}, a) + p \cdot rep(a)]}{\sum_{a_1} exp[\beta_{MF} \cdot Q_{MF}(s_{1,t}, a_1) + \beta_{MB} \cdot Q_{MB}(s_{1,t}, a_1) + p \cdot rep(a_1)} \tag{3.1}$$

where $rep(a)$ is 1 if a is the same as the action on the previous time step and 0 otherwise; β_{MF} (resp. β_{MB}) is the inverse temperature parameter for MF (resp. MB) valuation; and p is the perseveration bias.

At stage 2, the probability of choosing action a is computed simply with one term $Q_{MF}(s_{2,t})$ and one temperature parameter β_2:

$$P(a_{2,t} = a|s_{2,t}) = \frac{exp[\beta_2 \cdot Q_{MF}(s_{2,t}, a) + p \cdot rep(a)]}{\sum_{a_2} exp[\beta_2 \cdot Q_{MF}(s_{2,t}, a_2) + p \cdot rep(a_2)} \tag{3.2}$$

Larger values of β (β_{MF}, β_{MB} or β_2) favour exploitation of the corresponding system, while the perseveration bias reflects the tendency to repeat the same action regardless of the outcome.

Decker et al. (2016) fit the model to their participants' data to estimate the value of the six parameters: β_{MF}, β_{MB}, $alpha$, p, β_2, λ.

3.1.2.2 The characteristic patterns of choice or stay patterns

To observe the type of control (MF/MB balance) used by an agent in the task, a standard approach is to plot the aggregated proportions of stay choices (i.e., repetition of the preceding choice) as a function of the preceding trial's outcome and the preceding trial's transition type, as in Figure 3.2. Indeed, under MF control an agent is more likely to repeat a previously rewarded action than an unrewarded one, regardless of the second-stage state that led there (i.e., ignoring the transition type), as the left panel of Figure 3.2 shows. Under MB control, an agent will integrate the information provided about the second-stage state that was rewarding, and thus be more likely to repeat a rewarded action if it followed a common transition, or an unrewarded action if it followed a rare transition, than the two other cases, as shown in the right panel of Figure 3.2. Fitting a hybrid RL model trial-by-trial, Decker et al. (2016) found a significant increase in the contribution of the model-based system (β_{MB}) with age, but no significant change in the contribution of the MF system (β_{MF}) with age.

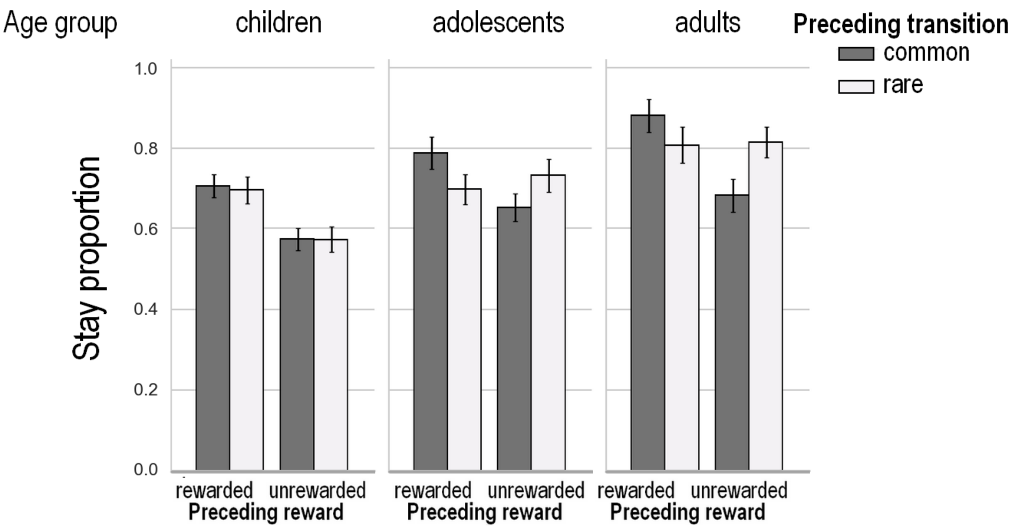


Fig. 3.2: **Proportion of stay as a function of age and transition (*stay patterns*) from Decker et al.'s (2016) data. The figure shows the proportion of stay (repeating the preceding trial's action) at first stage as a function of the preceding trial's reward and preceding trial's transition type, and age. Error bars represent the standard error of the mean.**

3.1.3 An alternative approach

Despite the apparent empirical success of Decker et al.'s (2016) account, there are several limitations of the work. For example, the conclusions are valid only if the hybrid RL model is valid for all age groups. Yet, Decker et al. (2016) do not report goodness-of-fit measures by age group, nor do they compare their model to alternative models. Thirdly, the use of separate inverse temperature parameters for each system, rather than a single parameter controlling the balance between MF and MB contributions, complicates the interpretation of the results. Fourthly, as others have pointed out (e.g., Ballard and McClure, 2019, Shahar et al., 2019), the identifiability and recoverability of parameters in such models is often problematic. Indeed, we have found similar problems in our attempt to replicate the parameter estimation with the same hybrid RL model specification as Decker et al. (2016) (cf. Section 3.2.2). Therefore, and combining these last two concerns, accounts other than changing the MB temperature parameter may explain the results equally well if not better. Furthermore, Decker et al.'s (2016) model provides a mathematical account of action selection in the task but does not provide a mechanistic or process-level account, or a way to generate response time data. The model we report in Section 3.3 attempts to do just this.

In addition to proposing a process-level account that makes contact with response time data, the new model will implement and test four developmental hypotheses to account for the changing model-free/model-based balance across the age groups in Decker et al.'s (2016) study. A first hypothesis is that the recruitment of the model-free and model-based systems is externally arbitrated and favours model-based more and more with age (*recruitment* hypothesis). In the other three hypotheses, the model-based system is assumed to be recruited at the same level throughout development while another mechanism explains the apparent increase in contribution of model-based with age. According to the second (*noise*) hypothesis, this mechanism is a reduction in the noise in model-based computations with age. According to the third (*bias*) hypothesis, the mechanism is an increase in speed of the model-based system computations. According to the last (*learning rate; lr*) hypothesis, the mechanism in an increase of learning rate with age. The background of each hypothesis will be presented in Section 3.3.2.3. The simulations according to the different hypotheses will be compared against choice data and response time data.

In order to prepare the design of the new model and attempt to obtain suggestive evidence for the developmental hypotheses, work was carried out on the original data and model and is presented in the next Section.

3.2 Original data and model: Exploratory work and replication

3.2.1 New analyses - RT first stage

This Section presents our exploration of the response times at stage 1 of Decker et al.'s (2016) task with the aim to gain information on the mechanisms underlying choice and build a more precise model of the task. Stage 1 was chosen given the central role of choices at stage 1 and the fact that such data did not appear in the original publication. Only analyses of stage 2 response times were present in the original publication but all the data was kindly made available to us by the authors.

3.2.1.1 Methods

We ran analyses guided by four different assumptions about potential processes underlying choices at stage 1.

(1) The reward at stage 1 may have been affected by the type of transition in the previous trial. Indeed, the rare transitions led to slower response time at stage 2 compared to the common transitions (as proposed by the author, reflecting surprise), and we assumed that that such effect might be carried on to the next trial. Thus, we explored response times at stage 1 (*RT_stage1*) as a function of the transition type of the preceding trial.

(2) A phenomenon known as *post-error slowing* (response time decrease following an incorrect trial) is commonly found in adults and children (e.g, Smigasiewicz et al., 2020). This has been explained as an adjustment in decision criteria (Laming, 1979) or longer processing of unexpected events (Wessel et al., 2012). If we assume that a

rewarded trial is perceived by participants as a correct trial, we could expect a slowing after low rewards compared to higher rewards. We thus examined *RT_stage1* as a function of reward on the preceding trial.

(3) Another hypothesis is that the response time increases with increasing conflict between the *preference* (i.e., Q-value) provided by the MF and the MB pathways. Indeed, it is well-known that response times increase with the difficulty of the decision (e.g., Ratcliff and McKoon, 2008). For example, if in the MF system the Q-value is much larger for option A than option B, but in the MB system the Q-value is much smaller for option A than option B, the *preferences* for both options go in opposite directions. This is what we will call *conflict*. We thus examined *RT_stage1* as a function of the normalised amount of conflict (*normedConflict*), which was computed as follows:

$$conflict = |Q_{MF}\_optionA - Q_{MF}\_optionB + Q_{MB}\_optionA - Q_{MB}\_optionB|$$

$$normedConflict = \frac{conflictPref}{Q_{MB}\_optionA + Q_{MB}\_optionA + Q_{MF}\_optionB + Q_{MB}\_optionB}$$

(4) The last hypothesis is that response time depends on the amount of involvement of the MB system. To evaluate this, we computed a score of degree of MB recruitment based on choice patterns, *MB score*, as follows:

$MBscore = |p_{rewarded,common} - p_{rewarded,rare}| + |p_{unrewarded,common} - p_{unrewarded,rare}|$, where p are the individual proportions in each condition.

We created this *MB score* to approximate the degree of model-based recruitment for an individual participant's stay pattern, for the purpose of this response time analysis. For an ideal MF learner, there is no difference according to transition type, so this score will be 0. For an ideal MB learner, the amount of probability difference between common and rare transitions in the rewarded case is equal to the (absolute) amount of difference in the unrewarded case. Thus, using this approximation, the MB score will increase when model-based involvement increases. As an example, the values of *MB score* computed on the replication of the participants' data shown in Figure 3.2 was 0.018 for the children, 0.142 for the adolescents and 0.192 for the adults.

Finally, we sort the data into four categories based on their MB score and we examine the *RT_stage1* as a function of the score category and age for the children's and adults' data (the age group with the most extremes scores).

3.2.1.2 Results

Figure 3.3 displays *RT_stage1* as a function of transition of the preceding trial. Figure 3.4 shows *RT_stage1* as a function of reward on the preceding trial. The response time at stage 1 appears unaffected by the preceding trial's transition type (unlike response times at stage 2). It also appears unaffected by the preceding trial's reward.

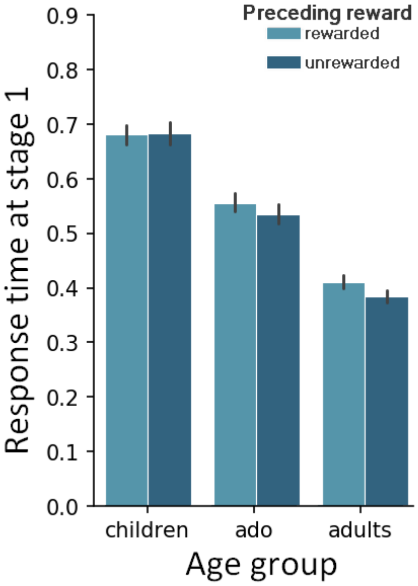


***Fig. 3.3:* Response time at stage 1 as a function of transition of the preceding trial's transition (*prevTrans*) using data not presented in Decker et al. (2016) but made available by the authors. The error bars represent the 95% confidence intervals. Ado = adolescents.**

Figure 3.5 represents *RT_stage1* as a function of *normedConflict*. Figure 3.6 presents *RT_stage1* as a function of the category of value of the MB score. Unsurprisingly, the response times are longer in children than adults. However, the response time differences between conditions (within a given age group) do not seem to differ between children and adults. Thus, the inspection of the response times at stage 1 did not reveal any meaningful patterns that could guide our modelling of the task.

3.2.1.3 Discussion

To conclude, the response times at stage 1 did not show any interesting patterns that provide any insights into the mechanisms underlying task performance in children or adults. It is likely that such underlying processes are too complex be revealed by simple assumptions. For example, it may be that there is an interaction of all factors that we considered separately, although we could still have observed the effect separately if they were large enough. The data at first stage could also be too noisy. However, knowledge from sources other than this data can inform the building of our new model. This is the approach we will take. Following this, further

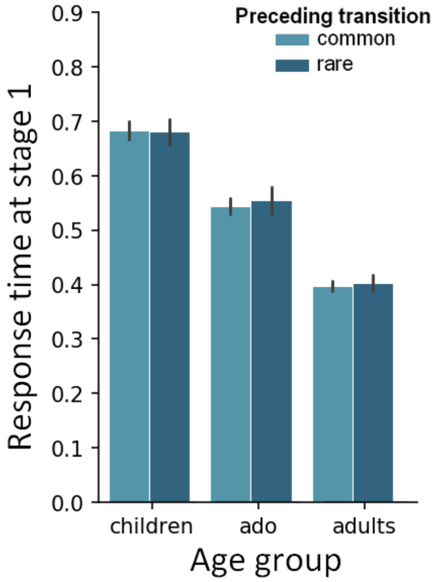


Fig. 3.4: **Response time at stage 1 as a function of reward on the preceding trial's reward (*prevReward*) using data not presented in Decker et al. (2016) but made available by the authors. The error bars represent the 95% confidence intervals. Ado = adolescents.**

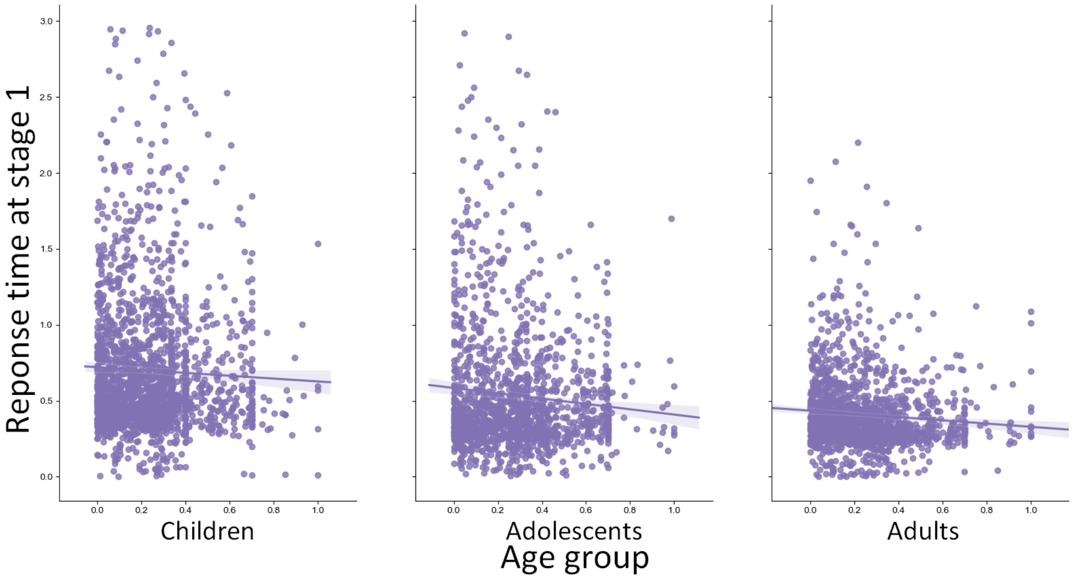


Fig. 3.5: **Response time at stage 1 as a function of *normedConflict* using data not presented in Decker et al. (2016) but made available by the authors. A simple regression line is plotted with the shaded area representing the 95% confidence interval.**

analyses using new model-driven variables might be able to uncover some information in response times at stage 1 (e.g., by confirming some model predictions on RT distributions).

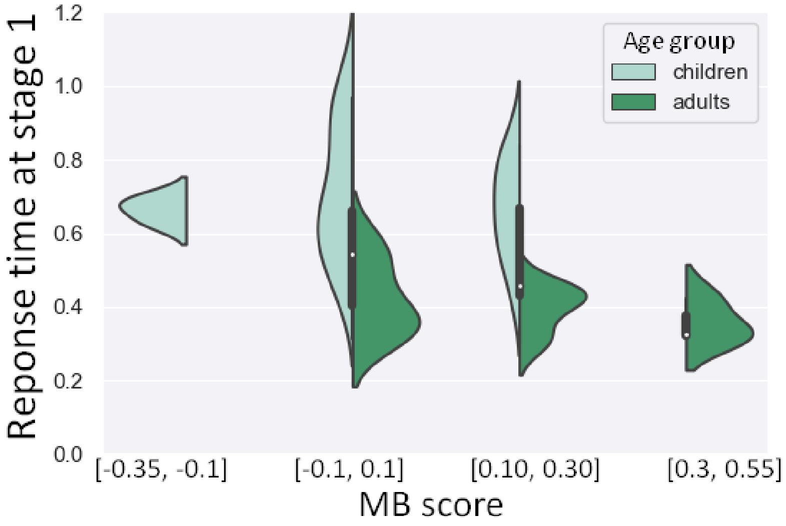


Fig. 3.6: **Response time at stage 1 as a function of the value of *MB score* using data not presented in Decker et al. (2016) but made available by the authors.**

3.2.2 Replication of the RL modelling in Decker et al. (2016)

3.2.2.1 Methods

We aimed to replicate the RL model of Decker et al. (2016) and its fitting to the data. We built the hybrid model-free/model-based RL model following the methods and supplementary methods of Decker et al. (2016), as well as the methods and supplementary methods of Otto et al. (2013). Our model-fitting (parameter estimation) procedure is described below. Contrary to the authors who used hierarchical Bayesian model-fitting, we used the more straightforward individual fitting approach used for example in Daw et al. (2011) and Palminteri, Lefebvre, et al., 2017, described below.

Fitting procedure

We estimated the six free parameters α, λ, p, β_{MF}, β_{MB} and β_2 for each participant separately. Specifically, we found the best-fitting values by minimising the negative log-likelihood of the data (obtained from the log of equation 3.1 summed over all trials of the participant). This was implemented with the Python function *basinhopping* from the module *scipy.optimize*, which uses the *L-BFGS-B* method (Byrd et al., 2006). The basin-hoping algorithm is a stochastic algorithm designed to find the global minimum of a function by iteratively giving random perturbation of the coordinates and performing local minimisation. The number of iterations was set as 20 or would stop if the global minimum candidate remained identical for 6 iterations of the algorithm. Following Daw et al. (2011), we constrained the learning rate α between 0 and 1, while the other parameters only had a lower bound of 0.

Model recovery

To assess the sanity of the model and reliability of our fitting approach, we ran a recovery analysis (following for example Ballard and McClure, 2019 and Palminteri, Lefebvre, et al. (2017)). This consists of generating an artificial dataset with a randomly defined set of generating parameters, applying the fitting procedure, and comparing the fitted parameters to the generating parameters (i.e., the ground truth). We repeated this 50 times with a new set of parameters drawn every time from: a beta distribution $B(1.5, 2)$ for α, λ and p; and a re-scaled beta distribution $B(2, 6) * 10 + 0.5$ for β_{MF}, β_{MB} and β_2. The distributions were chosen to match the expected distribution for the entire population following Ballard and McClure (2019). We also ensured that they included the estimates we found on the real dataset. We correlated the fitted parameter to the true, generating parameters. We repeated the procedure with a pure MF agent (i.e., the β_{MB} held fixed at 0) and a pure MB agent (i.e, the β_{MF} kept at 0).

Replication of behavioural patterns

Finally, following recommendations in Palminteri, Lefebvre, et al. (2017), we verified that we were able to reproduce the key behavioural patterns in the following conditions: (1) by re-simulating data with our replicated RL model, (2) while using the best-fitting parameters found via our estimation procedure. Specifically, we took the median of the participants' best-fitting parameters as set of parameters for the simulations.

3.2.2.2 Results

Model recovery

Figure 3.7 presents the fitted parameters plotted against the generating parameters. The correlations were high for α and β_2 (R=0.78 and 0.63), moderate for β_{MB} and β_{MF} (R=0.41 and 0.31) and poor for λ (R=0.17).

When holding the β_{MB} or the β_{MF} constant to simulate a pure MB or pure MF agent (figures 3.8 and 3.9), the recovery is better as evidenced with higher correlations (α, β_2 and p above 0.6, β_{MB} and β_{MF} between 0.4 and 0.5) except for λ which is still poorly estimated in the pure MB case (R=-0.2).

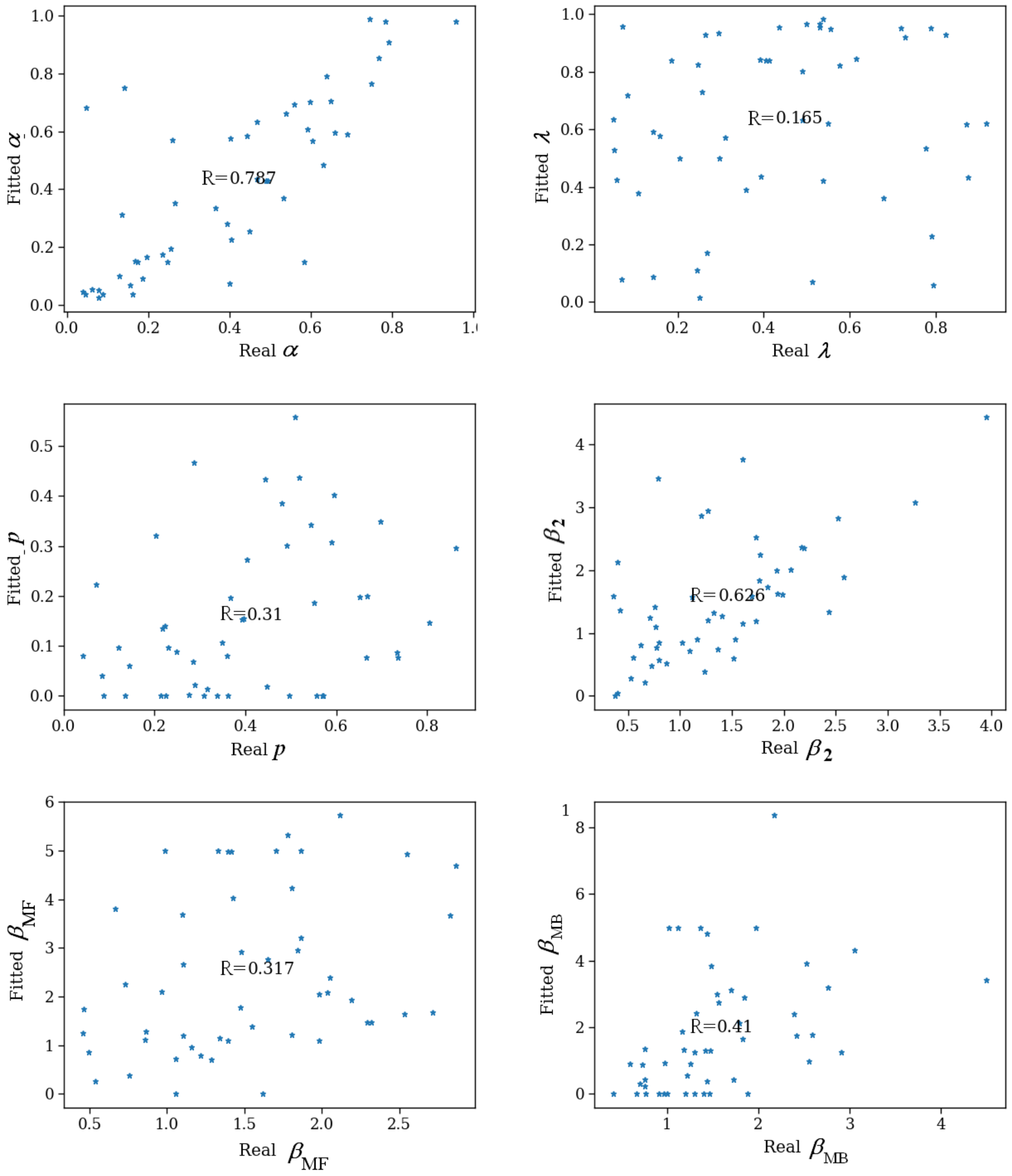


Fig. 3.7: Original RL model recovery: estimated (fitted) parameters values against the real, generating values. The value displayed on each scatter plot is the Spearman correlation coefficient R. The different panels are for the different parameters as follows: α: learning rate, λ: eligibility trace, p: persistence parameter, β_2: inverse temperature parameter at second stage, β_{MF}: inverse temperature parameter for model-free computations at first stage, β_{MB}: inverse temperature parameter for model-based computations at second stage.

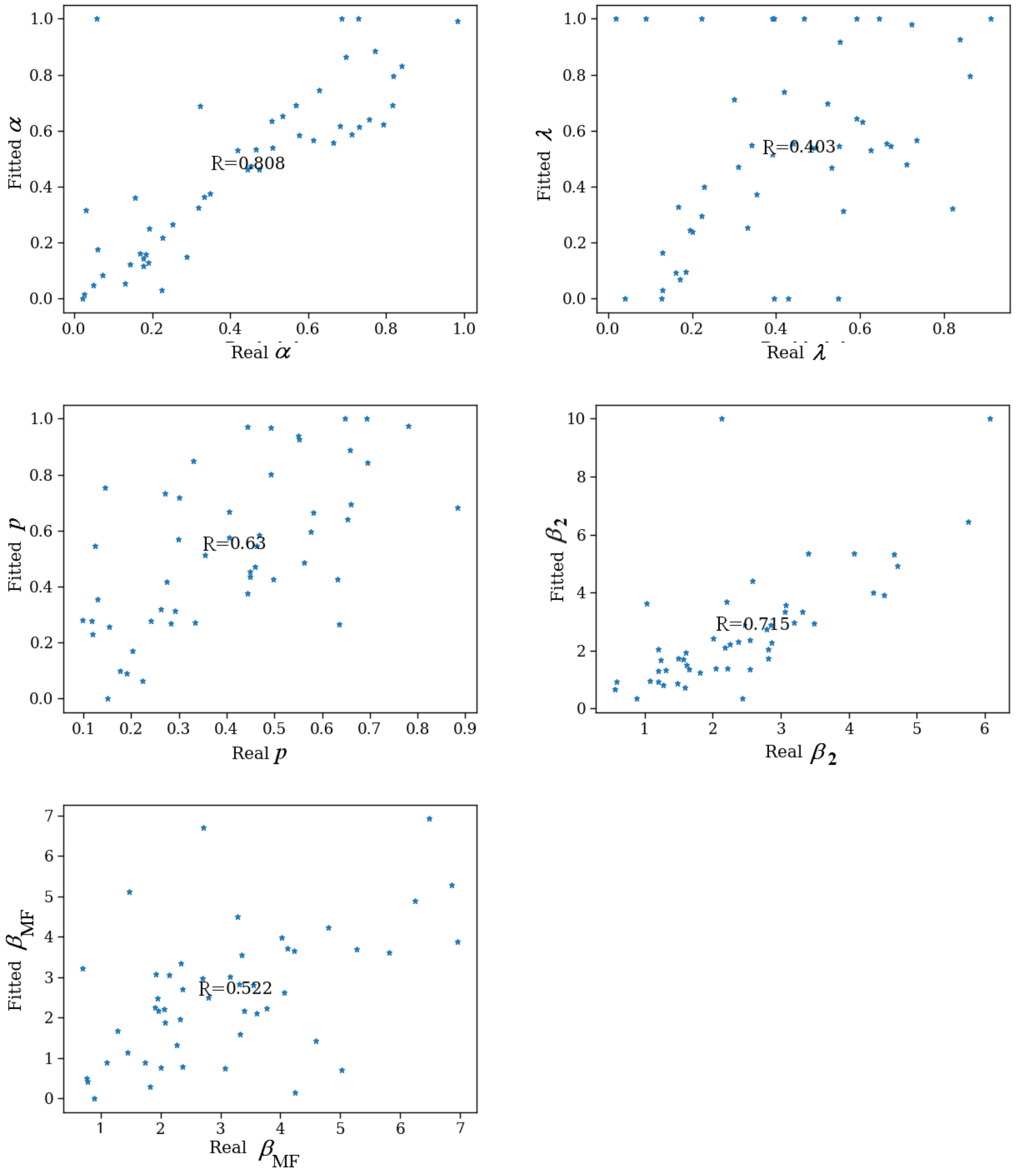


Fig. 3.8: **Original RL model recovery: estimated (fitted) parameters values against the real, generating values. Pure MF agent (β_{MB}, the inverse temperature parameter for model-based computations at second stage, was held at 0). The value displayed on each scatter plot is the Spearman correlation coefficient R. The different panels are for the different parameters as follows: α: learning rate, λ: eligibility trace, p: persistence parameter, β_2: inverse temperature parameter at second stage, β_{MF}: inverse temperature parameter for model-free computations at first stage.**

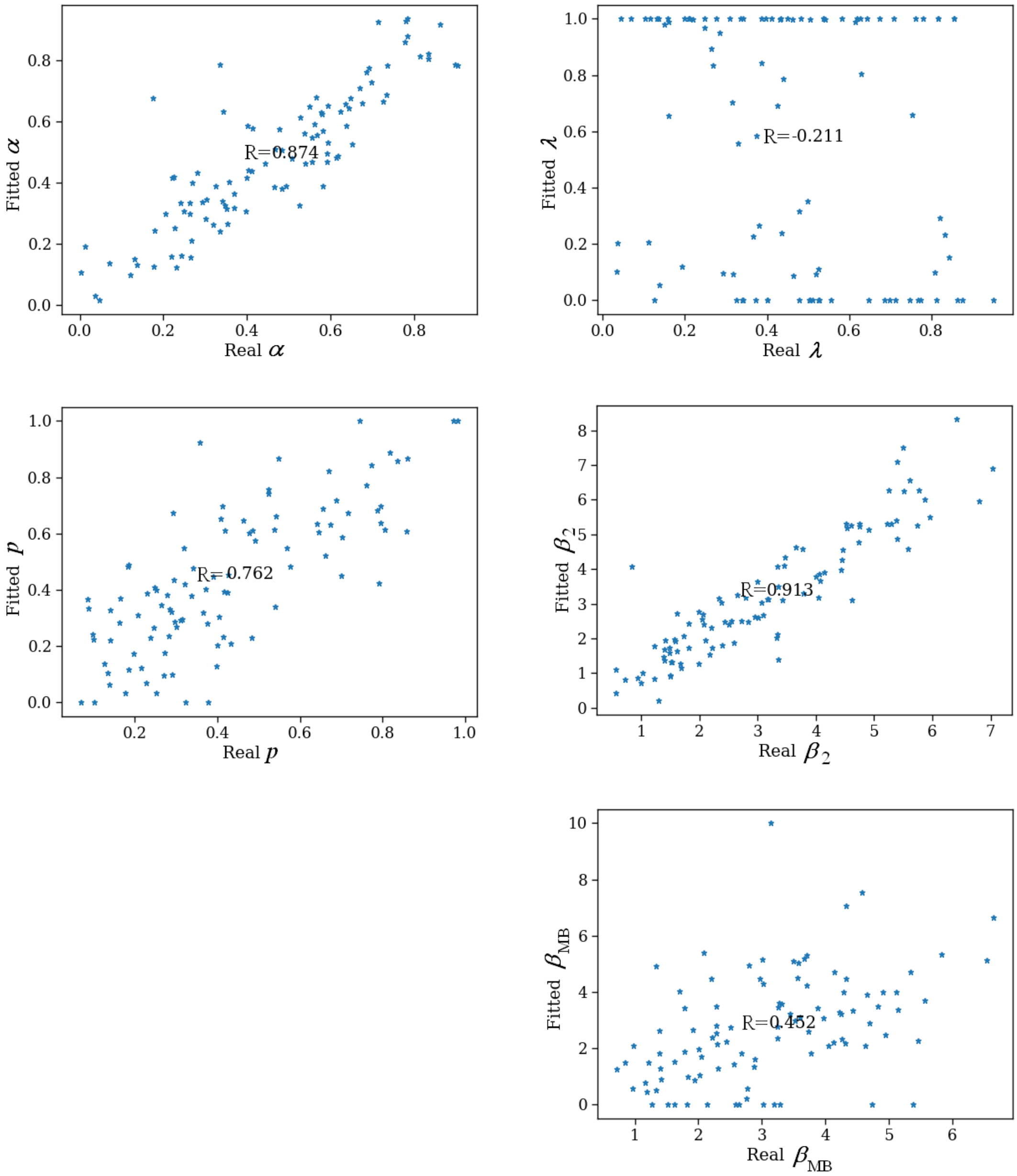


Fig. 3.9: **Original RL model recovery: estimated (fitted) parameters values against the real, generating values. Pure MB agent (β_{MF}, the inverse temperature parameter for model-free computations at first stage, was held at 0). The value display in each scatter plot is the Spearman correlation coefficient R. The different panels are for the different parameters as follows: α: learning rate, λ: eligibility trace, p: persistence parameter, β_2: inverse temperature parameter at second stage, β_{MB}: inverse temperature parameter for model-based computations at second stage.**

Best-fitting parameters

The values of the individual parameters estimated by our fitting procedure are shown in Figure 3.10 (all participants collapsed) and in Figure 3.11 (participants split by age group). For comparison, we overlay the estimates obtained by the hierarchical fitting procedure used in Decker et al. (2016). As can be seen, the parameter estimates we obtained do not match theirs in terms of median or dispersion. Most of the original study's values are comprised in the inter-quartile range (IQR) of values obtained by our procedure. However, our IQR is for most parameters an order of magnitude larger than the original study's IQR, which is a considerable difference.

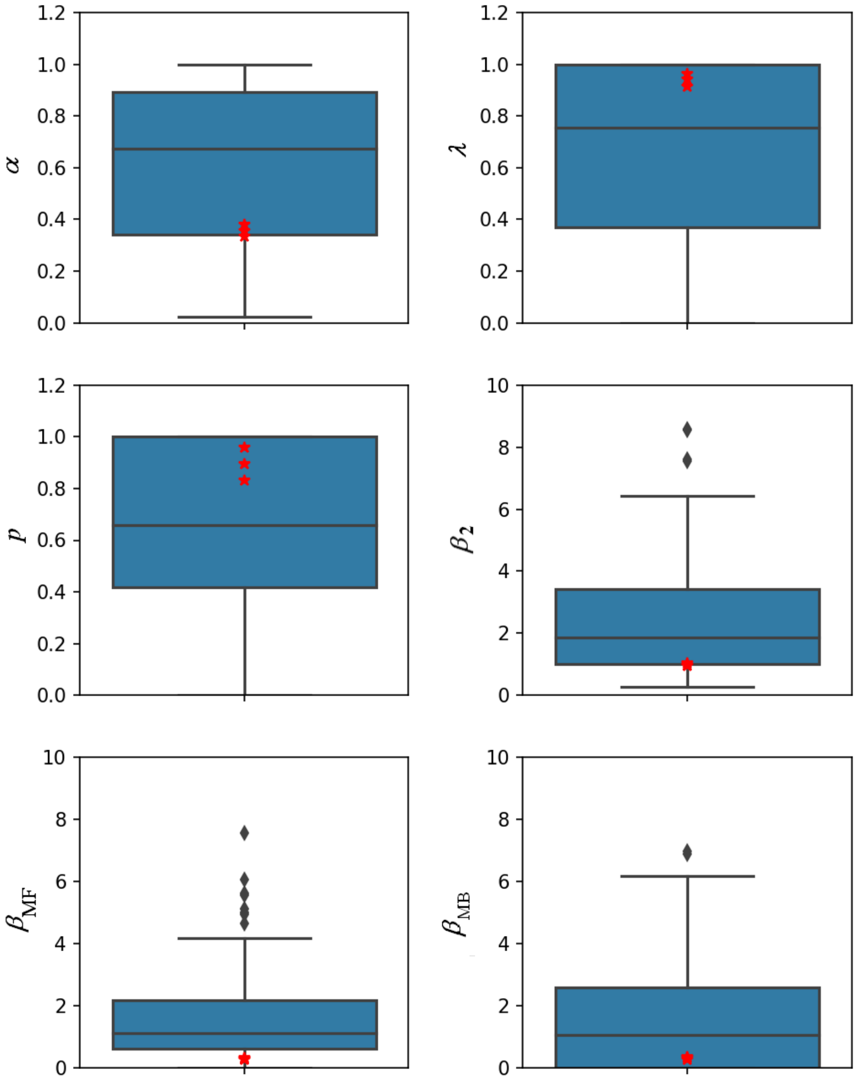


***Fig. 3.10:* Replication of the RL model: parameters estimated by a single-participants' fitting procedure on all ages grouped together. The boxplots follow Tukey's original definition. The red stars indicate (from bottom to top): the 25th percentile, the median and the 75th percentile of the group level estimated obtained by the hierarchical fitting procedure used in Decker et al. (2016).**

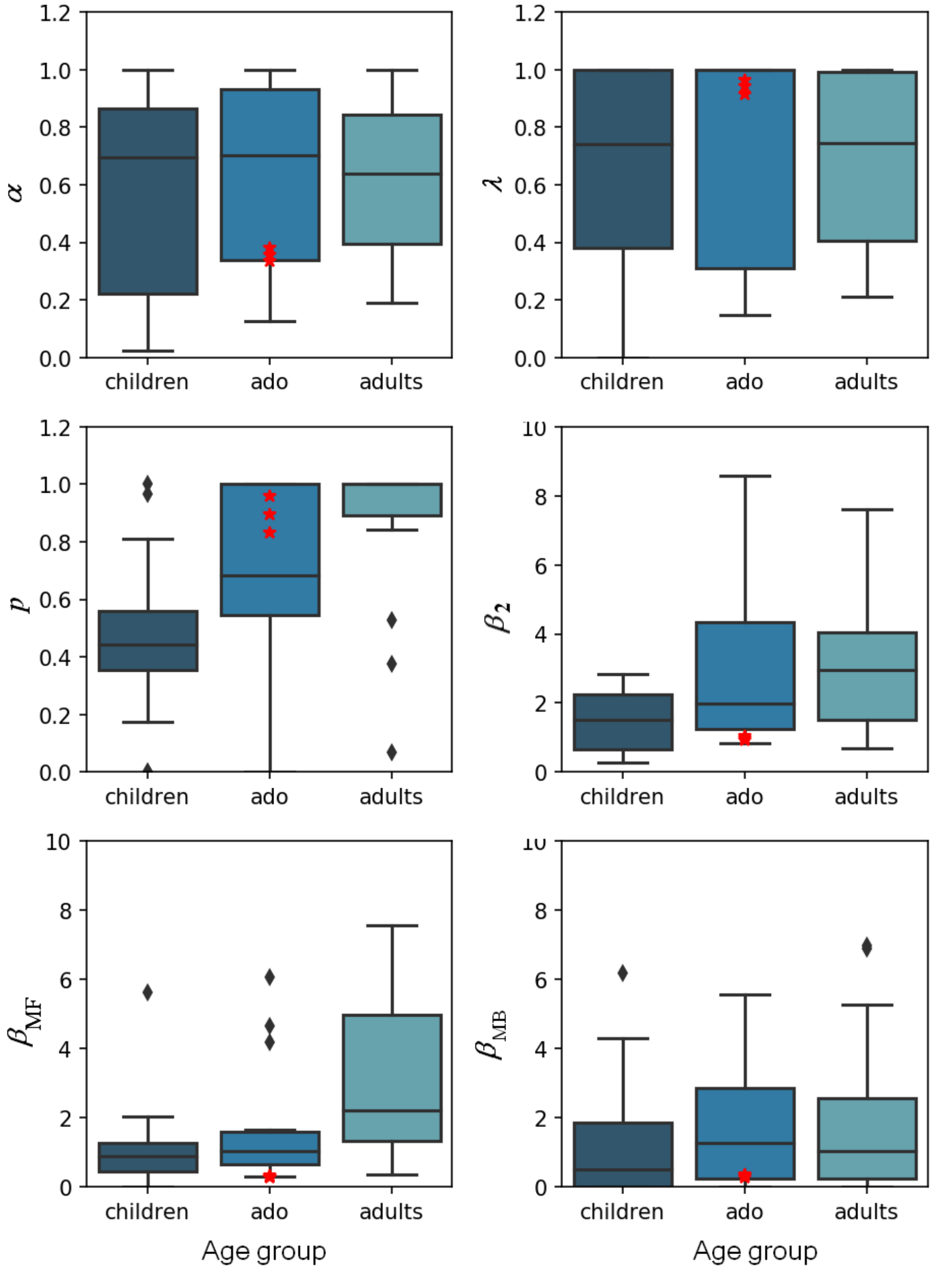


Fig. 3.11: **Replication of the RL model: parameters estimated by a single-participants' fitting procedure, shown by age group. The boxplots follow Tukey's original definition. The red stars indicate (from bottom to top): the 25th percentile, the median and the 75th percentile of the group level estimated obtained by the hierarchical fitting procedure used in Decker et al. (2016). The different panels are for the different parameters as follows: α: learning rate, λ: eligibility trace, p: persistence parameter, β_2: inverse temperature parameter at second stage, β_{MF}: inverse temperature parameter for model-free computations at first stage, β_{MB}: inverse temperature parameter for model-based computations at second stage. Ado = adolescents.**

Replication of behavioural patterns

Figure 3.12 presents the stay patterns simulated by using the median of the fitted parameters in each age group as an input parameter of our replicated RL model. It appears that the stay patterns are close to the original value, reproducing the typical pattern of the model-free/model-based balance in each age group.

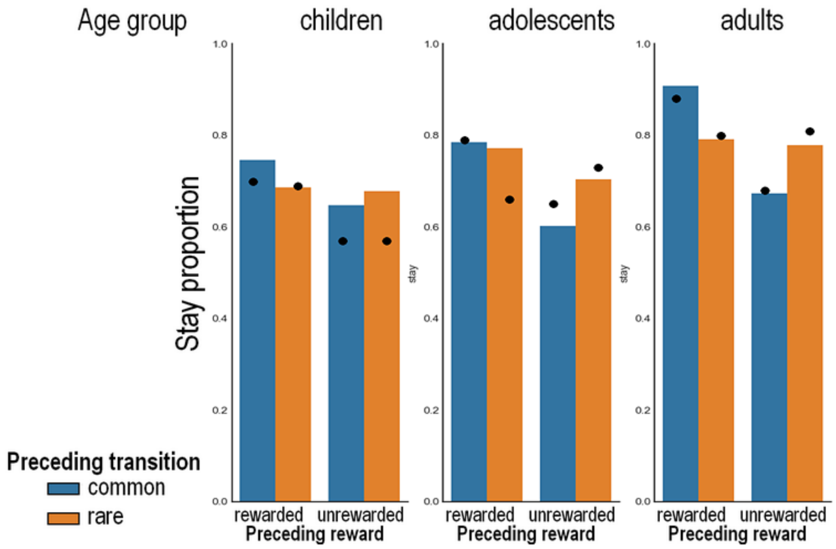


Fig. 3.12: **Generation of data with the fitted parameters in the three age groups. The bar indicates the simulated proportion of stay while the black dots indicates the data's proportion of stay. The different panels are for the different parameters as follows: α: learning rate, λ: eligibility trace, p: persistence parameter, β_2: inverse temperature parameter at second stage, β_{MF}: inverse temperature parameter for model-free computations at first stage, β_{MB}: inverse temperature parameter for model-based computations at second stage.**

3.2.2.3 Discussion

Our fitted parameters differed greatly from the median of Decker et al.'s (2016) fitted parameters. Nonetheless, the values of Decker et al. (2016) fall within our range of values. The large dispersion of our values was notable. However, a larger range of values for our procedure than for a hierarchical Bayesian procedure is not surprising. Indeed the hierarchical procedure estimates individual and group parameters in a mutually constraining fashion (M. D. Lee, 2011). Consequently, individual parameters tend to be less dispersed away from the group (Ahn et al., 2011).

Such large differences between our estimates and Decker et al.'s (2016) estimates raise the alarm concerning the sensitivity of findings to different approaches. This concern is important because the individual-participants fitting procedure used here has also been used by many in similar tasks. Such a procedure may be appropriate only when limited to a set of 2-3 parameters.

Some clarification is given by the recovery analysis. Indeed, when generating datasets with randomly chosen but known parameters, the fitting procedure did not always return the correct (generating) parameters. As expected, the recovery was improved when reducing the number of parameters (from 6 to 5). Even then, it remained unsatisfactory.

All in all, the results of our replication attempt raise two concerning points that are not unrelated. First, the estimates do not match the ones in the original study well. Second, the recovery procedure is sometimes accurate, but at other times gives parameter estimates that are too different from the ones used to generate the data. The recovery procedure indicates that the parameters estimated by the single-participant approach on the real dataset may not be reliable, which may explain the mismatch between our results and Decker et al.'s (2016).

The lack of match between the original study's and our parameter estimates has most likely one of two, or even both causes: (1) this dataset is too ambiguous and/or not rich enough to allow reliable estimation by a single-participant fitting procedure (as opposed to a hierarchical procedure); (2) our attempt at replicating the original model varies at the level of minor specifications that, even if capable of simulating similar behavioural patterns, changes the values of the parameters needed to generate such patterns.

The first cause is plausible given the inaccuracy of the parameter estimation revealed by the model recovery. Such concerns on the identifiability of parameters have been raised by others (Ballard and McClure, 2019, Nussenbaum and Hartley, 2019, Pedersen et al., 2017, and Toyama et al., 2019). This limitation of the RL model, to be able to fit only a certain range of behaviours, has been put forward by Daw (2011). The problem often arises when parameters are correlated. Ballard and McClure (2019) showed that there is a trade-off between the learning rate α and the decision noise β, where numerous combinations of α and β give the same log-likelihood. That is true even for RL models with 3 to 4 free parameters. The more complex the model, the more difficult it becomes to get reliable parameter estimates (Pedersen et al., 2017). Using additional data such as response times in order to constrain further the estimation and consequently improve parameter identifiability has been proposed, for example, by Ballard and McClure (2019).

Decker et al. (2016) used a hierarchical Bayesian fitting procedure. Such procedure is advised when the number of participants is large but their number of trials is small. It further constrains the estimation by having group trends inform the individual parameter values (Ahn et al., 2011, M. D. Lee, 2011). It is possible that parameter identifiability in complex models such as this one can only be resolved with such procedure. Nonetheless, the choice to use an individual fitting approach was motivated by the fact that it makes fewer assumptions on the data compared to the hierarchical fitting approach. Indeed, the latter assumes a common underlying distribution for participants within the same group. Although differences are expected when using the two fitting approaches, if the differences remain small, this indicates robustness of the RL model. If the differences are not small, this warrants some caution about drawing conclusions from the study alone.

The other possible cause relates to the more general problem of lack of specification of models when described in plain language as opposed to making the code available (see for example Cooper et al., 1996, Cooper and Guest, 2014 and Mikowski et al., 2018). For example, in our work, we noticed after running simulations that we were able to produce similar behaviour but with a range of β_{MF} and β_{MB} between 2 and 8, which are similar to the values in Daw et al. (2011), but far off from the values in Decker et al. (2016) (around 0.2 or 0.5). This was due to a crucial model specification described in Otto et al. (2013) (supplementary materials) which we initially missed: the Q-values of the non-chosen action were decayed after each step by multiplying them by $1 - \alpha$. Furthermore, some specifications like the initial Q-values may appear obvious, and thus writing them may be overlooked. We found no mention of the initialisation value in Otto et al. (2013) nor in Decker et al. (2016) but we assumed the default 0. However, while the default might indeed be 0, others initialise the Q-values at 0.5 (e.g., Ballard and McClure, 2019). Finally, Da Silva and Hare (2018) have shown how minor changes in the reward function can fundamentally affect behavioural patterns in the two-stage task. The extra specification (Q-value decay) affects the order of magnitude of βs but not the accuracy of the recovery. Thus, identifiability with such a large number of parameters remains an issue.

Small adjustments in the current RL model could slightly improve the fit of the data and perhaps even refine the developmental pattern. For example, keeping the β parameters constant assumes that the contribution of each system (plus the β at stage 2) is fixed throughout the task. A single β (per system) does make sense in terms of the overall balance of system for a given individual at a given age, but if this factor could vary from trial-to-trial, it might provide better fit. Importantly, this would be at the expense of the identifiability of the parameters as discussed previously. We know that children tend to explore more than adults (see Nussenbaum and Hartley (2019) for a review of reinforcement-learning accounts of exploration). Thus, we could expect the following developmental patterns on within-experiment change in β (all three β_{MF}, β_{MB} and β_2 indistinctly). Adults may have an initial high β which then decreases throughout the task (explore at first then exploit), while the children β may stay more constant (explore throughout). Such adjustments may be interesting to explore; however, we do not expect them to result in fundamentally different outcomes.

Rather than seeking to adjust the existing model or to replicate the exact fitting procedure as in the original study (which would retain the identifiability issues linked to the high number of parameters), we moved on to building a mechanistic model of Decker et al.'s (2016) task.

3.3 A novel model

3.3.1 Rationale and methods

3.3.1.1 Model and simulation of choices

Our specific concern is with the integration of the MF and MB systems for action selection, and in particular with action selection at the first stage of the task (the second stage being solely under control of the MF system). We model this integration in mechanistic terms through interactive activation and competition (IAC). We chose an interactive activation and competition type of model for its ability to implement a structure with diverse influences, where the evolution in time of the activation flow naturally generates a time dimension (number of cycles) that can be taken to represent the processing time of the system.

Figure 3.13 presents the architecture of our model. The model comprises three layers, and represents two pathways: the model-free (MF), and the model-based (MB) pathway.

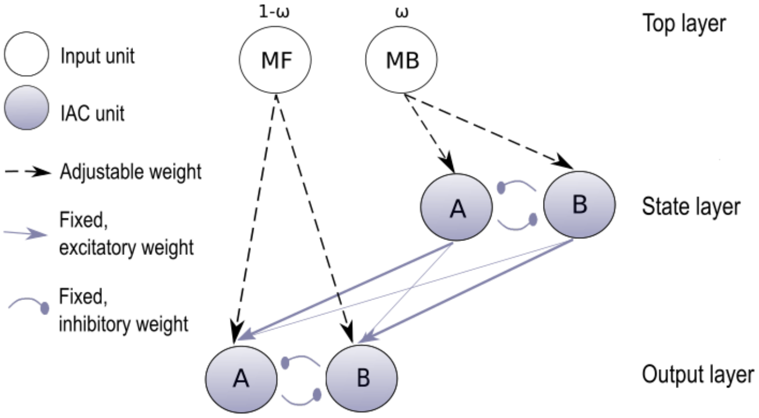


***Fig. 3.13:* Architecture of the model. The *input units* take a fixed value throughout one trial, determined by ω. The *IAC units* function according to IAC principles. Dotted, adjustable weights are initialised to 0. Thick weights are fixed at 0.7, and thin weights fixed at 0.3. Inhibitory weights are fixed at -2.**

The MF-MB layer, with an MF- and an MB-unit, feeds input to each pathway. The output layer represents the actions (unit A and B for actions A and B respectively). The MF-unit is connected to the output units by connections whose strength are adjusted as described below. The MB-unit is connected to an intermediate layer whose units (state-A and state-B) act as internal representations of the final states, with their activation rising or falling according to each state's appraisal. The weight of the connection from state-A unit to output-A unit is 0.7, and 0.3 from state-A to output-B (conversely for state-B), to embody the internal representation of the task transitions.[1] Finally, reciprocal inhibitory connections between state (respectively output) units facilitate the increase in activation differences (between output units A and B, or between state units A and B), and model the competition between alternative states (respectively actions). These lateral connections ensure that one element will eventually be sufficiently more active than the other so that the threshold is reached and a selection is performed.

Note that the MF layer is directly connected to the output units, while the MB pathway has an extra layer (implementing the task structure representation) re-initialised at each trial. This models the view that the MF system performs simpler and faster computations, while the MB system performs more time-consuming computations (Keramati et al., 2011a).

The operation of the model proceeds as follows. At the beginning of a trial, a constant activation is set in the MF-MB layer, shared across the two units according to a ratio determined by the ω parameter (0 corresponding to a pure MF pathway, and 1 to a pure MB pathway). The state units are initialised with zero activation. The output units are initialised with a proportion of activation remaining from the preceding trial, determined by the carry-over parameter (set to 0.1). This is equivalent to p in the RL model. Activity is iteratively propagated through all units via the IAC equations of Gilbert and Shallice (2002). At each processing cycle t, the net input $net(t)$ to each unit is the sum of the activations sent by connected units multiplied by the corresponding weight, plus the unit bias (set to -1.5). The unit activity at time $t+1$, $act(t+1)$, is then calculated as:

$$act(t+1) = \begin{cases} act(t) + s \cdot net(t) \cdot (max - act(t)) + \eta \\ \hfill \text{if } net(t) > 0 \\ act(t) + s \cdot net(t) \cdot (act(t) - min) + \eta \\ \hfill \text{if } net(t) < 0 \end{cases}$$

where $min = -1$ is the minimum activation, $max = 1$ the maximum activation, $s = 0.0015$ the step size (determining the processing speed) and η is a noise term, drawn from a Gaussian with standard deviation $\sigma = 0.01$.

The activations from each pathway are combined at the level of the output units. A trial terminates when one unit is inhibited below the response threshold (-0.9), and the action taken is the one corresponding to the unit with the highest activation. After taking an action, a state was set according to probabilistic transitions (taking action A led to state A with 0.7 probability and to state B with 0.3 probability, whereas taking action B led to state B with 0.7 probability and to state A with 0.3 probability). Each combination of state and action had a probability of reward. Like in the original study, the probability of reward drifted between 0.25 and 0.75 by addition of independent Gaussian noise, centred and with standard deviation 0.025. The trial was rewarded according to the action taken and state reached. The role of

---

[1] These are kept fixed, on the assumption that participants know the transitions from the start. A similar assumption is made in Decker et al. (2016).

the drifting probability was to encourage participants to keep learning throughout the task and thus keep a constant learning rate, which would otherwise generate parametric non-stationarity issues (cf. Daw, 2011).

Given our focus on stage one, we simplified the second part of a trial from the original task (presentation of the second state followed by the second action and the outcome) by having the first action leading to the second-stage state and the outcome simultaneously.

As in the RL conceptualisation, the two pathways independently learn from feedback. This is implemented by changing the weights within the MF and MB pathways. On receipt of reward, the weights between the MF unit and the selected output $O_i \in \{A, B\}$ and between the MB unit and the reached state $S_i \in \{s_A, s_B\}$ are updated as follows:

$$w_{MF-O_i}(t+1) = w_{MF-O_i}(t) + \alpha \cdot (r - w_{MF-O_i}(t))$$

$$w_{MB-S_i}(t+1) = w_{MB-S_i}(t) + \alpha \cdot (r - w_{MB-S_i}(t))$$

where t is the trial index, $r \in \{0, 1\}$ is the observed reward, and $\alpha = 0.6$ the learning rate.

The update for the MF weights resembles the Temporal Difference (TD) learning implemented in the MF system in Decker et al. (2016). We model the MB system as learning what state is rewarded rather than what action is rewarded, hence updating the weights according to both the received reward and the state reached, regardless of which state unit had stronger activation. In the processing, the weights influence how much each state is sought.

The ω parameter reflects the overall tendency of an agent to favour one strategy or another at the beginning of processing. The σ parameter in the output units acts as decision noise. In contrast to Decker et al. (2016), the noise in the action selection units (taken to integrate the degree of exploration) is separated from the contribution of the MF and MB systems. For example, an agent with $\omega = 0$ and $\sigma = 0$ would have the output units' activations driven purely by the MF connection weight values and systematically take the action fed by the strongest connection. Increasing σ increases the tendency to explore.

The above-describe model constitutes the baseline that will be referred to later as the *recruitment* variant. In Section 3.3.2.3, we present variants of this model to test several developmental hypotheses.

3.3.1.2 Model stay patterns evaluation

We evaluate the novel model by examining the stay responses' pattern to quantify the model's ability to generate MF or MB prototypical behaviours, in line with Decker et al.'s (2016) analytical approach. We fit simple logistic regression models to our simulated data with *stay* (repeat preceding trial's action or not) as the dependent variable, and as explanatory variables: *reward* (rewarded or unrewarded) from the preceding trial, *transition* (common or rare) from the preceding trial, and their interaction. Thus, the MF signature is present if the coefficient estimate for reward is significantly different from 0 (indicating a reinforcement effect) but not the other terms. The MB signature is present if the reinforcement effect is combined with a significant interaction between reward and transition. We will report the main effect of reward first and the interaction first, since they are the key predictors differentiating between MF and MB signature, followed by the main effect of transition, which is predicted to be non significant in both cases. The regressions were performed with the *glm* function of Python's package *statslearning* (Seabold and Perktold, 2010).

3.3.1.3 Response time simulations

As mentioned, we chose an IAC model for its ability to generate response time directly from its processing (it does not need additional steps to link the model to response time data such as *Drift diffusion modelling*; used for example in Seabold and Perktold (2010)). Response time is expected to hold crucial information about the MF-MB interaction given the characteristics of each system (Daw et al., 2005, Keramati et al., 2011a). Indeed, the MF is computationally simple so would by itself lead to very fast responses, while it takes more time for the MB system to simulate or search within its representation of the world. Furthermore, Dezfouli et al. (2014) suggested that the type of recruitment of each system directly depends on the time pressure assigned to a decision.

To generate response time, we first took the number of cycles until reaching response threshold at each trial. We then converted the number of cycles into RTs (in seconds) by regressing individual simulations against individual child distributions. Specifically, we assumed that for each distribution, the number of cycles and the real data were linked by a linear regression. We thus estimated the regression coefficients and transformed the number of cycles into RTs with these coefficients.

We finally fit exGaussians to the RT histograms of individual participants or the RT histograms simulated by the model variants. An exGaussian (exponentially modified Gaussian distribution) is the sum of independent normal and exponential random variables. It is well-suited to account for long tails in RT distribution because of the exponential component (Heathcote et al., 1991). It has the following probability density function:

$$f(x;\mu,\sigma,\lambda) = \frac{\lambda}{2} e^{\frac{\lambda}{2}(2\mu+\lambda\sigma^2-2x)} erfc(\frac{\mu+\lambda\sigma^2-x}{\sqrt{2}\sigma})$$

where *erfc* is the complementary error function: $erfc(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

The fit was done with the function *annealing* from the python module *curve-fit*[2] with the following bounds $\mu = [-1, 1]$, $\sigma = [0.01, 1]$, $\tau = [0.01, 1]$. We will use the three parameters (μ, σ, τ) of these RT-fitted exGaussians to compare the different

---

[2] https://pypi.org/project/curve-fit.annealing/

simulations to the data (to assess fit to data) and also compare them between each other (to assess the discriminative power of the simulations).

When looking at Figure 3.21, it seemed that none of the models generates data sufficiently close to the empirical data. Indeed, the simulated RTs looks more alike to each other than to the empirical data. Beyond the absolute fit to data, it was interesting to *(1)* assess whether the different model variants actually generate distinguishable RTs, that had the potential to tease the different hypotheses apart, and *(2)* see which model variant accounts better for the child data than the others.

To test whether the simulated RT data from the different models are discriminable on the basis of their distributions *(1)*, we tested the differences in estimated parameters using independent samples t-tests comparing the *recruitment* model with the other model variants for each parameter (μ, σ and τ).

To find out which model is the closest to the children's RT data distribution *(2)*, we calculated the Euclidian distance between the exGaussian parameters from the model variants to the exGaussian parameters of the child data in the average-parameter space.

3.3.2 Results

3.3.2.1 Functioning of the model

Examples of activations evolving in a trial are shown in Figure 3.14. This illustrates the mechanisms in case of, and in the absence of, conflict. The top panels represent what we can define as *conflicting* cases, where the option favoured (i.e., with the strongest weight) by MF-pathway is opposite to the option favoured by MB-pathway (approximating the weight from MB-unit to state A as a preference towards output A). In the bottom panels, both pathways favour the same output. Consequently, the selection threshold is reached more rapidly in the latter, non-conflicting case, than in the conflicting case. In the bottom examples, the contribution of the MB-pathway arrives late, due to the time of building up in the state layer. Consequently, the output unit crossing the threshold is the one favoured by the MF-pathway.

Figure 3.17 presents the simulations of the number of cycles (acting as a proxy for RTs). The age categories were approximated by different ω parameters, by matching the MB scores (defined in Section 3.2.1.1) of participants to simulated choice patterns. For simulations, the RTs increase with MB contribution, which is not what is observed in participants (see Figure 3.21). This is because our model currently does not account for the effect of age on motor responses, which is likely the main reason why children are slower than adults. Although the fit of our RT simulations to the data is poor, the primary goal is to show how such process model can (1) generate response time data, (2) later discriminate between different mechanistic hypotheses and (3) get a qualitative comparison to suggest the most suited hypothesis among them.

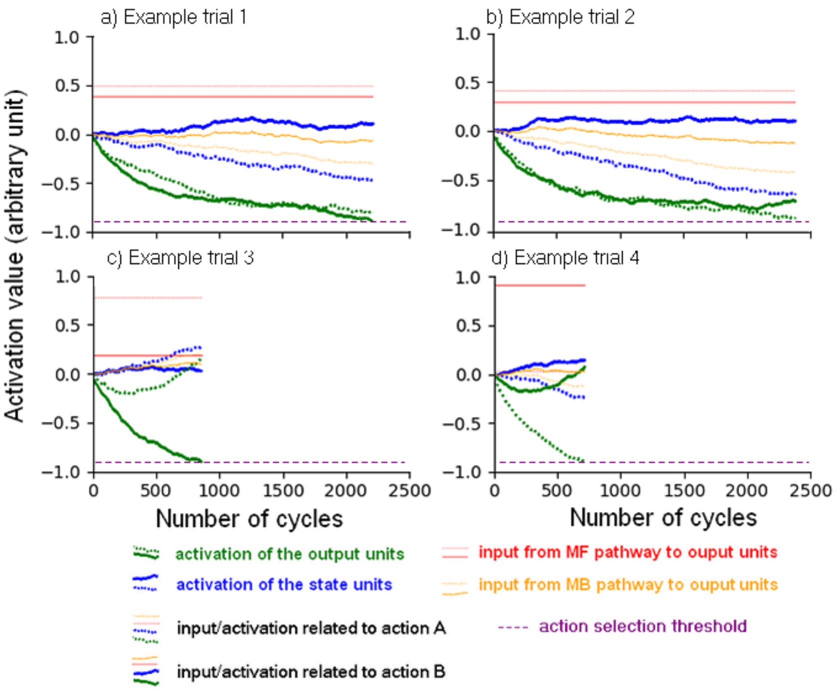


Fig. 3.14: **Activations time course in four trials with ω=0.5.**
The trials were taken from various parts of the task. Dotted lines represent an output unit or the input to this same unit; plain lines represent the alternatives. The constant inputs from the MF and MB pathways contribute to driving the evolution of activations in the state and output units. A trial stops when the activation of one of the two output units (in green) crosses the threshold.

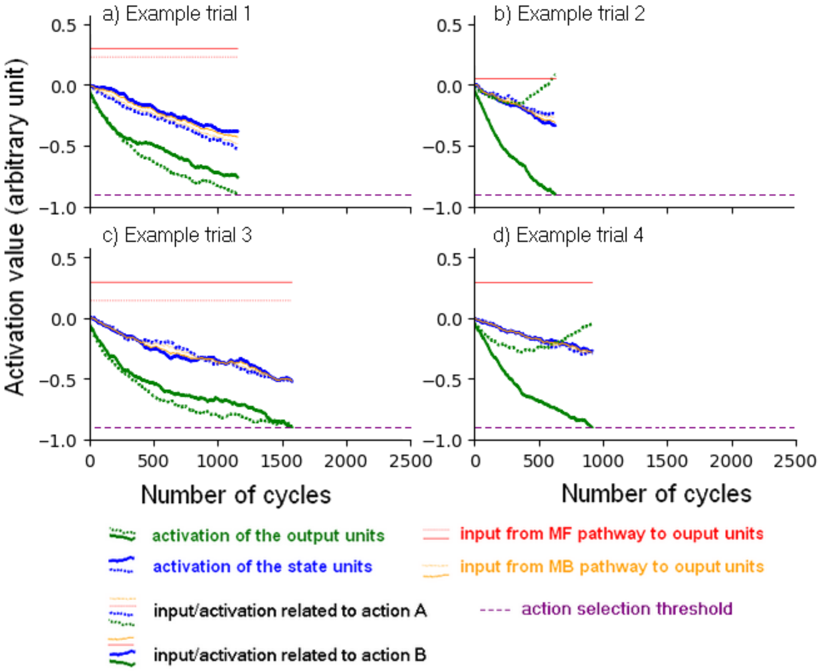


Fig. 3.15: **Activations time course in four trials with ω=0 (i.e., pure MF).**
The trials were taken from various parts of the task. Dotted lines represent an output unit or the input to this same unit; plain lines represent the alternatives. The constant inputs from the MF and MB pathways contribute to driving the evolution of activations in the state and output units. A trial stops when the activation of one of the two output units (in green) crosses the threshold.

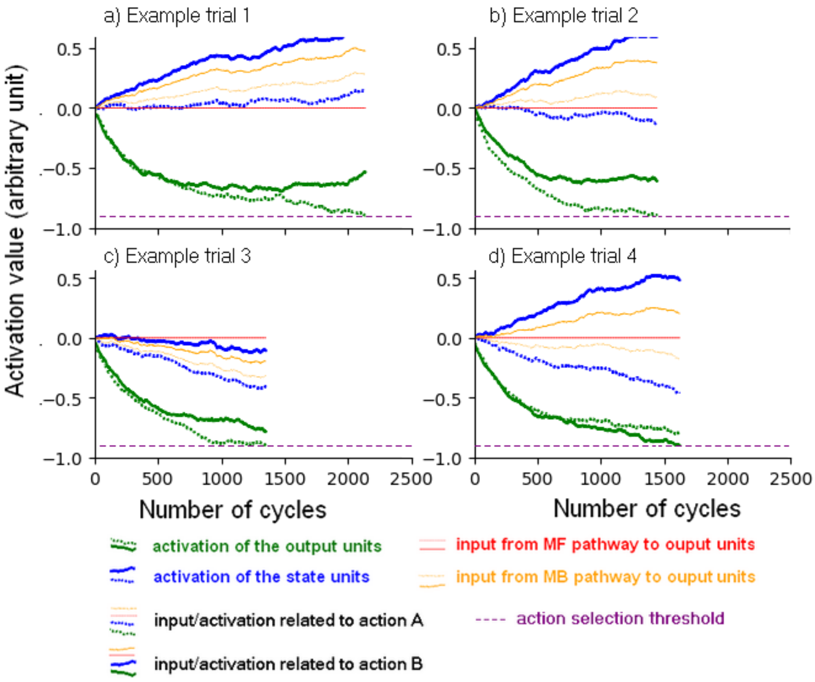


***Fig. 3.16:* Activations time course in four trials with ω=1 (i.e., pure MB).** **The trials were taken from various parts of the task. Dotted lines represent an output unit or the input to this same unit; plain lines represent the alternatives. The constant inputs from the MF and MB pathways contribute to driving the evolution of activations in the state and output units. A trial stops when the activation of one of the two output units (in green) crosses the threshold.**

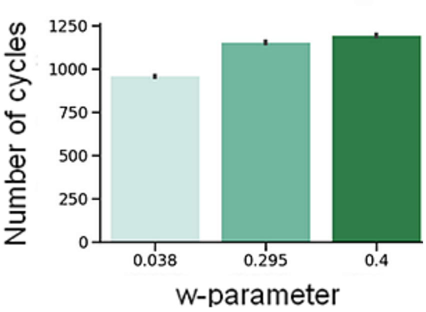


***Fig. 3.17:* Simulated RTs (number of cycles) for different ω reflecting the age groups, each simulated 20 times. Error bars represent 95% confidence intervals.**

3.3.2.2 Standard behaviour of the model

We first explore the outputs of the model and its fit to the Decker et al. (2016) data. The model fit was evaluated by its ability to reproduce the MF signature (significant reward effect and no other effect) or the MB signature (significant reward effect and significant interaction) accordingly (cf. Section 3.3.1.2).

We started by varying the relative strength of the systems. In the simulations of the stay patterns shown in Figure 3.18, we reproduced the theoretical behaviour of an agent under pure MF control. For this and the following simulations, we report the following statistics: β (current predictor's regression estimate), SE (standard error), p (p-values) and 95% CI (95% confidence interval on the regression estimate). For $\omega = 0$, we obtained a main effect of reward: β = 0.84, SE = 0.14, $p < 0.001$, 95% CI [0.55, 1.12]; but no interaction: β = -0.1, SE = 0.1, $p = 0.326$, 95% CI [-0.31, 0.1]; and no main effect of transition: β = 0.26, SE = 0.17, $p = 0.141$, 95% CI [-0.08, 0.6]. We also reproduced the behaviour of an agent under pure MB control (for $\omega = 1$, main effect of reward: β = 0.37, SE = 0.15, $p < 0.05$, 95% CI [0.08, 0.66]; and interaction: β = 0.39, SE = 0.18, $p < 0.05$, 95% CI [0.04, 0.73]; but a main effect of transition: β = -0.15, SE = 0.12, $p = 0.209$, 95% CI [-0.4, 0.09]), similar to Figure 1 in Decker et al. (2016). Gradually increasing ω led to a graded increase towards the prototypical MB pattern.

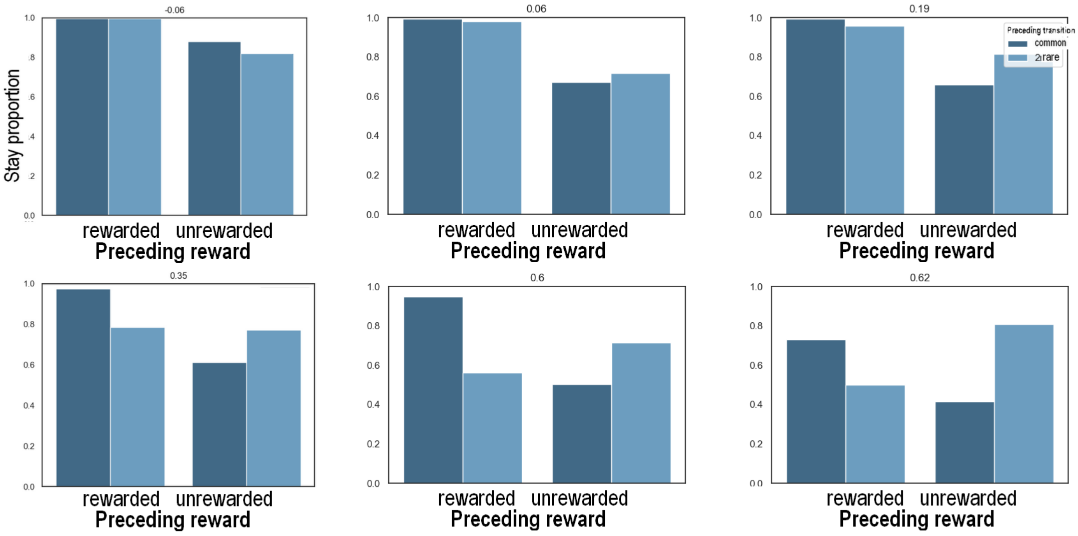


Fig. 3.18: **Model simulations of the stay pattern (proportion of stay as a function of the preceding trial's reward and transition type) for values of ω between 0 and 1 increasing by steps of 0.2. Each plot is based on 200 trials averaged over 20 runs (reflecting the number of participants per age group in the original study).**

We subsequently tested the model fit to Decker et al.'s (2016) data. That is, we ran the model by setting the parameter ω to reproduce the relative reliance of model-free and model-based that characterises the three age groups: children rely on model-free ($\omega = 0$), adolescents rely on a mixture of model-free and model-based ($\omega = 0.5$) and adults predominantly rely on model-based ($\omega = 0.7$). Figure 3.19 presents the stay patterns of three sets of simulations which fit the children's, adolescents' and adults' behaviours (with $\omega = 0$, 0.5 and 0.7 respectively). The MF signature is present for $\omega = 0$ (main effect of reward: β = 0.84, SE = 0.14, $p < 0.001$, 95% CI [0.55, 1.12]; but no interaction: β = -0.1, SE = 0.1, $p = 0.326$, 95% CI [-0.31, 0.1]; and no main effect of transition: β = 0.26, SE = 0.17, $p = 0.141$, 95% CI [-0.08, 0.6]), while the MB signature is present for the other two ω values (for $\omega = 0.5$: main effect of reward: β = 1.07, SE = 0.14, $p < 0.001$, 95% CI [0.8, 1.35]; interaction: β = -0.23, SE = 0.11, $p < 0.05$, 95% CI [-0.45, -0.01]; and no main effect of transition: β = 0.31, SE = 0.17, $p = 0.062$, 95% CI [-0.02, 0.65]); for $\omega = 0.7$: main effect of reward: β = 0.41, SE = 0.15, $p < 0.01$, 95% CI [0.12, 0.7]; interaction: β = 0.47, SE = 0.18, $p < 0.01$, 95% CI [0.12, 0.82]; and no main effect of transition: β = -0.18, SE = 0.13, $p = 0.175$, 95% CI [-0.43, 0.08]). .

As mentioned, the same MF and MB signature were found on Decker et al.'s (2016) data: for the children's group, the MF signature only is present (main effect of reward: β = 0.30, SE = 0.08, $p < 0.001$ but no interaction: β = 0.02, SE = 0.04, $p = 0.65$; and no main effect of transition: β = 0.02, SE = 0.04, $p = 0.79$). However for the adolescents' group, the MB signature is present (main effect of reward: β = 0.22, SE = 0.08, $p < 0.01$ and an interaction: β = 0.35, SE = 0.10, $p < 0.01$; and no main effect of transition: β = 0.09, SE = 0.06, $p = 0.13$). For the adults' group, the MB signature is also present (main effect of reward: β = 0.56, SE = 0.11, $p < 0.001$ and an interaction: β = 0.49, SE = 0.13, $p < 0.001$; and a main effect of transition: β = 0.07, SE = 0.08, $p < 0.05$).

In other words, the model reproduces the key stay patterns observed with increasing age by increasing values of ω. We have thus shown that is it possible to provide a mechanistic account of the dynamic integration between systems and how this balance changes with age.

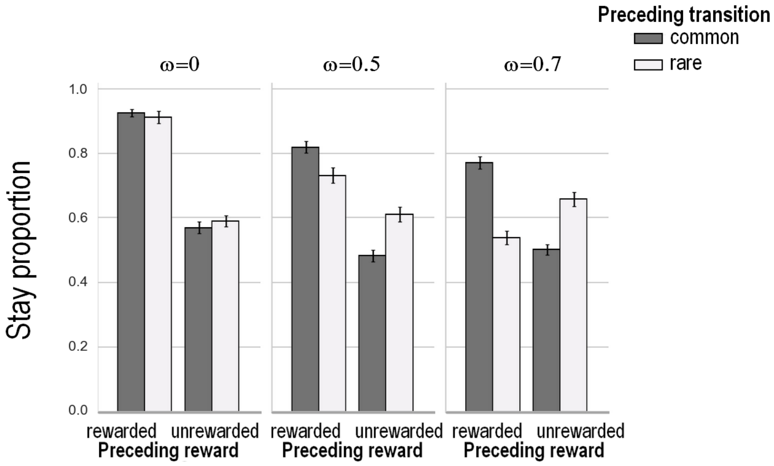


Fig. 3.19: **Model simulations of the stay pattern (proportion of stay as a function of the preceding trial's reward and transition type) for values of ω fitting Decker et al.'s (2016) three age groups (*Recruitment* hypothesis). Here and in all subsequent figures each plot is based on 200 trials averaged over 20 runs (reflecting the number of participants per age group in the original study), and error bars represent the standard error of the mean.**

3.3.2.3 Developmental changes: competing hypotheses

In this Section, we explore additional mechanistic hypotheses of the age effect on behaviour by leveraging specificities of the model. Clues in Decker et al.'s (2016) analyses rule out the possibility that the youngest group did not understand the task transition structure. Indeed, all groups were significantly slower following rare than common transitions (with no age-transition interaction), and, when prompted, all groups expressed equal awareness of the task structure. Thus, altering the 0.7 and 0.3 weights of the model is a very unlikely explanation of the changes that occur across development compared to the most plausible explanations that follows.

We now turn to the meaning of the ω parameter. Adults' behaviour in similar tasks fit by a hybrid RL model suggests a form of metacontrol of the MF/MB balance, e.g., via uncertainty-based competition (Daw et al., 2005) or a cost-benefit assessment (Kool et al., 2017b). In this vein, the change in MF/MB trade-off with age may derive from metacontrol mechanisms that vary (through different assessment and/or bias towards one strategy or the other) with age. We will refer to this as the 'recruitment strategy' hypothesis.

However, there are alternative explanations for the absence of the MB pattern in the youngest group, motivated by the MB system being arguably less mature in childhood. Indeed, the MB system has been proposed to involve prefrontal cortical areas based on both neuroimaging (e.g., Jueptner et al., 1997) and neuropsychological (e.g, Duncan, 1986) evidence, and the prefrontal cortex shows protracted maturation until early adulthood (Lenroot and Giedd, 2006). Our mechanistic approach allows exploration of different hypotheses concerning the fine-grained nature of maturation of the MB system, and we consider three such hypotheses.

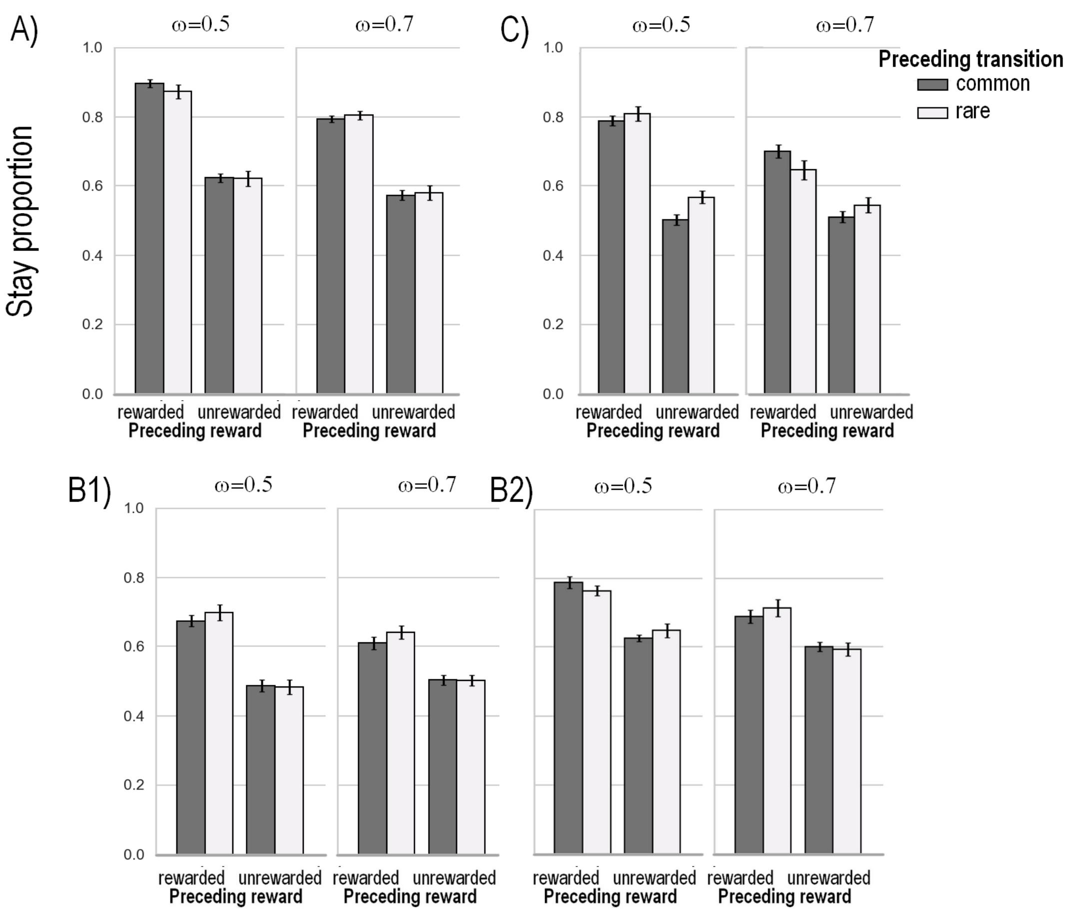


***Fig. 3.20:* Model simulations of the stay pattern for the different manipulations, each set repeated for two different ω values. A) *Noise*: increased noise in the state layer ($\sigma = 0.2$); B1) decreased bias in the state layer (-4.5); B2) *Bias* decreased bias in the state layer (-4.5) and increased carry-over (0.4); C) *LR*: altered learning rates: 0.8 in the MF-pathway, 0.2 in the MB-pathway. Each plot is based on 200 trials averaged over 20 runs, error bars represent one standard error of the mean.**

First, delayed maturation could be linked to noisier computations in younger participants. Implementing higher σ in the state layer may, overall, reduce the MB contribution at the expense of the MF contribution to choice in children. Indeed, strong noise may partly mask any existing difference in the connection weights in the MB pathway, hindering the divergence in activation levels of the state units (where the activation of one unit would normally take over). This would lead to smaller variation between the state's input to one compared with the input to the other output unit. Ultimately, this would limit the MB pathway's influence on the output, relative to the MF pathway (because the MF pathway is direct and rapid).

Secondly, and not mutually exclusive with the first hypothesis, MB processing in adults may benefit from extensive experience (such benefit arguably having reached ceiling in MF system). Under this scenario, children's computations would therefore start with a disadvantage in MB pathway at each trial, implementable via a high magnitude of negative bias in the state layer. Such altered biases would delay computations in the MB compared to the MF pathway. Such a delay may prevent the state units' activations from diverging sufficiently from each other. If the delay is too big, this may prevent the stats units from influencing the output units before the end (i.e. before one output unit reaches the threshold for action selection).

A third hypothesis is that children's MB system learns less efficiently, i.e. that the system has a lower learning rate. Indeed, there is empirical evidence, although conflicting, that learning rates change with age (see the review of Nussenbaum and Hartley, 2019).

3.3.2.4 Developmental changes: simulations

The rationale of the following simulations is that the children stay patterns, initially assumed to originate from the recruitment of the model-free system only (that is, with $\omega = 0$ as on the left panel of Figure 3.19), can originate from alternative mechanisms. Thus, we implemented the alternative mechanisms (variants) by keeping the same parameters as for the baseline (*recruitment*) simulation, except for the parameters linked to the alternative mechanism, and with both $\omega = 0.5$ and 0.7 which simulates adolescents' and adults' data with the recruitment hypothesis. We thus compare the results of the *variants* to the stay patterns obtained with $\omega = 0$ (mimicking children's behaviour; left panel of Figure 3.19), referred to as the pure MF behaviour.

For the first hypothesis, we increased the noise to $\sigma = 0.02$ in the state units (Figure 3.20.A). For $\omega = 0.5$, the stay pattern closely resembles pure MF behaviour. The regression analysis confirms the presence of the MF signature (main effect of reward: β = 1.16, SE = 0.15, $p < 0.001$, 95% CI [0.87, 1.45]; ; but no interaction: β = 0.27, SE = 0.18, $p = 0.133$, 95% CI [-0.08, 0.62]; and no main effect of transition: β = -0.02, SE = 0.11, $p = 0.871$, 95% CI [-0.23, 0.19]). Similar comments apply for $\omega = 0.7$ (main effect of reward: β = 0.49, SE = 0.13, $p < 0.001$, 95% CI [0.23, 0.75]; but no interaction: β = 0.09, SE = 0.16, $p = 0.56$, 95% CI [-0.22, 0.41]; and no main effect of transition: β = 0.01, SE = 0.1, $p = 0.957$, 95% CI [-0.2, 0.21]).

Second, we decreased the bias in the state layer to -4.5 (Figure 3.20.B1). The MF signature is again present (for $\omega = 0.5$, main effect of reward: β = 0.96, SE = 0.14, $p < 0.001$, 95% CI [0.69, 1.22]; but no interaction: β = -0.15, SE = 0.16, $p = 0.367$, 95% CI [-0.46, 0.17]; and no main effect of transition: β = 0.16, SE = 0.11, $p = 0.145$, 95% CI [-0.05, 0.37]; for $\omega = 0.7$, main effect of reward: β = 0.52, SE = 0.13, $p < 0.001$, 95% CI [0.27, 0.78]; but no interaction: β = -0.05, SE = 0.16, $p = 0.752$, 95% CI [-0.36, 0.26]; and no main effect of transition: β = 0.02, SE = 0.11, $p = 0.855$, 95% CI [-0.19, 0.23]). However, the stay proportions are relatively low (cf. left panel of Figure 3.19). We can raise the model's perseveration, a behaviour common in children, with *carry-over* $= 0.4$ (Figure 3.20.B2). The results then match the pure MF plot, with the MF signature (for $\omega = 0.5$, main effect of reward: β = 0.53, SE = 0.14, $p < 0.001$, 95% CI [0.26, 0.8]; but no interaction: β = 0.07, SE = 0.17, $p = 0.693$, 95%

***Table 3.1:* Significance levels of the logistic regression predictors of interest (the table presents the full p-value or: $* : p < 0.05$, $** : p < 0.01$, $*** : p < 0.001$). The model variants represent the different developmental hypotheses modelled.**

Model variant	Recruitment			Bias		Noise		Lr	
ω	0	0.5	0.7	0.5	0.7	0.5	0.7	0.5	0.7
Reward	$***$	$***$	$**$	$***$	$**$	$***$	$***$	$*$	$**$
Reward*Transition	0.326	$*$	$**$	0.693	0.59	0.133	0.56	0.287	0.334

CI [-0.26, 0.4]; and no main effect of transition: β = -0.05, SE = 0.11, $p = 0.646$, 95% CI [-0.26, 0.16]; for $\omega = 0.7$, main effect of reward: β = 0.44, SE = 0.14, $p < 0.01$, 95% CI [0.17, 0.7]; but no interaction: β = 0.09, SE = 0.16, $p = 0.59$, 95% CI [-0.23, 0.41]; and no main effect of transition: β = -0.13, SE = 0.12, $p = 0.27$, 95% CI [-0.36, 0.1]).

Finally, we changed the learning rates to $\alpha_{MB} = 0.2$ and $\alpha_{MF} = 0.8$ (Figure 3.20.C). For $\omega = 0.5$, the pattern presents no prototypical signature. We find a main effect of reward (β = 0.32, SE = 0.14, $p < 0.05$, 95% CI [0.06, 0.59]), no interaction (β = -0.11, SE = 0.1, $p = 0.287$, 95% CI [-0.31, 0.09]), but also a main effect of transition (β = 0.32, SE = 0.16, $p < 0.05$, 95% CI [0.0, 0.64]). The same goes for $\omega = 0.7$: main effect of reward (main effect of reward: β = 0.39, SE = 0.13, $p < 0.01$, 95% CI [0.13, 0.66]), no interaction (β = -0.1, SE = 0.1, $p = 0.334$, 95% CI [-0.3, 0.1]), but also a main effect of transition (β = 0.42, SE = 0.16, $p < 0.01$, 95% CI [0.1, 0.74])

Table 3.1 presents a summary of the findings. All in all, two mechanisms produced stay patterns indistinguishable from the 'recruitment strategy' with $\omega = 0$: noisier computations in the MB pathway (*noise* variant), and less efficient processing in the state layer with higher perseveration (*bias* variant).

3.3.2.5 Response time simulations

For the response time simulations, we retain the variants that produced the patterns of interest as per the previous Section. That is, we kept the four developmental hypotheses implemented as follows:

- *Recruitment:* the relative recruitment of the contributions of MF/MB systems for action selection is externally arbitrated and increasingly in favour of MB with age: ω increases with age.
- *Noise:* the MB system is recruited to the same level but computations are nosier in children than adults: ω held fixed but noise in the state units increased from $\theta = 0.01$ in the *recruitment* variant to $\theta = 0.02$.
- *Bias:* the children MB system is less experienced than adults and slower to initiate computations, and have higher perseveration: ω held fixed but amplified negative bias in the state layer from -1.5 to -4.5, plus the model's perseveration is raised with *carry-over* from 0.1 in the *recruitment* variant to 0.4.
- *LR:* children MB learning is slower (but the MF learning, which is simpler, is at the adult level): ω held fixed but learning rates decreased from 0.6 in both channels in the *recruitment* variant, to $\alpha_{MB} = 0.2$ (very slow) and $\alpha_{MF} = 0.8$ (slightly faster).

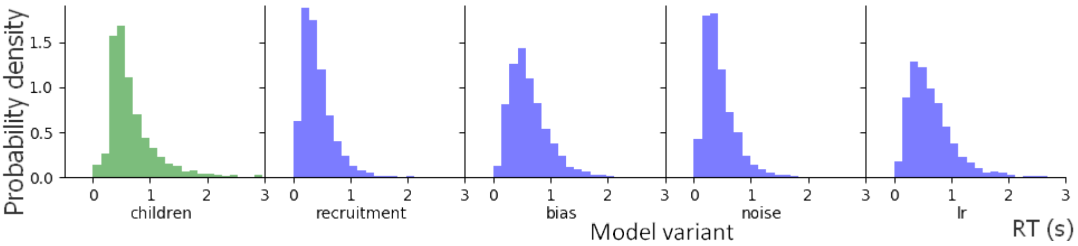


***Fig. 3.21:* Response time of the children and four simulations. *Children* refers to the original dataset. The names refer to 4 variants representing the developmental hypotheses modelled and were generated with the following changes in parameters: in *Recruitment*, the relative contributions of MF/MB system ω was modulated (increasing ω represents increased contribution of the MB system); in *Noise*, ω was held fixed but noise θ in the state units was increased from 0.01 to 0.02; in *Bias*, ω was held fixed but the negative bias in the state layer from -1.5 to -4.5 was amplified, and the model's perseveration (consistent with children's behaviour in general) *carry-over* was raised from 0.1 to 0.4; and in *lr*, ω was held fixed but the learning rates was decreased from 0.6 (for both channels) to $\alpha_{MB} = 0.2$ and $\alpha_{MF} = 0.8$.**

Figure 3.21 shows response time distributions for the four model variants' data and the real children's data. The corresponding exGaussians, reconstructed with the parameters fit to the simulated or real data, are shown in Figure 3.22A. We plot the fitted parameters of the four model variants in Figure 3.22B. When comparing the variants against the *recruitment* model, there were significant differences for two parameters of the *bias* model (μ and σ both $p = 0.002$). Post-hoc pairwise comparisons between the other models also revealed significant differences, namely between *bias* and *noise* variants, whose μ and σ differ (both $p = 0.002$). The other comparisons were not different from each other.

So far, the results suggest that at least a few hypotheses can be distinguished from each other on the basis of their relative fit to the RT distribution. This ability of the RT data to discriminate from the model variants at this stage strongly suggests that the RT contains informative characteristics that can be leveraged via our model to support the most plausible developmental hypothesis.

Figure 3.22C shows the fitted exGaussian parameters in the 3D parameter space. The 3D Euclidian distance d_{3D} of the variants to the children's data were: 0.124 for *recruitment*, 0.132 for *bias*, 0.118 for *noise* and 0.129 for *lr*. The lowest value of the 3D distance of *noise* variant from the child data ($d_{3D} = 0.118$), combined with the fact that parameters for *noise* were significantly different from the ones for *bias*, indicates that *noise* hypothesis is more plausible than the *bias* hypothesis. However the parameters for *noise* and *recruitment* were not significantly different from each other, and the *recruitment* variant had the second smallest distance to the child data ($d_{3D} = 0.124$). Therefore, at this stage, both hypotheses (*noise* and *recruitment*) are equally likely, but both are better than the *lr* or *bias* hypotheses.

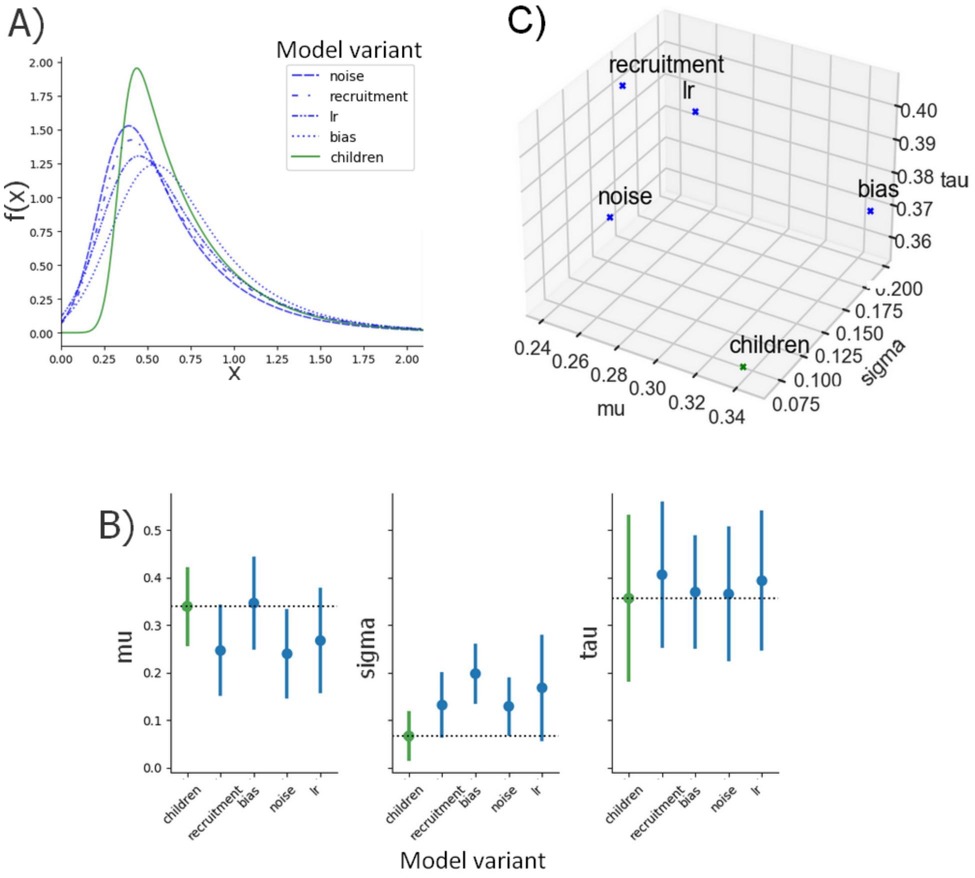


Fig. 3.22: **A comparison of child and simulated RT distributions: (a) exGaussian curves generated with the average estimated parameters of the exGaussian distribution (mu, tau and sigma), (b) relationship among distributions in the exGaussian's parameters space, (c) estimated parameters (the error bars are the 95% C.I. on individual participants or simulations, the dotted line represent the children's values; lr=learning rate).**

3.4 Discussion

3.4.1 Our model

We have proposed a mechanism for the integration of the MF and MB systems in the widely used two-stage task, and shown different ways in which the MF/MB trade-off may emerge across age.

The model offers a level of analysis for understanding action selection different from Decker et al.'s (2016) study. Indeed, important problems tackled by the RL approaches, such as the exploration-exploitation dilemma, are fundamentally at the computational level in Marr's terms (Marr, 1982). For example, many reinforcement learning tasks have found that children tend to explore more than adults (e.g., Sumner et al., 2019). Others have looked at the relationship of the MF/MB balance with stress and working memory (Otto et al., 2013), or the balance of exploitation and exploration as a function of environment volatility (Speekenbrink, 2015). Such RL approaches makes contact to some extent with the algorithmic level. Indeed, they describe the inputs used for computation (outcomes from past choices in given contexts, etc.) and the operations performed: essentially learning (e.g., Q-learning) and the sampling strategy (e.g., softmax choice rule). This approach informs the extent to which variables of interest (e.g., age) are linked to different RL parameters (e.g., learning rate), which can provide insight into what operations of RL the variables are modulating, but not precisely into the processes by which the variables modulate behaviour. Our approach centres on the algorithmic level. We propose a way in which several RL components (Q-values and probability distributions) are represented (discrete units with dynamic activation) and connected, as well as how the within-trial dynamic computations are carried out by each system and subsequently integrated.

The model accounts for the key behavioural patterns attributed to varying MF/MB balance. One way to generate such stay patterns is to vary the relative strength of the input sent to one or the other pathway (via the ω parameter), reproducing Decker et al.'s (2016) developmental data by increasing the ratio in favour of the MB pathway with age. In theoretical terms, this is consistent with Decker et al.'s (2016) account, where the behaviour of children is captured by lower MB inverse temperature than adults' (and not by the MF temperature parameter), although our model alters the trade-off between MF and MB. However, when exploring mechanistic accounts of the age differences, we found that the model was also able to account for the younger group's behavioural patterns through other theoretically-motivated manipulations. Namely, with values of ω initially associated with adolescents and adults, altering key processes of the MB pathway (with higher computation noise, or with lower bias combined to higher carry-over) yielded patterns hardly distinguishable from children's pure MF behaviour. Hence, the stay by reward and transition pattern does not allow us to discriminate between the *recruitment strategy* hypothesis and one of the three alternative mechanisms (related to increased noise).

When using response time, although the fit to the children's data may not be sufficient, it does show the ability of the model to simulate both behavioural and response time patterns consistent with the data, and to discriminate between some model variants. Further refinement in the model, such as allowing the temperature β or the learning rate α to vary for a given individual within the task, age may improve both aspects. Relatedly, our simulations of response time have focused on simulating different mechanisms for the same age group (children). To use response time simulations for different age groups, the increasing motor speed (and general processing speed) with age must be integrated into the model.

Another interest of such models, if developed further, would be to inform future task design. Indeed, the model makes sufficiently distinct predictions for the different developmental hypotheses to disentangle them. For example, one could think of a study with a similar task but two conditions. One condition would include stimuli that adults are familiar with but the children are not (*unfamiliar condition*). Another condition would include stimuli where children are as familiar with the stimuli as adults (*familiar condition*). In such a study, the *bias* (efficiency) hypothesis would predict that children display less MB contribution than adults in the *unfamiliar condition* but not in the *familiar condition*. Indeed, familiarity would boost the bias in state units. On the contrary, the *noise* hypothesis would predict no difference between the *familiar* and *unfamiliar* conditions, i.e. children would have less MB contribution than adults in both conditions, because noise is independent of the knowledge of or familiarity with stimulus. Other interesting paradigms may manipulate time pressure (e.g., by giving a short amount of time to reply). Accordingly, with a shorter time to respond, the *Bias* hypothesis would predict that children would increase reliance in MF but that adults would not or would to a lesser extent. On the contrary, the *(Noise)* hypothesis would predict similar effects of time pressure on adults and children. Finally, the *recruitment* hypothesis may be tested by using different tasks that measure the relative recruitment of model-free and model-based in different ways. If such a recruitment is the same in different tasks, this would be in favour of the *recruitment* hypothesis. Moreover, the different hypotheses may not be mutually exclusive.

An important limitation of the present study is that the model space (Palminteri, Wyart, et al., 2017), or the number of developmental hypotheses explored, may not represent the entirety of plausible hypotheses. Therefore, it is a first step towards distinguishing candidate hypotheses, but there may be other candidate mechanisms that have not been included in this study and should not be discarded.

3.4.2 Response time in classical RL fit

Previous studies have included response time in modelling of RL (e.g., Ballard and McClure, 2019, Gershman, 2018, and Fontanesi et al., 2019). Shahar et al., 2019 used response time specifically in the two-stage task, but were doing so for the purpose of better parameter estimation within the same RL variant. By contrast, our usage of response time was not concerned with parameter estimation but was aimed at separating competing model variants.

Our model has the advantage of having response times generated straightforwardly from the processing of the model (only converting the number of cycles to response times), without any extra step such as Drift Diffusion Modelling (e.g., Fontanesi et al., 2019). Nonetheless, regardless of the aim, all approaches mentioned here benefit from using response times as additional data. In general, more data provide additional constraints, and can be any form, such as kinematics, eye-tracking, or neuroimaging (e.g., Borst et al., 2015).

3.4.3 Model-free/model-based balance and Executive functions

Given the relevance of executive function in the rest of this thesis, it is worth noting studies that have linked the model-free/model-base balance to executive function tasks. Otto et al. (2014) have compared the inter-individual differences of adults' performance on the two-stage task and the same adults' performance on the Stroop task (Cohen et al., 1999). They found that the susceptibility to the Stroop interference was negatively correlated with model-based use in the two-stage task. Although not on a developmental sample, the finding is compatible with the idea that development in inhibitory control may be a key driver of improvement in action selection. Nonetheless, Otto et al. (2014) further propose that cognitive control may favour model-based relative to model-free recruitment either by boosting the relevant task representations or by inhibiting model-free responses. This view puts the role of inhibitory control as inhibiting model-free representations, that is, state-actions associations learned by simply caching previous rewards obtained by trial-and-error. The view in Chapter 2 will be slightly different. There, inhibitory control is applied on irrelevant distractors, that is on lower-level object representations rather than on state-actions associations. The view is also slightly different from that of the model of Cooper and Shallice (2000), where cognitive control or supervisory control influences the appropriate action schemas rather than inhibits response favoured by a different system.

Potter et al. (2017) had 9- to 12 year-olds perform the two-stage task as well as three separate tasks to measure working memory, statistical learning and fluid reasoning. Among the three, they found that only fluid reasoning predicted higher-recruitment of model-based strategy. They propose that the ability to integrate learned associations is central to the developmental differences found in model-free-/model-based balance. It is indeed likely that several components (in addition to inhibition) may be developing and driving change in action control types.

3.4.4 Comparison with other theories

The appropriateness to model the habitual system by model-free RL, and the goal-directed system by model-based RL, has recently been questioned.

In the same line as previous accounts (e.g., Kool et al., 2017a, S. Lee et al., 2014), Miller et al., 2019 argue for a weighted average of habitual and goal-directed systems which depends on the reliability of estimate or a cost-benefit assessment of the systems. Their model, however, considers the strength of the habitual system as proportional to how often an action has been taken, regardless of its outcome. This is against the classical associative learning account, on which the RL theories are aligned, where the learning is contingent on the outcome.

In opposition to Miller et al.'s (2019) account, Perez et al. (2020) stand by the classic learning view where habit strength is modulated by actions' outcomes. They argue that the balance of habitual and goal-directed systems is best understood as their relative strength across training, which follows directly from the linear sum of their strength. Perez et al.'s (2020) account is partly consistent with our model where the 'winning' system is determined by the system providing the strongest activation early enough (so that its favoured action reaches threshold before the other- in case of conflict). An important difference between Perez et al.'s (2020) model and our model is that in our model the combination of outputs is not a linear sum unlike in Perez et al.'s (2020).

Another view is proposed by a recent strand of research that explains habitual behaviour in terms of Hierarchical Reinforcement Learning (Dezfouli and Balleine, 2013). Based on a variation of the two-stage task, Dezfouli and Balleine (2013) found that a hierarchical model fit the data better than the MF-MB hybrid RL model. Specifically, their model explained the choice patterns at first stage just as well as the alternative model, but explained better second-stage choice patterns. Specifically, the Dezfouli and Balleine's (2013) model explained the fact that, following a rewarded trial, and having reproduced the same first-stage action as the previous trial, participants also tended to reproduce the preceding second-stage action. The participants repeated first- and second-stage actions even after rare transitions, suggesting the execution of a sequence could be insensitive to within-sequence feedback (when, in that case of rare transition, a flat goal-directed system would have taken a different second-stage action). Dezfouli and Balleine (2013) concluded that what is taken as model-free control may be a form of model-based selection of action sequences. A. Morris and Cushman (2019) used multiple other variations of the two-stage tasks and found, like Dezfouli and Balleine (2013), that humans employ a form of model-based control of sequences. However, A. Morris and Cushman (2019) further argue that humans still show model-free control in addition to model-based selection of sequences.

3.4.5 The two-stage task in general

After discussing the modelling aspects, we return to the experimental design itself. The task has been found to show significantly different relative MF/MB contribution following what may seem as minor changes in the paradigm: changing the range of possible reward probabilities (Kool et al., 2016), changing the rate of reward drift

(Kool et al., 2016) or changing whether the final probability across all final states do sum to 1 or not (Da Silva and Hare, 2018).

Consequently, it is possible that the differences observed across age groups are not due to fundamental developmental changes in the processes underlying the combination of MF and MB system, but are due to different understanding of the rules of the task, motivation, and/or valuation of reward (the gamified treasure and cash reward) of children compared to adults. Beyond basic comprehension checks, it is difficult to ensure that children have a deep understanding of the task or were executing it very well (that is, trying to consider rewards and transitions). Although the stimuli and the instructions have been adapted to children, such experimental design (repeated sequential choices by pressing keys) might not be the most engaging, nor the most ecologically-valid way to assess children's behaviour. Additionally, the necessary step of mapping the stimuli on screen to the keys on the keyboard adds further noise between the action selection process and the response.

Nonetheless, as mentioned earlier, this type of two-stage task is widespread in the adult literature and its variables have shown significant correlations with neural correlates (e.g., role of dopamine: Wunderlich et al., 2012) or constructs in multiple domains (e.g., clinical: Montague et al., 2012) or in various fields. Despite limitations, the replication of findings in multiple studies with variants of the task does suggest that the task overall taps into a meaningful balance between two forms of learning and action selection. Furthermore, this allows borrowing further findings from the task, that even within an adult population, may inform development such as the earlier-mentioned finding that cognitive control abilities predict recruitment of model-based control (Otto et al., 2014).

In summary, the results do suggest strongly that there are quantitative changes of children recruitment of model-based behaviour across development. Such finding should be corroborated with more tasks, that are child-friendly to ensure the behaviour of the children on the task reflects the recruitment of model-free and model-based type of control.

3.5 Conclusions

To conclude, the novel IAC model introduces a mechanism for the dynamic interaction between model-free and model-based pathways, where the systems' relative contributions can vary from trial to trial. The model successfully reproduced key behavioural patterns. It was able to do so when implementing different theoretically-motivated accounts. As such, it goes a step beyond the standard RL model which is rather at the computational level (following Marr, 1982's levels) by allowing one to implement different algorithmic accounts.

Based on choice patterns alone, it was not possible to discriminate well amongst the different accounts (only the *learning rate* produced choice patterns incompatible with the data and could be excluded on this basis). We further simulated response times. First, the response time data allowed good discriminability between competing models. Second, on the basis of this work, the following hypotheses appeared the two most plausible to explain the apparent increasing contribution of model-based strategy to choices with age: *(a)* that computations in the MB system are noisier in children than adults (but not in the MF system), and (*b*) that there is a default recruitment of one system over another with age. Further models should be developed to explore the many plausible developmental hypotheses and further experiments should discriminate between the candidate hypotheses. Given the difference in complexity and speed of computations in model-free and model-based systems, experimental manipulations of time pressure should be particularly helpful in differentiating the systems' signatures. Furthermore, given the difference in flexibility between the two systems, manipulations that occur in the middle of the planning or execution process should impair the systems differentially.

All in all, this model provides a proof of concept of the suitability of modelling the mechanism by which model-free and model-based control are combined. It also provides evidence for the novel model's ability to tease apart candidates developmental hypotheses and/or ultimately to conceive new tasks that can be used to better understand changes in children's use of habitual / model-free and goal-directed / model-based strategies. The following chapter further explores the relationship between habitual / model-free and goal-directed / model-based action selection in children. Specifically, it looks at how an unexpected change, which requires recruitment of a goal-directed strategy, affects children's behaviour in a simple routinised task.

Chapter 4

Action Sequence Control following an Unexpected Change: Two Touchscreen Studies

Abstract

This chapter investigates action control in 5- to 9-year-olds with a touchscreen game and two executive function tasks. The touchscreen game consisted of making repeated action sequences. After a training phase, the game included unexpected perturbations mid-sequence to investigate the ability of children to adjust an ongoing sequence, which presumably reveals model-based or supervisory recruitment. In order to dissociate between subprocesses of supervisory control (principally monitoring and inhibitory control), the unexpected perturbations varied at two different levels (state or goal) and two types of errors were made possible.

Younger children performed worse than older children when facing certain perturbations, which is consistent with the view that a model-based or supervisory system improves beyond 5 years of age. Furthermore, the within-trial variability of speed is proposed to be a kinematic marker of the recruitment of model-based or supervisory control when required in the task. The younger children also struggled more when facing a change of goal than when facing a change of state that was a consequence of their action, suggesting that the ability to monitor high-level goal-related cues develops across school-age years. A predominant error in both age groups and both types of transitions was the failure to update an ongoing action sequence. Taken together with the stronger link between the inhibition score and the action control optimal selection rate (compared to age and switching score), this suggests an important role of inhibition in the emergence of model-based type of control.

All in all, the chapter suggests that monitoring and inhibition of an ongoing sequence are both critical skills underlying model-based or supervisory type of control that develops throughout the school-age years.

4.1 Introduction

This chapter further investigates children's use of two systems for action control; whether conceptualised as *'model-free/model-based'* systems as in Decker et al.'s (2016) study used in chapter 3, or *routine/supervisory* systems following Norman and Shallice's (1986) dual-systems theory. In the previous chapter, Decker et al.'s (2016) task had the merit of measuring the apparent balance of *model-free* and *model-based* systems in children with a task that can be linked to many other studies with adults using a similar paradigm (the so-called *'two-stage task'*). However, we saw that the *two-stage task* was perhaps not well suited for understanding the mechanisms underpinning children's behaviour. Furthermore, the task's rules and concept are probably too abstract to use the task with children younger than Decker et al.'s (2016) sample (8-12 years old). Crucially, as argued in chapter 1, the period around the start of school (5-6 years old) and beyond is particularly interesting in terms of changes in sequential action control. This led us to design a new task, that would dissociate the recruitment of *model-free* control and *model-based* control, and be inspired by the *two-stage task* while making up for the shortcomings of the latter.

The newly-developed task was engaging for children and more concrete through gamification and the use of a touchscreen. The touchscreen also had the advantage of making the task relatively more ecologically relevant than responding on a keyboard. Contrary to Decker et al. (2016), our approach did not aim to find a 'default balance' or strategy that children would use across identical trials. Instead, we aimed to find how much children of various ages would be able to recruit *model-based*-like control following unexpected changes in the task, which made a change of action plan necessary to fulfil the instructed goal. In this task, the goal and how to achieve it was made explicit, easy to understand, and rewarded throughout the game by various animations.

4.1.1 Multiple processes interacting in action selection

To investigate action control as a whole, one must consider together multiple underlying processes and their interactions. Firstly, the routine system (contention scheduling in the original terms Norman and Shallice, 1986) is held responsible for selecting actions through schemas. A schema is a unit representing a subsequence of actions, or at a lower level, a motor skill that fulfils an action. We will talk indistinctively about schema selection or action plan, to designate an action or a sequence of action selected at a given point in time. An *action* in this computerised experiment is defined as reaching for a given icon.

The next processes we mention are part of the supervision system, which modulates the routine system. We ignore the goal generation process, since our task will have an instructed goal, where the sequence to be generated to achieve the goal will be explicitly instructed and presumably sufficiently easy to remember for our age range. The detection/monitoring and inhibition functions are central to our study.

We know that the perceptual salience of items can affect action selection by capturing attention even if the items are irrelevant to the goal (cf. Chapter 2). Our study will attempt to control for purely perceptual salience to focus more on detection of cues as a result of monitoring at various strengths, or differential allocation of attention to specific cues. We use monitoring in a broad sense (e.g., as used in Carter et al., 1998 and Pouget et al., 2017), which involves (a) anticipating or predicting the outcomes of action, and (b) comparing such predicted outcome to the response in order to detect a potential mismatch. We broaden it slightly to include predicting possible changes of the environment that may not be a consequence of the actions. Note that the monitoring signal can be described in reinforcement learning terms (difference between expected and obtained reward) and thus correspond to the reward prediction error used in chapter 3.

Some aspects of monitoring in action production have been found to evolve around school ages. In a problem-solving task, children as young as 4-6 years old were able to efficiently anticipate the consequences of potential actions and use this to guide their choices (Kaller, Rahm, Spreer, and Mader, 2008). Furthermore, to successfully detect discrepancies from an expected response, one must maintain information in mind, specifically maintain the goal and the actions to be carried out. Marcovitch, Boseovski, et al. (2007) has shown that a simple manipulation can disengage 4- to 5-year-olds' attention from the goal, impairing their ability to maintain the tasks' goal in mind. As noted by Stedron et al. (2005), the development of working memory and selective attention likely share common mechanisms and thus undergo the same developmental progression, and working memory is known to improve throughout school-age years (Diamond, 2013). If there are developmental differences related to maintaining goal-relevant information in mind, the differences may be when it comes to managing the overarching goal which is more distal rather than monitoring the perceptual consequences of action (Freier et al., 2017).

Following the successful detection of a mismatch, an ongoing action must be inhibited in order to be corrected. We therefore hypothesise that inhibitory control plays an important role in error correction (this is expanded on in Section 4.3.3). Inhibitory control also undergoes important changes in our age range as shown by many such as Reck and Hund (2011), Carver et al. (2001), and Lewis et al. (2017).

Finally, following successful detection of a needed update- and successful inhibition of a previously planned action (schema), a mechanism to select the replacement action is needed. The strategy needed to correct varies greatly according to a given task and will thus be discussed in Section 4.4 in the specific context of our experiment.

To our knowledge, no prior study has studied such processes integrated together within a sequence of several actions in children.

4.1.2 SuperCook: A novel action control task

4.1.2.1 Aims of the SuperCook task

This study aims to identify the behavioural signature of the two systems of action control (model-free or model-based following the terms of Decker et al. (2016), or routine and supervisory following the terms of Norman and Shallice (1986)), and investigate their interaction in the control of a sequence of actions by young children. Specifically, it looks at whether children can judiciously recruit one system or the other, as well as how this varies across school-age years.

We designed a gamified task reproducing key elements of sequential action control and allowing the investigation of age differences in action control under different degrees of influence of the two systems. Namely, the task involved practising appropriate sequences of actions varying between one of two goals, practised a large number of times (during a *training phase*), proposedly transferring the control to the routine system. In a second phase (*test phase*), the task had trials similar to the training phase (*common transitions*) as well a small percentage of unpredictable perturbations (*rare transitions*) in the middle of the sequence. The perturbation was designed to change the nature of the second action that would have led to the best reward in the absence of that perturbation.

We were interested in whether children of different ages would tend to continue with their (assumed) initial sequence plan, or update the sequence to appropriately obtain the maximal reward. We propose that such an update of plan requires higher recruitment of the supervisory system than in the absence of perturbation. By the rare transition manipulation, we thereby attempt to dissociate *model-free/routine control*, which would be mostly responsible for executing sequences after the *training phase* and during *common transition*, from *model-based/supervisory control*, which would be needed after *rare transitions*. The task also goes further than identifying the relative amount of model-free/model-based control. Specifically, assuming that lower model-based/supervisory control is observed in younger children, the task attempts to dissociate between the roles of monitoring functions and inhibitory function in the action sequence update.

Additionally, the task had two subtypes of rare transition (one changing the goal cue: '*rare_goal*' transition, one changing the action-effect cue: '*rare_state*' transition) to better understand the developmental change in monitoring the environment and the overarching goal during the execution of action sequences. Therefore, we can assess to what extent the supervisory system can regain control when external events require the replacement of an action after a sequence was initiated; and if so, under which conditions (e.g., following what kind of environmental change).

4.1.2.2 Description of the SuperCook task

We fully describe the task here because it is fundamentally similar for the two experiments in this chapter (Section 4.2 and Section 4.3). The differences will be described in the Methods of each section.

Design of the task

The game was played on a touchscreen. We chose the context of baking, which would be familiar enough for this age range (indeed, at preschool children are already familiar with watching meal preparing activities: Freier et al., 2015). This would facilitate learning of the overall rules, while the combination of goal and action sequence was completely arbitrary to limit confound with prior knowledge.

Here we describe the main part of the SuperCook task, that was split into a *training* followed by a *test* phase. In the task, a cat and a rabbit cartoons were to be fed many cakes, each character having their favourite cake recipe. A recipe was made of a chocolate bar (black or white chocolate) and a fruit as shown in Figure 4.1.

Fig. 4.1: **Mapping of the action sequences (favourite cakes) per goal cue (character) according to one version (left panel) or the other (right panel), the versions being randomised across participants. This image was also used as instruction materials.**

The two correct recipes were made of equal intuitiveness (that is, both fruits were not the most common fruits found on chocolate cakes). The mapping between a character and its 'favourite' recipe (e.g., white chocolate with pear for the rabbit, and dark chocolate with grape for the bunny) was pseudo-randomised across participants. A trial's goal was cued from the start, and throughout the trial, by a character shown in the centre (e.g., in Figure 4.2 one had to *make the cat's favourite cake*). Icons representing ingredients were also present on screen, as well as a central preparation bowl which reflected the state of the cake preparation.

Icons were to be collected with a spoon icon then dragged-and-dropped to a central location on the screen. A sequence consisted of three steps, but we will be interested in a two-step action plan since only the steps number 1 and 2 required action selection and the third one acted as a terminating action, as will be detailed later.

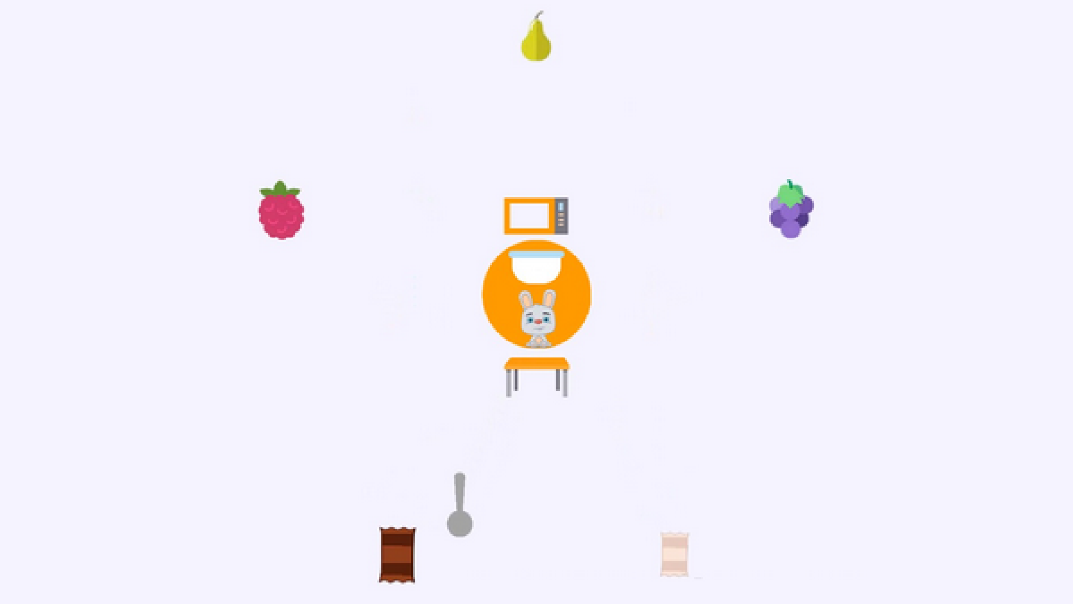


Fig. 4.2: **Training and test phase: Layout of the game at the start of a trial (in this example, the type of goal is 'make the *bunny*'s favourite cake') for steps 1 and 2.**

Transitions and reward structure Different types of trials were included in the test phase. The type of trial occurrence was made unpredictable by pseudo-randomisation. There were 70 % *'Common trials'*, which were the same as trials in the training phase. There were 15% *rare_state* and 15% *rare_goal*. In *rare* types of trial, the course of a trial was the same for *common* trials up to the moment when the first step was completed (i.e., first icon brought to the table). At this moment, an unexpected change happened: in *rare_state* trials, the bowl showed the colour of the unselected chocolate (e.g., dark chocolate colour when the white chocolate had been drag-and-dropped; while in common transitions the bowl would have shown the white chocolate colour), while in *rare_goal* trials, the character displayed was swapped with the other one.

The rule was to always try to make the character's favourite sequence (indicated by a reward 2 in Figure 4.3) but when an unexpected transition happened, to make the second favourite recipe of the character (indicated by a reward 1 in Figure 4.3), which had the raspberry as second action. The entire reward structure is shown in Figure 4.3 for a given character, and for the other character the structure was symmetrical starting with dark chocolate as first action.

This third option was introduced to understand better the process behind sequence failure. Indeed, when on rare transitions the second-stage action was not the correct (maximally-rewarded) one, it left out two incorrect options whose selection would be informative of the process behind the incorrect choice. One type of error, that we denote as *'failing to update'* error, could be to select the second-step icon that would have led to the best reward if there had not been a rare transition. For example, in the case illustrated on 4.3 (the goal cue is the rabbit, the maximal

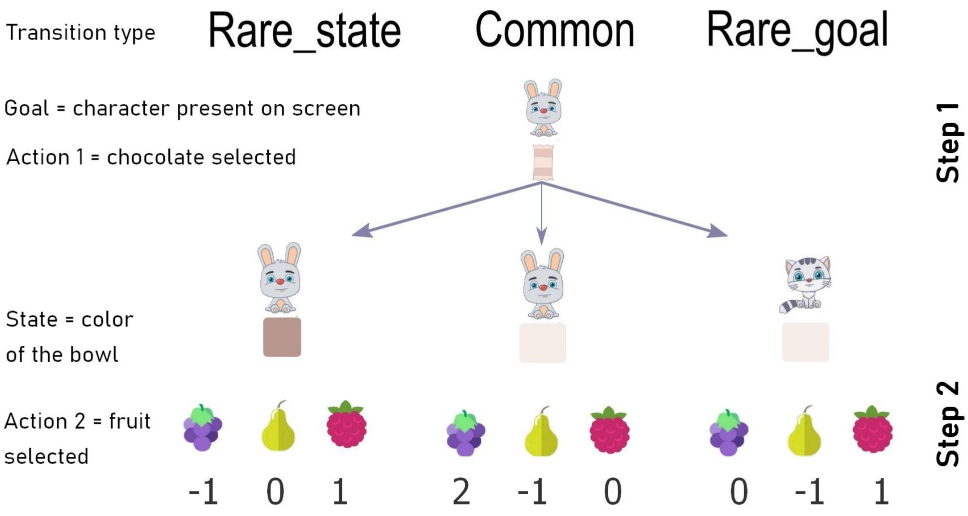


Fig. 4.3: Reward according to the three types of transition that may occur and the sequence carried out. In the illustrated example, the type of goal is 'make the *bunny*'s favourite cake', and the random mapping is such that the best cake is the one with white chocolate, and we display the cases following selecting the white chocolate at step 1.

reward sequence has white chocolate and grape), selecting the grape instead of the raspberry would constitute such error. Another error could be what we will call *'activating the alternative prepotent plan'*, that is updating the second-step action by using the other sequences' best action. In the case illustrated on 4.3, selecting the pear instead of the raspberry would constitute such error.

Trials in the *training* and *test* phase The course of a trial in the training phase was the following. To begin with, the spoon icon had to be dragged-and-dropped to the chosen chocolate icon. To select the spoon, the screen had to be touched within a radius of 3.8 cm around the centre of the spoon (all selection parameters were fined-tuned to obtain a good balance between smooth use of the game and precise touch-trajectory recordings). If the participant released the touch (intentionally or not) on a given icon, the icon would stay at its location and the movement could resume by touching back and continuing the drag-and-drop action. When the spoon icon would touch an icon (that is, when the spoon got within the 'selection radius' of 1.9cm around the icon), the icon would become selected i.e. visually attached to the spoon (see the right side in Figure A.1) and displaced together with the spoon. Once selected, an icon would be detached when placed to a target location (which we describe hereafter). The attached spoon-icon was to be drag-and-dropped onto the table on the centre (correct target location). Alternatively, a food icon could be swapped with another food icon by getting into the selection radius of the other icon (a process repeatable if the touch went more than 7.1 cm away from the swapping location and back). When participants brought the spoon and icon to the table (within 3.5cm), the chocolate icon would disappear and a change of bowl state would be triggered, representing the colour of the chocolate 'placed in it' (this, and more illustrations of the appearance of the game being played can be seen on the right side of Figure A.1). This marked the end of *Step 1* and the beginning of *Step 2*. If the finger stayed on the touchscreen, the spoon was still selected. At that time the spoon had to be brought to a fruit icon, which had to be brought to the oven in the centre. When the spoon with fruit icon reached the radius of the target oven, step 3 would be triggered, starting with an entire change of screen as described in Section 4.1.2.2 (with an icon of a baked cake in the centre, and the icon of the character on one side, as can be seen at the bottom right of Figure A.1). There, the baked cake icon had to be dragged-and-dropped onto the character ('to feed the animal'). This would terminate the trial and play the feedback for 1000 ms. The feedback showed one the characters of Figure 4.4 depending on the sequence carried out.

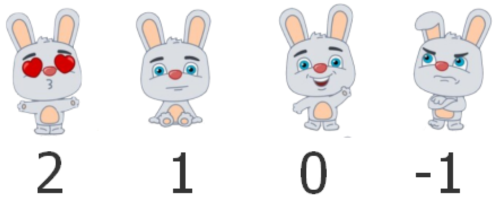


***Fig. 4.4:* Images of the four possible types of feedback. The mapping between the goal cue, the sequence carried out and the reward value (which defined the feedback image as shown) is presented in Figure 4.3.**

Although the game allowed participants to select and move any food item at any point in the sequence, we reduced the degrees of freedom in how the task could progress. At the start, only the spoon could be selected (not the other icons). Only placing the chocolate icons on first step would update the bowl's state, and only a fruit could be placed in the bowl after a chocolate had been placed. When invalid actions were attempted, the icons would return to their initial location and the trial would go back to where it was prior the incorrect attempt. Trials with an incorrect type of icon selection attempt (chocolate instead of fruit or fruit instead of chocolate) were flagged for analyses as 'incorrect icon step attempt'. The spoon icon, necessary to select food items, was designed to be able to record the kinematics of selection of the food icons. Our interest was to analyse this 'selection' segments of the game. To mimic the fluidity of real-life action sequences, we ensured the steps could follow one another without lifting the spoon or hiding the view of relevant icons, which is why the first step icons where on the bottom, the target table was below the character and the selection distance were large enough. To ensure that we captured selection at the intermediate action level, and not purely at the motor level, the location of the icons on the screen changed (among their upper or lower half slots) randomly at each trial. Finally, the sequence of goal presentation was pseudo-randomised.

Trials in the familiarisation phase The familiarisation phase was only aimed at letting the child get used to the touchscreen (level of pressure needed to act on the touchscreen, reactivity, etc.) The start screen for the familiarisation phase is shown in Figure 4.5, had one spoon, one central area (same as for the non-familiarisation trials) and a chocolate and fruit that were resembling but different from the non-familiarisation trials icon. A spoon had to be drag-and-dropped towards the chocolate icon, which would result in attaching the chocolate to the spoon, then it had to be brought to the central area and then the spoon to the fruit, which also had to be brought back to the central area, resulting in finishing the trial.

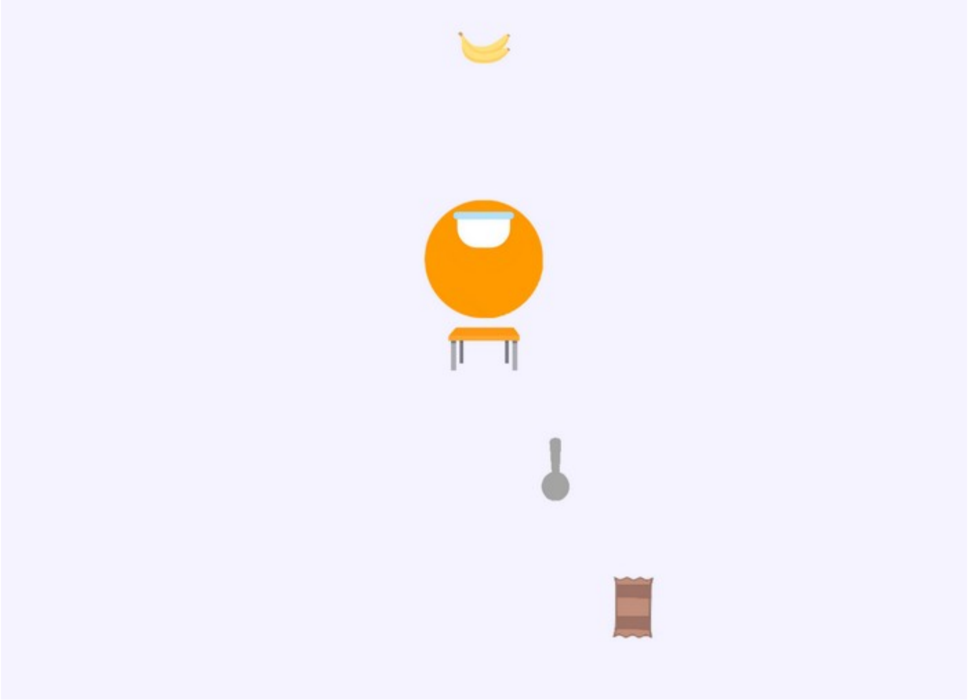


Fig. 4.5: **Familiarisation phase: Layout of the game at the start of a familiarisation trial.**

Materials

Stimuli The contrast between the colours of the chocolate and the fruits icons, as well as the overall visual impression were matched among icons (Figure 4.2 represents the layout during Step 1 and Step 2 of the sequence in the training and test phases). The position of the icons on the screen was adjusted to get sufficient surface to travel for richness in kinematic data while allowing comfort of touchscreen use given the arm length of the shortest participants. More precisely, the three fruit icons were placed on three possible slots in the upper half, while the two chocolate icons were placed on two possible slots in the lower half. The locations of fruits on the three slots, and of chocolates on their two slots were randomised at every trial. We will give the dimensions in cm, and the display ratio was 1 pixel = 0.026cm. For the coordinates, we take the centre of the spoon icon as the origin (0,0), with the x-axis horizontal and pointing to the right and the y-axis vertical and pointing upwards. The spoon icon's dimensions were 1.2x2.7cm. Overlaid on the spoon icon was a table pictogram of dimension 2.8x2.1cm. The food icons were equidistant from the centre, with a spacing of 10.9 cm, and there was a spacing of 11.8cm between each pair of neighbouring items within their lower or top half. Hence the slots for initial start were located at (in cm): (-10.9,4.55), (0,10.9), (10.9,4.55), (-5.9,-9.2) and (5.9,-9.2). The moving icons and the bowl were sized 2.5x2.5cm. The character was sized 2.5x2.5cm, located 2.2cm above the spoon, and 2.4cm below the bowl; both characters and the bowl were displayed on top of an orange disk of diameter 4.9cm. The close location of the character and the bowl was an important feature: it allowed us to limit the difference in perceptual salience between both cues (as compared to having one central and the other peripheral). Step 3 of the sequence happened on a different screen, which appeared at the end of step 2. That screen displayed the 'resulting' baked cake (as visible in Figure 4.1) in the centre (size 2.9x3.8cm), and the character, of size 3.4x3.4cm, either at the (-10.9,4.55) or the (10.9,4.55) location (pseudo-randomised across trials).

Apparatus We used a 23-inch (48.77 x 27.43 cm) touchscreen monitor (ProLite T2336MSC iiyama). The game was programmed in Python, and the interface to the touchscreen was carried out by embedding the code in OpenSesame 3.2.8 with the *inline_script* functionality. The x, y coordinates were sampled at 20 Hz. Participants sat in front of a child-sized table facing the screen laid at an angle of 30°. The position of the chair was adjusted so that the alignment of the dominant hand was centred when reaching towards the screen.

Procedure

The game had three phases: a familiarisation, training and test phase. The familiarisation phase included 4 trials; the training phase 4 practise plus 16 trials; the test phase 8 practise plus 52 trials.

After the familiarisation phase, the main rules were introduced with a cover story by pre-recorded instruction videos (cf. script in appendix A.1.1). The video showed the game's screen on one side, and a person talking through the game on the other side. There was an instruction video before the training phase, and another one just before the test phase. The video script introduced the cover story and the associated actions to be made, which was illustrated in parallel by a recording of the game being carried out.

In the familiarisation phase and in the first round of practice trials (after the training phase's instruction video), the experimenter would play along with the child to help them start and invite them to continue by themselves, giving help when needed. In the second round of practice trials (after the test phase's instruction video), the experimenter would only assist, correct or repeat the rules when necessary (after the video for that part had been viewed). The task ended with a final thumb up screen and congratulating the child on their performance.

We aimed to encourage the children to do the game not too slowly. This was (1) to ensure that we could capture within-trial dynamics with kinematics on the touchscreen, (2) to maximise the chances of observing errors, and (3) to make it possible for the entirety of the study being carried out in the time frame of a single visit. Thus, the instructions (cf. appendix A.1.1) encouraged children to go as fast a possible, and there was a timeout if any step lasted over 15 seconds. A time out screen, showing a clock watch, was displayed for 500 ms and the same trial (same goal cue) was repeated.

To facilitate task learning and to speed up automatisation, the correct mapping was explicitly instructed at the beginning of the task. We verified that the children had understood and memorised it by asking them to tell us the favourite cake of each character prior to starting the game. In addition, feedback was given after each trial (as described previously) so that the children knew how accurate they were. To ensure that children would stay engaged and keep pursuing the maximal reward (as instructed) despite the repetitive nature of the task, we supplemented the feedback character image with 5 different approving cartoon sounds (e.g., 'yipee') randomly selected at each trial. In addition, every 5 trials a gif animation of the characters 'dancing' was shown for 4 seconds while entertaining music was played provided the preceding sequence was correct, if not correct it would be postponed until the subsequent correct sequence.

4.2 Experiment 1: A study Of Kinematic Measures in the SuperCook task

4.2.1 Introduction

This Experiment had two purposes. The first one was to validate the novel SuperCook task (in order to ensure that the children in our target age range would understand the rule and stay engaged with the game, as well as to adjust the task).

The second aim was to identify appropriate kinematic features for experiment 2, with data independent from the final sample of Experiment 2. Our proposed feature selection procedure aims at circumventing challenges of analysing kinematic data while avoiding over-exploration of the dataset which could lead to false positives as well as keeping interpretable features.

4.2.2 Kinematics measurement

Measurements such as errors and response time may be insufficient to distinguish among candidate mechanisms involved in sequential action control, given the multiplicity of cognitive processes interacting together. Kinematic measures have proven informative in studies of action control with children as presented in chapter 1. The SuperCook task was performed on a touchscreen in order to measure the kinematics parameters of action selection which holds the potential to be an informative complement to response time and choice data. Apart from providing additional measures, the touchscreen has the advantage of being more engaging and ecologically relevant than a task with a computer and keyboard.

To tackle the high-dimensionality of kinematics data, several analytic approaches can be taken, as presented in Section 1.4.2. We will illustrate how two of the approaches would work on our data. The *machine learning approach* involves using all the data points (unreduced dataset); it is ideal if one wishes to extract a maximal amount of information from the data and is less concerned by interpretability. Applied to our study, let us assume that the kinematic features are expected to differ pre- and post-training (due to quantitatively or qualitatively different modes of action control). We will call this theoretical dataset ’training set’. Let us assume we want to see how a given experimental manipulation, post-training, influences the mode of action control (embodying our ’test set’). We could use machine learning to train a classifier on the training set, with instances labelled as pre- and post-training, and subsequently use it on the test set to classify the experimental conditions of interest (thus allowing to characterize whether the experimental manipulation yields to one or the other pre-defined action control modes, that matches the pre- and post-training data). This approach has the advantage of using all the available information, but the application of such a classifier would be limited to the exact same experimental paradigm. Therefore, the approach lacks generalisability. Another drawback of such an approach is the lack of interpretability: the mass of data points itself does not tell us anything about what trajectories are like. The approach would tell us that some experimental condition resembles pre- and post-training, but not tell us *how* they resemble each other (i.e., what kinematic features they share).

Another approach mentioned in Section 1.4.2 is the *single feature approach*. It would consist in pre-identifying a few features that we assume would be informative based on the literature. A drawback is that no study of the kind of SuperCook has analysed kinematics and thus this approach would risk missing out on very informative features. Aside from the literature, we could imagine that a variety of features might be informative of the action processes at play in the task. Testing all of them would cause a multiple comparisons problem.

A data-driven approach to feature selection would be optimal but as mentioned this faces the challenges that kinematics inherently differ for any different task setup. This study started with a substantial pilot, which allowed us to circumvent this problem by using this pilot as a dataset for feature selection for the SuperCook experiment. Our feature selection approach involved defining a large number of features, based on the literature and based on their potential to capture meaningful effects in the task, followed by deriving a machine learning approach to perform a selection of interpretable features.

4.2.3 Methods

4.2.3.1 Participants

A sample of 13 children aged between 4 and 9 years old took part in this pilot of the SuperCook task. Children were encouraged to complete all trials but the experimenter would stop the session if the child was not happy to cooperate, or would wait until the child was bored if the child was not following the rules of the game. Given that the pilot experiment aimed at assessing the suitability of the task, and at performing broad analyses on kinematic data indistinctly of the age, every pilot child's data was included even if the child had not completed the entirety of trials.

All participants were recruited via Birkbeck's Babylab database, and were tested individually at the Babylab. Written consent was obtained from the parents as well as verbal consent from the children. The experiment was conducted after receiving approval from the institutional ethics committee according to the principles of the Declaration of Helsinki. Families received a thank-you gift for taking part in the study.

4.2.3.2 Materials and procedure

Section 4.1.2 described the latest version of the task and procedure that was used in the main experiment, benefiting from experience with the version in this Section. The differences from the task described in Section 4.1.2 are presented below.

In the present pilot version, the upper- and lower-half locations of the fruits and chocolate icons are swapped compared to the version in experiment 2. Additionally, in the pilot there is no microwave icon. Finally, all instructions in the pilot were given verbally by the experimenter (instead of given by a pre-recorded video).

4.2.4 Results

We followed the following procedure of kinematic feature selection. First we transformed the set of (x,y) coordinates at each time point into a set of 43 features (for a given trial). These features are defined in table 4.1.

Table 4.1: **Kinematic features computed and included in the PCA. Every feature was computed on single trials' data. The expression *chosen icon* designates the icon that is selected at the end of the trial. The features originated from the literature or from intuition about potentially relevant metrics. Table part 1.**

Feature short name	Feature definition
maxOrthogDist	Maximal orthogonal distance between the trajectory samples and a straight line between the start and chosen icon.
spatialDist_atMaxDist	Maximal distance between the trajectory samples and a straight line between the start and chosen icon.
time_atMaxDist	Time at the point where spatialDist_atMaxDist is reached.
pathLength	Sum of the absolute difference between every pair of adjacent trajectory samples.
closestPoint	Smallest distance to the closest alternative icon (among the non-chosen icons).
closestPointRel	Smallest value of the ratio of the distance to the closest alternative icon to which the trajectory gets closer (among the non-chosen icons) over the distance to the chosen icon.
maxLogRatio	Maximal log ratio of distance to the closest non-chosen icon over the distance to the chosen icon.
proportionPointsQuadrant	Proportion of the samples spent in the quadrant of the chosen icon.
avTrajAngle	Average of the instantaneous angle between the trajectory and the most direct path to the chosen icon.
trajAngle_10	avTrajAngle over the first 10% samples of the trajectory.
angleAtPercentInterpolSample_10	*avTrajAngle* over the first 10% samples of the spatially-interpolated trajectory (that is, trajectory transformed so that it contains 100 data points spaced evenly)
trajAngle_20	avTrajAngle over the first 20% samples of the trajectory.
angleAtPercentInterpolSample_20	*avTrajAngle* over the first 20% samples of the spatially-interpolated trajectory (cf. *angleAtPercentInterpolSample*_10).
trajAngle_30	avTrajAngle over the first 30% samples of the trajectory.
angleAtPercentInterpolSample_30	*avTrajAngle* over the first 30% samples of the spatially-interpolated trajectory (cf. *angleAtPercentInterpolSample*_10).
av2dSpeed	Average of the speed's norm.
varSpeed	Variability of the speed's norm within a trial.
avXSpeed	Average speed along the horizontal axis.
avYSpeed	Average speed along the vertical axis.
varXSpeed	Variability of the horizontal speed within a trial.
varYSpeed	Variability of the vertical speed within a trial.
av2dAccel	Average acceleration.
avXAccel	Average acceleration along the horizontal axis.
avYAccel	Average acceleration along the vertical axis.
accelExtrema	Maximum of the absolute value of acceleration.

Table 4.2: **Kinematic features computed and included in the PCA. Every feature was computed on single trials' data. The expression *chosen icon* designates the icon that is selected at the end of the trial. The features originated from the literature or from intuition about potentially relevant metrics. Table part 2.**

Feature short name	Feature definition
avOrthogDist	Average of the signed orthogonal distance the trajectory samples and a straight line between the start and chosen icon.
AbsavOrthogDist	Average of the absolute orthogonal distance the trajectory samples and a straight line between the start and chosen icon.
maxTrajAngle	Maximal angle between the trajectory and the most direct path to the chosen icon.
areaTrajDirectpath	Area under the curved formed by the trajectory and the most direct path to the chosen icon.
nVerticCrosses	Number of times that the trajectory crosses the vertical axis coming through the centre of the screen.
nHorizCrosses	Number of times that the trajectory crosses the horizontal axis coming through the centre of the screen
nChangeXdirection	Number of change of direction in the horizontal axis.
epochSpeed_2	Average speed in the first 2 seconds of the trial.
epochXSpeed_2	Average horizontal speed in the first 2 seconds of the trial.
epochYSpeed_2	Average vertical speed in the first 2 seconds of the trial.
epochSpeed_3	Average speed in the first 3 seconds of the trial.
epochXSpeed_3	Average horizontal speed in the first 3 seconds of the trial.
epochYSpeed_3	Average vertical speed in the first 3 seconds of the trial.
rt	Response time, i.e. duration of the trajectory from the centre to the chosen icon.
epochSpeed_percEnd_70	Average speed in the last 70% of the trajectory.
epochSpeed_percEnd_50	Average speed in the last 50% of the trajectory.
epochSpeed_percEnd_30	Average speed in the last 30% of the trajectory.
epochSpeed_percEnd_20	Average speed in the last 20% of the trajectory.

Having considered many potential cognitive processes relevant to the task and their translation in (x,y,t) space, we did not believe we were discarding important information at this stage. However, we did not select among the features based on intuition, to make sure that we made the most of the remaining information.

We extracted the features from the data of step 2 of the test phase (indistinctly on common and rare transitions), because this will be the main interest in the final analyses.

Having defined and extracted all the features from the dataset, we sought to select among them (in other words, further reduce the dimensionality of our data) using our own variant of a Principal Components Analysis (PCA). Although PCA is traditionally used to extract and use components that are linear combinations of the original variables, called Principal Components (PCs), we took a slightly different approach.

The first step was the same as the traditional approach; that is, obtaining linear combinations of the input features (the PCs) such that the PCs are orthogonal (i.e. uncorrelated) to each other. We then ranked the principal components such that PC1 is the component with the largest proportion of variance explained, PC2 with the second largest proportion, and so on. Then, still following the traditional approach, we selected the minimal number of components that accounted for more than 50% of the overall variance. Hence, looking at the scree plot (Figure 4.8), we decided to retain 3 first PCs (they together explain 55.9% of the overall variance).

Now we turn to retaining the features that explain as much variance as possible. If we look at the input features that constitute the 3 PCs: the higher the features' contribution (loading) to the same component, the higher their correlation together; and features with high loading in a cluster will be less correlated with features with high loading in another cluster.

Having kept 3 PCs, we aimed to select 3 features. We wanted them to satisfy two constraints. The first is explaining a lot of variance, by taking a feature with a high loading in each PCs; with the rationale that if a feature has a large contribution to a PC, and we know that this PC contributes a relatively large amount to the overall variance, then this feature contributes largely to the dataset's variance (albeit less than the original PC). This choice was guided by the values shown in Figure 4.6. Secondly, we wanted these 3 features to be as little inter-correlated as possible. Thus, among the features that had quasi equal contributions to a given PC, we chose a combination of 3 features that had as little correlation to each other as possible: this was done by inspecting Figure 4.7 where features whose lines are perpendicular are uncorrelated, and the correlation increases with a decrease in angle below 90 degrees (or the negative correlation increases with an increase of angle away from 90 degrees).

Following this procedure, the chosen features were *avTrajAngle* (Average of the instantaneous angle between the trajectory and the most direct path to the chosen icon), *varSpeed* (Variability of the speed's norm within a trial) and *maxLogRatio* (Maximal log ratio of distance to the non-chosen icon over the distance to the chosen icon).

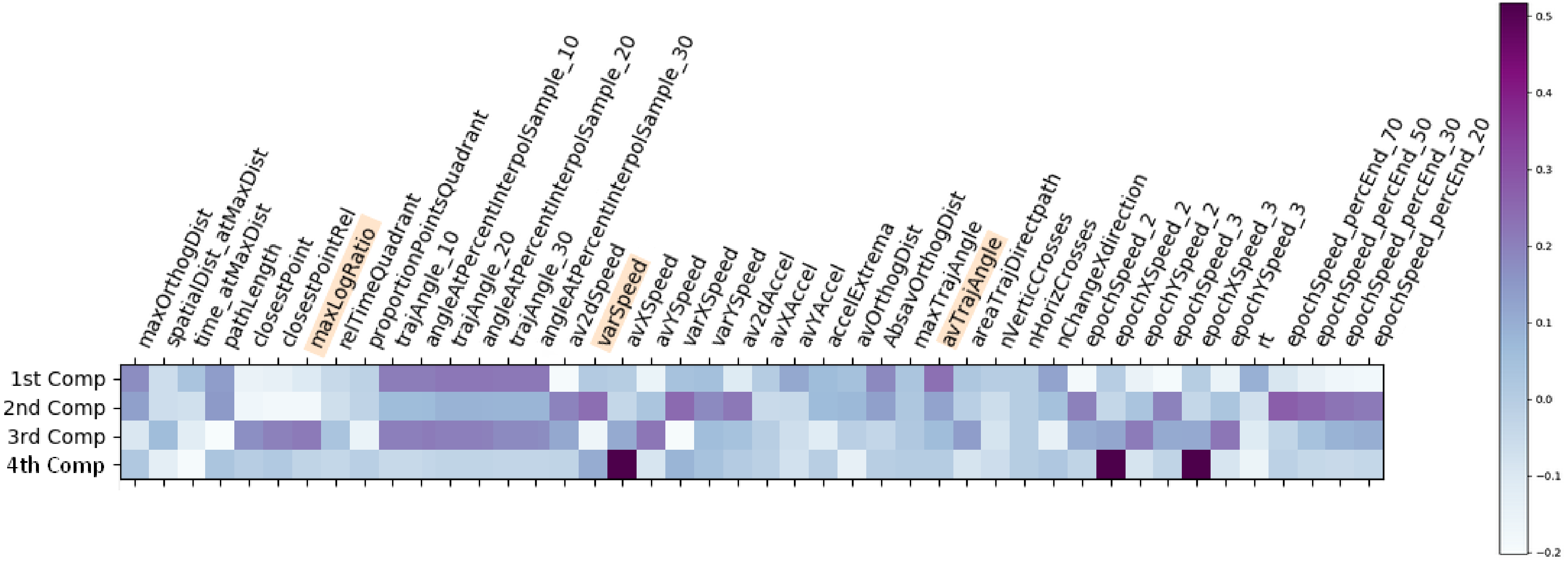


Fig. 4.6: **PCA (principal component analysis) output: contribution of features to the 4 principal components. The meaning of the features is defined in table 4.1. The highlighted features are the ones that were retained for further analyses.**

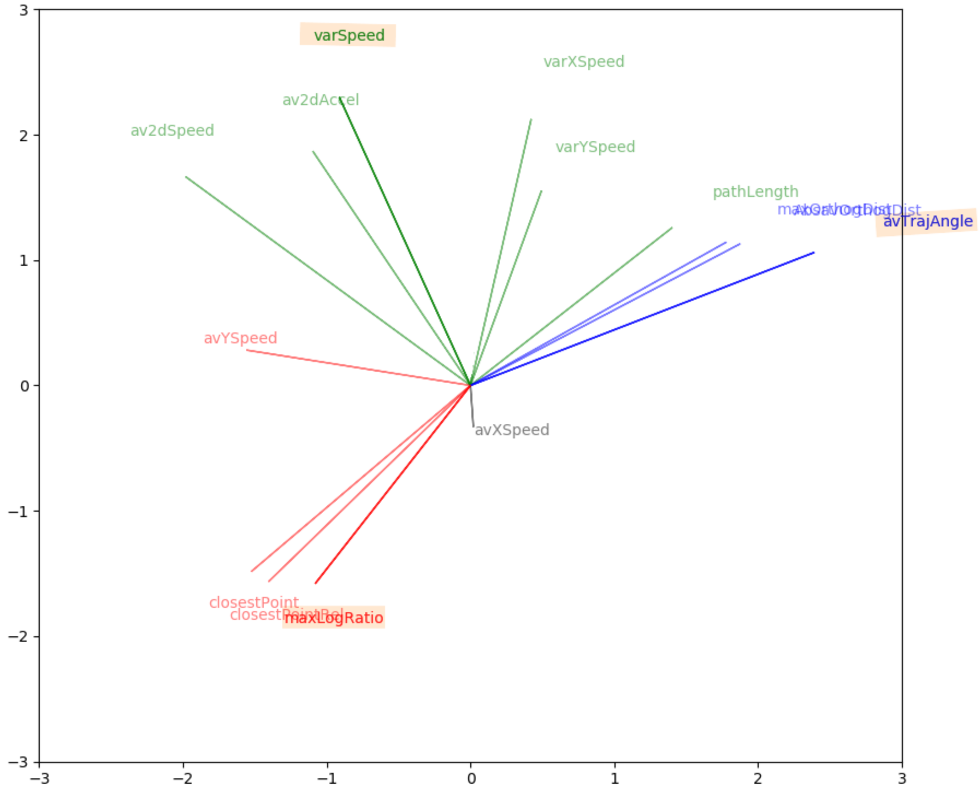


Fig. 4.7: **PCA output: correlation among features (an angle of 90°between two features means they are uncorrelated, acute-angle features have positive correlation, and obtuse-angled features have negative correlation). For visualisation purposes, only a subset of the features are shown. The entirety of the features entered in the PCA are listed in table 4.1 along with their meaning. The highlighted features are the ones that were retained for further analyses.**

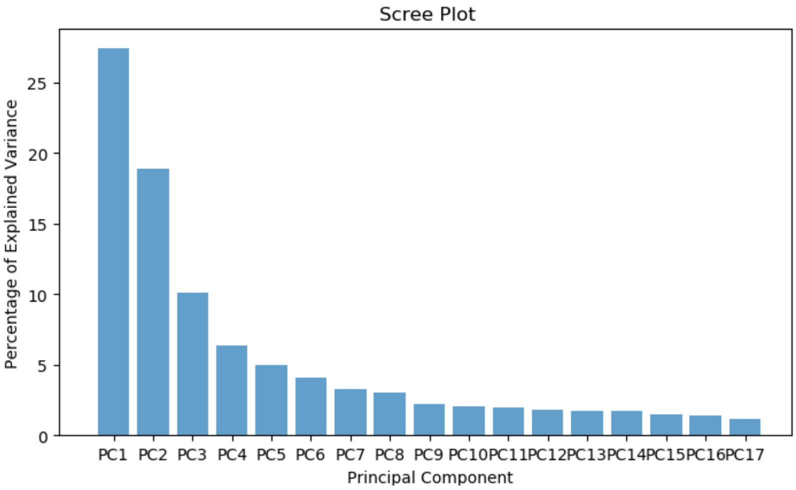


Fig. 4.8: **PCA scree plot: percentage of explained variance per PC (principal component) for the first 70 PCs.**

4.2.5 Discussion

This pilot experiment achieved two aims. One aim was the selection of kinematic features for experiment 2 (section 4.1.2). The 3 kinematic features *avTrajAngle*, *varSpeed* and *maxLogRatio* hold little redundancy (i.e. they are not strongly inter-correlated), they explain a good amount of variance in this dataset, and they are intuitively interpretable. Therefore, they will be used to test for the presence of meaningful developmental changes in the SuperCook experiment 2.

The other aim was to refine the materials and procedure and to find an age range suitable for the task. From observing the participants during the testing session, it appeared clearly that 4 year-olds were too young to complete the task while 5 years old were able to do so. Thanks to this pilot experiment, the age range as well as minor material changes will be used in the experiment 2.

4.3 Experiment 2: Effects of Unexpected State / Goal Transitions in the SuperCook task

4.3.1 Introduction

This experiment investigates the use of model-free (routine) and model-based (supervisory) types of action control in school-aged children. The dissociation between model-free and model-based types of control uses probabilistic transition, inspired from the study of Decker et al. (2016), but goes further in separating the subcomponents of supervisory control that might be at play and developing around that time. Namely, it investigates the difference in development between monitoring a cue that is more distal (the goal) versus a cue that arises as a consequence of the children's action (the state). It also looks at the different roles of monitoring and interrupting an ongoing sequence, as well as selecting the appropriate replacement action. Finally, it relates development in action control together with development in executive functions.

This experiment involves the final version of the SuperCook task, as described in Section 4.1.2.2. As a reminder, the SuperCook tasks involves performing two short action sequences. A training phase is followed by the test phase, which introduces key experimental manipulations. We analyse optimal action selection, response time and kinematics. The experiment also includes two separate executive function tasks.

4.3.2 Adjustments of the SuperCook task

The SuperCook task was adapted following observations on the pilot experiment (section 4.2). Firstly, the lower age bound was changed to 5 years old. Secondly, in the pilot experiment it appeared that when the chocolate icons (first step) were on the upper-half of the screen, the hand of participants was sometimes hiding the bottom-half of the screen, in a way that could interfere with the second stage's action selection. Consequently, one change made was to swap the upper-half or lower-half location of the fruits and chocolate icons. In this way, the two steps then consisted of reaching successively to the bottom and then to the top.

Another observation in the pilot experiment was that some participants were going around the orange disk (cf. Figure 4.2) to drag the second-step icon to the same bottom-location at step 1 (instead of taking the shortest path directly to the disk). Even though this section of the trajectory was not analysed directly, we wanted to standardise this across participants. Thus, in addition to the table icon below the disk, a microwave icon was added above the orange disk. With this new version, after dragging the chocolate icon to the centre, participants were prompted to go towards the table (i.e. the bottom of the disk because the trajectory started at the bottom), and when dragging the fruit icon to the centre, they were prompted to go towards the microwave (i.e. the top of the disk).

A last change was to add video-recorded instructions (along with adjustments in the instructions script) to ensure all participants received the same instructions for the task.

4.3.3 Executive function (EF)

In this experiment, we aimed to measure EF by using standard tasks in order to relate the action control processes more specifically to EF abilities and to their changes across childhood. We chose an inhibition task and a switching task. Each task is designed to measure more specifically one component, however all EF are known to be highly correlated (Miyake et al., 2000a). Thus, we do not claim to measure a ’pure inhibition component’ or a ’pure switching component’. Rather, when talking about inhibition and switching measured by these tasks, what we mean is ’EF with a dominance of inhibition’ and ’EF with a dominance of switching’.

To fit the main game and EF tasks within the time frame of a single testing session, we had to limit ourselves to two EF tasks. We chose inhibition and switching for the following reasons. Our primary hypothesis was that inhibition would be the most critical component linked to the use of supervisory control. Indeed, we assumed that one would prepare a three-step action plan at the beginning of the task, and that this plan would need to be inhibited following rare transitions to allow selection of the alternative action. However, the crucial process to successfully update the action plan following rare transitions might resemble switching between task sets more than inhibiting an action plan. Hence, we included a switching task to assess the separate impacts of developing switching and inhibition abilities on the SuperCook task’s variables.

It is important to note that there is a variety of tasks aimed at measuring inhibition or switching, and that they may conceptualise the construct differently from each other. Some inhibition tasks draw on a pre-existing prepotent response. For example, a child-friendly Stroop-like inhibition task require to respond "big" when seeing a small circle and "small" when seeing a big circle (e.g., Ikeda et al., 2014).

Other inhibition tasks build a new prepotent response during the task. For example, in the Go-NoGo/whack-a-mole types of task (Shapiro et al., 2013, Petitclerc et al., 2015) that we will use, a first phase trains the participant to respond to a certain stimulus, making the response prepotent, which must later be inhibited. Because our sequence production task involves repeating a previously unknown action sequence, it more closely resembles this latter type of inhibition task.

Similarly, switching tasks vary. For example, in Chevalier and Blaye (2009), the Advanced Dimensional Change Card Sort examines switching between rules, while other tasks such as in Crone et al. (2006) look both at rule switching and response switching (that is, measuring whether a response is different relative to the preceding trial, compared to a response repeated from one trial to the next). Our switching task will involve specifically *task-switching*, that is, switching between stimulus features and response rules, as will be detailed in the Methods (section 4.3.5.2).

4.3.4 Hypotheses

For the training phase, we hypothesised that all children would show improvements in response time with practice. Our key hypotheses concern the test phase. After completion of the training phase, the rare transitions will require changing what would have been the maximally rewarded action sequence if there had not been a rare transition (something that we coin *'initial action plan'*). In the framework of Decker et al. (2016), the rare transitions require the use of a model-based type of control. We hypothesised that younger children would overall struggle more than older children to update sequences after a rare transition, following the finding of Decker et al. (2016) that model-based use increases with age (albeit Decker measured the change between children and adults, we believe we would find a similar trajectory within a child sample). Similarly, we predicted that all age groups would show the same level after common transitions, linked to the finding that model-free use remained constant in Decker et al. (2016).

Furthermore, looking at whether errors (regardless of the type of errors) are more frequent following one or the other rare transition type will shed light on the relative importance of higher-level goal information compared to perceptual (state) information in driving the correct choice in a developmental population. Following the findings of Freier et al. (2017), we predicted that all ages may be able to monitor and react appropriately to a change of state but that the younger children may struggle more with monitoring the cues for the higher-level goal compared to older children.

Furthermore, having three options of second-step actions leading to two possible types of errors can shed light on the mechanisms behind the error (that is, the error pattern can reflect the subprocess(es) of supervisory control that is/are changing across development and explain the changes in action selection patterns with age). The possibilities are that, depending on the developmental stage: (1) the change following a transition is not detected; (2) a sequence that is initiated cannot be interrupted (hence not corrected) due to a lack of inhibitory control; (3) that it can be interrupted (as a reaction to events) but the best replacement-action is not always selected, due to less good replacement-action selection process; or (4) that it both can be interrupted and replaced correctly (indicating both sufficient inhibitory control and mechanisms of replacement action).

Therefore, the patterns of errors as a function of transition type, and the type of errors within each transition will be able to support one or more hypotheses. A possibility is that the error patterns are the same type regardless of the type of rare transition, which would indicate the presence of developmental differences in the *detection or monitoring processes* between transitions (the downstream processes being unaffected). Another possibility is that error patterns differ according to the type of transition, which would indicate that the *detection or monitoring* is the same for all transitions, but that the downstream processes for *action update* are differentially affected. These processes may be inhibitory control (to interrupt the sequence) or the selection of the appropriate replacement action. There, the type of errors ('failure to update' or 'alternative action plan') may indicate whether inhibitory control or the alternative action selection processes is more important.

Finally, we hypothesised that the inhibition score would strongly, and more strongly than the switching score, predict action selection patterns and response time, indicating that improvements in inhibitory control abilities across development are underlying improvements in the appropriate recruitment of supervisory control in action sequence update.

4.3.5 Methods

4.3.5.1 Participants

Twenty-three participants took part in the study. Three of them failed to complete the entire SuperCook task and thus were excluded from the analyses. The 20 remaining participants were between 5 years 1 month and 9 years 5 months, with a median age of 6 years 10 months. The number of participant age by age is shown in table 4.3. Due to an initial technical problem in the executive function tasks, data from the first participant is missing for the switching task, and data from the first 3 participants is missing for the inhibition task.

Table 4.3: **Participants' age in the final sample. The *younger* and *older* children groups are defined as below or above the median (below 6 years 10 months).**

Age	5 y.o.	6 y.o.	7 y.o.	8 y.o.	9 y.o.	*Younger children*	*Older children*
N	6	5	5	1	3	10	10

All participants were recruited via Birkbeck's Babylab database and tested in the same lab. Written consent was obtained from the parents as well as verbal consent from the children. The experiment was conducted after receiving approval from the institutional ethics committee and conformed with the principles of the Declaration of Helsinki. Families received a thank-you gift for taking part in the study.

The study originally planned to collect 60 participants, twice the amount of participants tested in the study of Decker et al. (2016) because we were interested in comparison of children age groups while Decker et al. (2016) compared a child group to another two groups (adolescent and adults). Unfortunately, the data collection for this study had to be cut short because of the COVID-19 pandemic.

4.3.5.2 Materials

SuperCook task

The experiment materials are fully described in Section 4.1.2.

Executive Functions (EF) tasks

Switching task We developed a child-friendly version of a switching task in which both stimulus and response sets are changed simultaneously, and the classification dimension is intrinsically defined by the stimulus. An equivalent version of the task for adults (Rogers and Monsell, 1995a) requires classifying a number as odd or even (task set 1), and a letter as vowel or consonant (task set 2). Our task sets were *classifying animals as sea or land animal* and *classifying objects as sports or food-items*. Thus, the classifying dimension was given by the stimulus (animal or object). Responses were given by pressing a key on the left or right of a keyboard. Stickers were placed to indicated the two possible response keys (and the other keys were disabled). The left/right key response mapping was given by pictograms of sea/land (task set 1) or sports/food (task set 2) displayed on the left/right sides of the screen (cf. Figure 4.9). This means that, when the central icon was an animal, a sea pictogram was always presented on the right while a land pictogram was always presented on the left. When the central icon was an object, a sports pictogram was always presented on the right while a food pictogram was always presented on the left. The tasks included 6 different animals and 6 different objects.

Inhibition task The inhibition task was of the *Go-NoGo* type introduced by Casey et al. (1988). Our task was a variant of the whack-a-mole task used by Petitclerc et al. (2015), itself inspired from the task developed and made available online by Sarah Getz and the Sackler Institute for Developmental Psychobiology. We adapted slightly the number of trials and stimulus presentation time to accommodate our study's age range and constraints. The child-friendly instructions asked participants to catch the moles in a garden (*Go stimuli*) by pressing the space key when a mole was presented on screen, and to not crush the aubergines by not pressing the key for the aubergines (*NoGo stimuli*). The stimuli (the mole exemplary without disguise and the aubergine) are shown in Figure 4.10.

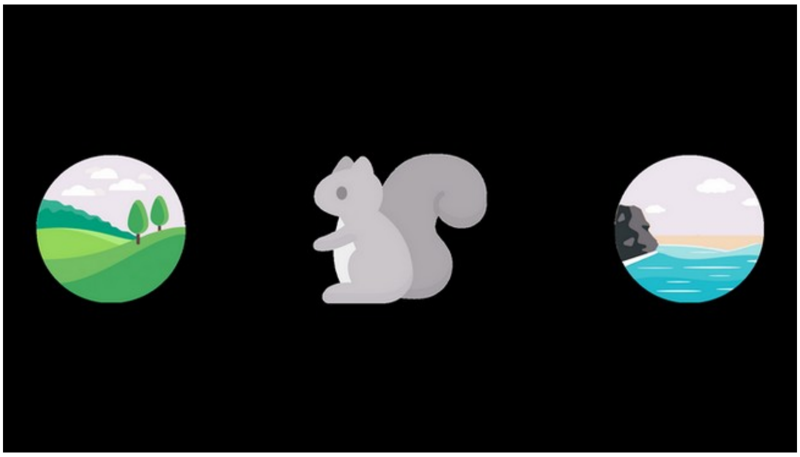


Fig. 4.9: **Example of a display of the Switching task. In this example, the task set is 'classifying the animal as a sea or land animal'. The icons on the side indicates that one should press: left for land, right for sea.**

To maximise errors on our variable of interest (false alarm rate) and thus not be at floor level, we varied the Go stimuli (the mole was disguised with different hats) as suggested by Young et al. (2018). To diminish anticipatory responding, target appearance was made unpredictable with inter-stimulus intervals randomly selected among 2000, 2100, 2200, 2300 and 2500 ms.

The stimulus presentation was initially set to 1000 ms. However when observing the first 6 participants' performance, the false alarm rate seemed consistently low hence we feared it would not capture well inter-individual variability in inhibitory control. The stimulus presentation was thus changed to 750ms (as in Abdul et al., 2017) to enhance difficulty for the rest of the participants. The age of participants was not correlated with the order of testing thus this change should not cause a systematic bias in the data, nonetheless the stimulus presentation variable is included as an independent variable in the analyses.

Fig. 4.10: **Stimuli used in the Inhibition (Go-NoGo/whack-the-mole) task. Left: go stimulus, right: no-go stimulus.**

4.3.5.3 Procedure

The SuperCook task was always administered first, always followed by the Switching task and finally the inhibitory control task.

EF Tasks

Switching task The task started with practice trials that introduced the task as well as each of the 6 possible stimuli, and ensured instructions were understood. Then a block of 10 trials of the animal task set was presented, followed by 10 trials of the objects task set (block 1 and 2; *non-switch trials*), before a block of 20 trials alternating between each task set (block 3; *switch trials*).

Inhibition task The task started with 7 practice trials, which were repeated when necessary to ensure the task was understood. To build up a prepotent response, a series of 15 Go trials was administered first. After that, four blocks of 25 trials were administered, with breaks in between each block. Each block contained 6 NoGo trials, with the sequence pseudo-randomised while ensuring that there was never two successive NoGo.

4.3.5.4 Analyses

In the SuperCook task, we analysed the non-practice trials; that is, 16 trials for the training phase, 52 trials for the test phase. We excluded the trials flagged as 'incorrect icon step attempt' as defined earlier. We also excluded trials with a timeout. The understanding of the rules was assessed during the experiment with the practice trials. When analysing the data, to ensure that the participants were playing along throughout the game (despite the repetitive aspect), we verified that each participant carried out the maximally rewarded sequence in more than 60% of the trials (chance level being at 16%) in the training phase and did not have to exclude any on that criterion.

Dependent variables (DVs)

The response time variables consist of the duration of the action selection portions of the game (referred to as 'selection portion'): for step 1, the time from the start of the trial (stimulus display) to the moment when the first icon is selected; for step 2, from the time when the first icon has been dropped at the central location to the time when the second icon is selected.

Step 3 was a terminating action, so that the main step of interest (number 2) did not have the particularity of being the first or last action of the sequence.

Training phase As a measure of performance, we defined *optimal selection* as the proportion of correct trials throughout the training phase. We coded trials as correct if the sequence was the trial's maximally rewarded sequence (i.e. correct combination of icons according to the trial's goal), and incorrect for any other combination. We split the trials of the training phase into two halves, that we call *early* and *late* part. We regressed the optimal selection against age, training part (early/late) and their interaction; unless otherwise specified the regressions we mention in this chapter follow the mixed model procedure described in the *Mixed model* Section below. The errors were classified as explained in the Section 4.1.2.2. In all analyses other than optimal selection (both in training and test phases), we sought to gain further information in control processes that were not already contained in the error patterns. We thus excluded the trials in which the sequence carried out was not the maximally rewarded one. To investigate more closely the type of planning that participants use in the task, we also analysed the RT separately for step 1 and step 2. In particular, we were interested in seeing if there were stronger practise effects on one step than the other. We regressed the response time against step number, trial and age (and all the possible interactions).

Test phase To assess the effects of the test phase's key experimental manipulation (introducing rare transitions) as a function of age, we looked at optimal selection in trials type: common, rare_state and rare_goal. There, *optimal selection* was defined as the proportion of maximally rewarded sequence among the trials that had the first correct step (which represent 70% of the trials for the lowest-performing participant). Note that the maximal reward differed according to the transition type (2 in common transitions, 1 otherwise). We will refer to the execution of the non-maximally rewarded sequence as an 'error'.

Given the small number of errors, we will look at them in median-split age groups (with the median being 6 years 10 months). To investigate the effect of rare transitions on continuous DVs (RT and kinematic variables), we look at the selection portion of the trajectory and we focus on step 2 only as the manipulation occurs in between steps 1 and 2. We regressed each variable against age and transition type. To assess the effect of the subtype of rare transitions (rare_goal and rare_state) we additionally ran each model on the subset of data with rare trials.

The kinematic variables were the three selected following the pilot analyses presented in Section 4.2.4: *avTrajAngle* (Average of the instantaneous angle between the trajectory and the most direct path to the chosen icon), *varSpeed* (Variability of the speed's norm within a trial) and *maxLogRatio* (Maximal log ratio of distance to the non-chosen icon over the distance to the chosen icon).

EF tasks The switch cost, or *EFsw score*, was computed as the difference between the mean response time on switch trials (block 3) and the mean response time on non-switch trials (block 1 and 2). It represents the additional processing cost of switching among the task sets, and taken as inverse proxy of switching abilities (a larger cost indicating less efficient switching abilities).

The DV of interest for the switch task, that we will call *EFwh score* was the false alarm rate. The false alarme rate was computed as the proportion of *NoGo* trials in which a response was given (instead of inhibited), in the phases outside the practise. To account for the differing stimulus presentation durations (750 and 1000ms) across participants, we included stimulus presentation as a continuous predictor whenever we included the EFscore. The false alarm rate was used as a proxy for inhibitory control, with higher false alarm rates indicating lower inhibitory control.

Mixed models

The main analyses consist in linear Mixed Models regressions (MMR) or logistic Mixed Models regressions (logMMR), also known as mixed-effects regression, with a random intercept per subject to account for the within-subject design (Steele, 2014). The other predictors were taken as fixed effects. For the binary DV (optimal selection), we carried out logMMR with the function *mixed_model* of the R package *GLMMadaptive*. For all the other DV, that are continuous, we first log-transformed the DV to correct for asymmetrical distributions. We then ran MMRs with the function *lme* of the R package *nlme*. When describing results, predictors are named as such: *age* for the continuous age predictor, *trans* for the type of transition (levels: common, rare_goal, rare_state), *trial* for trial number (ordinal variable), *EFwh* for the continuous EFwh score, and *EFsw* for the continuous EFsw score.

For exploratory purposes, we denote the FDR-corrected p-values that fall below 0.1 as *notable*. Such data is not to be treated as significant results, yet is noted and considered together with the rest of the evidence. Indeed, such an experiment relies on crossing information from several dependent variables that bring complementary information and can better be interpreted when considered as a whole.

Non-parametric tests

To ensure the robustness of our results, we ran additional analyses on the DV showing significant effects in the MMRs. These were non-parametric tests on the raw (not log-transformed) data summarised at the individual level with the trials' median, and discretised to assess age effects. We discretised the data with a median split, resulting in a *'younger'* group (5 years 1 months and 6 years 10 months) and an *'older'* group (6 years 10 months to 9 years 5 months). We used Wilcoxon-Mann-Whitney test when comparing age groups and Wilcoxon signed-rank test when comparing transitions. For ease of reading, the text and figures depict the untransformed variables, as well as show the age difference in median-split group. Such plots match the non-parametric tests. As mentioned, for the MMR age was entered as a continuous factor, for complementarity of the approaches and thus higher reliability of the results.

Relative importance analysis

A preliminary analysis (Pearson correlation) between age and the EF scores indicated that the Switching task score (switch cost) was not significantly correlated with age ($r = -0.353$, $p = 0.107$), but the Inhibition task score (false alarm rate) had a significant negative correlation with age ($r = -0.504$, $p = 0.024$). Hence using age and EF scores as predictors in a same regression introduces a problem of multicollinearity (see for example Tonidandel and LeBreton, 2011). Even though only the Inhibition task score was correlated with age, we treated both tasks' scores with the same procedure. When a DV was found to change significantly with age, we quantified to what extent the differences were attributable to changes in inhibitory function, switching function or by functions evolving with age other than the two. To do so, we ran relative importance analyses in complement to the MMR (we took the same approach as Magimairaj, 2017).

Relative importance analysis partitions the explained variance among correlated predictors (Tonidandel and LeBreton, 2011). It works by predicting the DV with the orthogonally transformed predictors, and standardising the regression coefficients so that we can obtain the contribution of a predictor considering a predictor alone or combined to the other predictors. We used the metric *lmg* calculated by the function *calc.relimp* of the R package *relaimpo*. The *lmg* metric yields the percentage of variance explained by each predictor so that they sum to 100%. We used the parameter *EFwh_tStim* (value = 1000 or 750) as a binary predictor to control for the different stimulus presentation times across participants. Thus, the predictors entered were age, EFwh, EFsw. To ensure convergence in the relative importance analyses regressions, we standardised all continuous predictors (EF scores and age). We performed this analysis only on the participants with entire datasets.

4.3.6 Results

4.3.6.1 Training phase

Optimal selection

The average optimal selection (proportion of trials with the maximally rewarded sequence) during the training phase was 0.89. The lowest individual average optimal selection was 0.69. The details are presented in table 4.4.

Table 4.4: **Optimal selection rate in the training phase, for all children combined and by age group.**

	All	Younger children	Older children
mean	0.887	0.9	0.875
std	0.094	0.099	0.093
min	0.687	0.687	0.687
max	1	1	1

We computed the proportion in the first or second half of training (predictor *half*) in order to look at the evolution of optimal selection across training. The average optimal selection did not significantly vary with age or with the training half, as detailed in table 4.5.

***Table 4.5:* Results of the logistic Mixed Model Regression on Optimal selection with the predictors age and half of the training phase (*early* for the first 8 trials, *late* for the last 8 trials). The baseline for *half of training* was the *early*-trials' half, compared against *late*-trials' half.**

Predictor	β	CI (l)	CI (u)	SE	p	p (FDR)	sig
Intercept	0.87	0.57	1.17	0.14	<0.001	<0.001	***
age	0.00	-0.04	0.05	0.02	0.861	0.861	
half(late)	0.18	-0.19	0.55	0.18	0.333	0.444	
age:half(late)	-0.03	-0.08	0.03	0.02	0.292	0.584	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p: (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR p-values: ***: $p < 0.001$.

RTs

For all Mixed Model Regressions, we report the FDR-corrected p-values. When looking at the RT as a function of age, step and half of training trials (Figure 4.11, table 4.7), there was a non-significant but notable main effect of age where RT decreased with age ($p = 0.064$). There was a non-significant but notable interaction trial x step ($p = 0.061$), as well as a non-significant but notable interaction age x step x trial ($p = 0.090$).

From Figure 4.11, it looks like for younger children, the RT in step 1 is larger early than late in training, suggesting that practise speeds up step 1, and that this effect of practise is absent in other steps and in older children. We tested this pairwise comparison with a non-parametric test. When comparing the RT at each step between the halves of training, none of the pairwise comparison between the first and second half were significant (table 4.6).

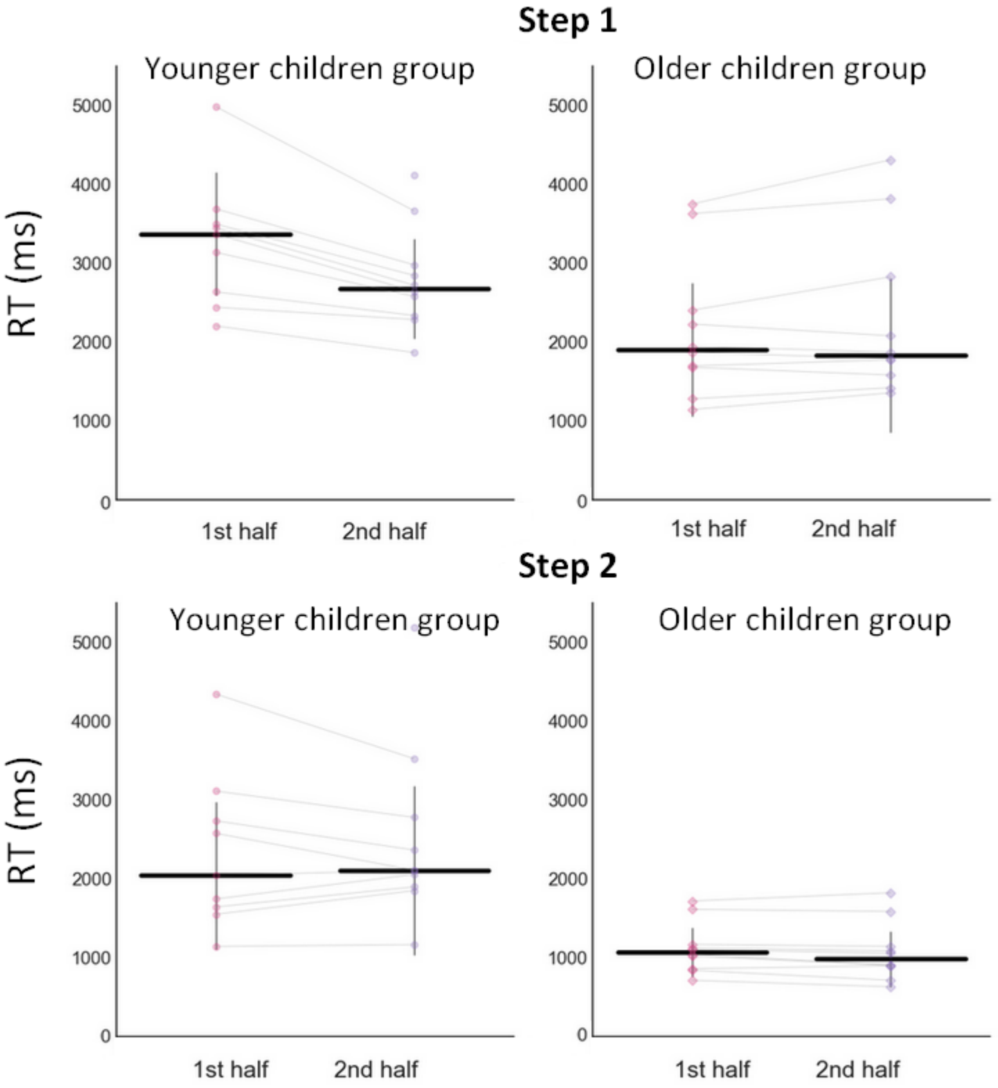


Fig. 4.11: **Response time (RT) as a function of step, part of the training (first or second half of trials) and age (left panel: younger children, right panel: older children). A pair of individual dots joined by a line represents the data of one participant.**

Table 4.6: **Wilcoxon signed-ranks test comparing the RT on the first and second half of trials (excluding practice trials) of the training phase. The figures in the table are: Statistic (sample size with non-zero difference), p-value.**

	Younger	Older
Step 1	W(10)=3, p=0.156	W(10)=30, p=0.173
Step 2	W(10)=6, p=0.438	W(10)=48, p=0.808

Table 4.7: Linear Mixed Model Regression on RT in the training phase with the predictors age, trial and step number. The baseline for step was the *step 1* compared against *step* 2.

Predictor	β	CI (l)	CI (u)	SE	p	p (FDR)	sig
Intercept	9.03	8.20	9.87	0.43	<0.001	<0.001	***
trial	-0.02	-0.07	0.03	0.03	0.421	0.561	
age	-0.16	-0.29	-0.03	0.06	0.016	0.064	.
step(1)	-0.80	-1.72	0.12	0.47	0.088	0.141	
trial:age	0.00	-0.01	0.01	0.00	0.517	0.591	
age:step	0.02	-0.11	0.15	0.07	0.726	0.726	
trial:step	0.09	0.01	0.16	0.04	0.023	0.061	.
trial:age:step	-0.01	-0.02	-0.00	0.01	0.045	0.090	.

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: FDR (false discovery rate)-adjusted p-value, *sig.*: significance levels based on FDR p-values: .: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

4.3.6.2 Test phase

Optimal selection

The overall optimal selection is shown on the estimated marginal plots in Figure 4.12. As clearly visible in Figure 4.12, the optimal selection at common transitions was quite even across ages, while the optimal selection in rare_goal was near the level of common transitions for the older children but decreased with decreasing age. Consistent with this, the Logistic Mixed Model (table 4.8) showed a main effect of rare transitions ($p < 0.001$) and a significant age-by-transition interaction ($p < 0.001$). In other words, younger children made more errors at rare_goal than in common transitions contrary to older children.

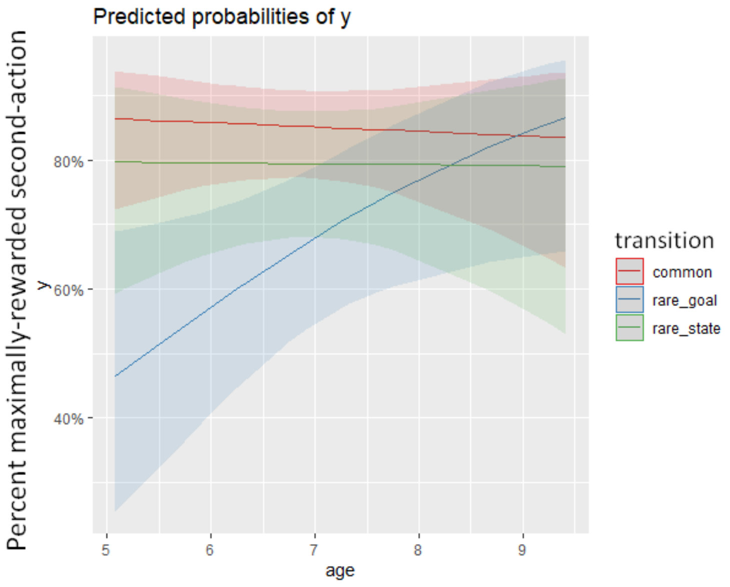


Fig. 4.12: Logistic regression's predicted probability of maximally-rewarded second choice according to age and type of transition. The shaded area represents the 95% confidence intervals.

Table 4.8: **Results of the logistic Mixed Model Regression on Optimal selection at test phase, with age and transition as predictors. The baseline for transition was *common* transitions.**

Predictor	β	SE	z-value	p	p (FDR)	sig
Intercept	2.11	1.38	1.53	0.1251	0.25	
age	-0.05	0.19	-0.27	0.7881	0.788	
trans(rare_goal)	-4.60	1.01	-4.34	<0.0001	<0.001	***
trans(rare_state)	-0.69	1.07	-0.65	0.5150	0.772	
age:trans(rare_goal)	0.51	0.15	3.54	0.0003	0.001	**
age:trans(rare_state)	0.04	1.45	0.29	0.7750	0.930	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: FDR (false discovery rate)-adjusted p-value, *sig.*: significance levels based on FDR p-values: .: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Error types

The types of choice (including correct choices, which shows optimal selection within the trials with first correct choice) as a function of transition type is shown in Figure 4.13. As mentioned earlier and shown in the central panel, the errors in common transitions were infrequent regardless of age, so we will not investigate them further. When looking at the types of errors made in rare transitions, it seems like the most common error is the *'failing to update'* error, that is the one where the icon selected is the one that would have led to the maximal reward if there had not been a rare transition regardless of the subtype rare_goal or rare_state (grape, purple dots on the figure). A Wilcoxon signed-rank test on this data (N=20) revealed that indeed the percent of *failing to update* error was higher than the other type of error (*'activating the alternative prepotent plan'*): W(20)=14.0, $p < 0.001$. The results still held when separated by age group (younger children: W(10)=1.0, $p < 0.05$, younger children: W(10)=1.0, $p < 0.01$), although statistics on small sample size (N=10 in each group) should be taken with caution.

RT

For *RT*, when looking at the entire dataset (with common transitions as baseline), there was no significant effect of group or conditions nor interaction between these variables (Figure 4.14 and table 4.9). When looking at the rare transitions only, there was a main effect of transition (p=0.010), a main effect of age (p=0.007) and a significant interaction age x transition (p=0.040) on RT.

When looking at non-parametric tests for the interaction age x transition (table 4.10), younger children had significantly longer RT at rare_goal than rare_state transitions (p=0.0371), but not older children (p=0.8457).

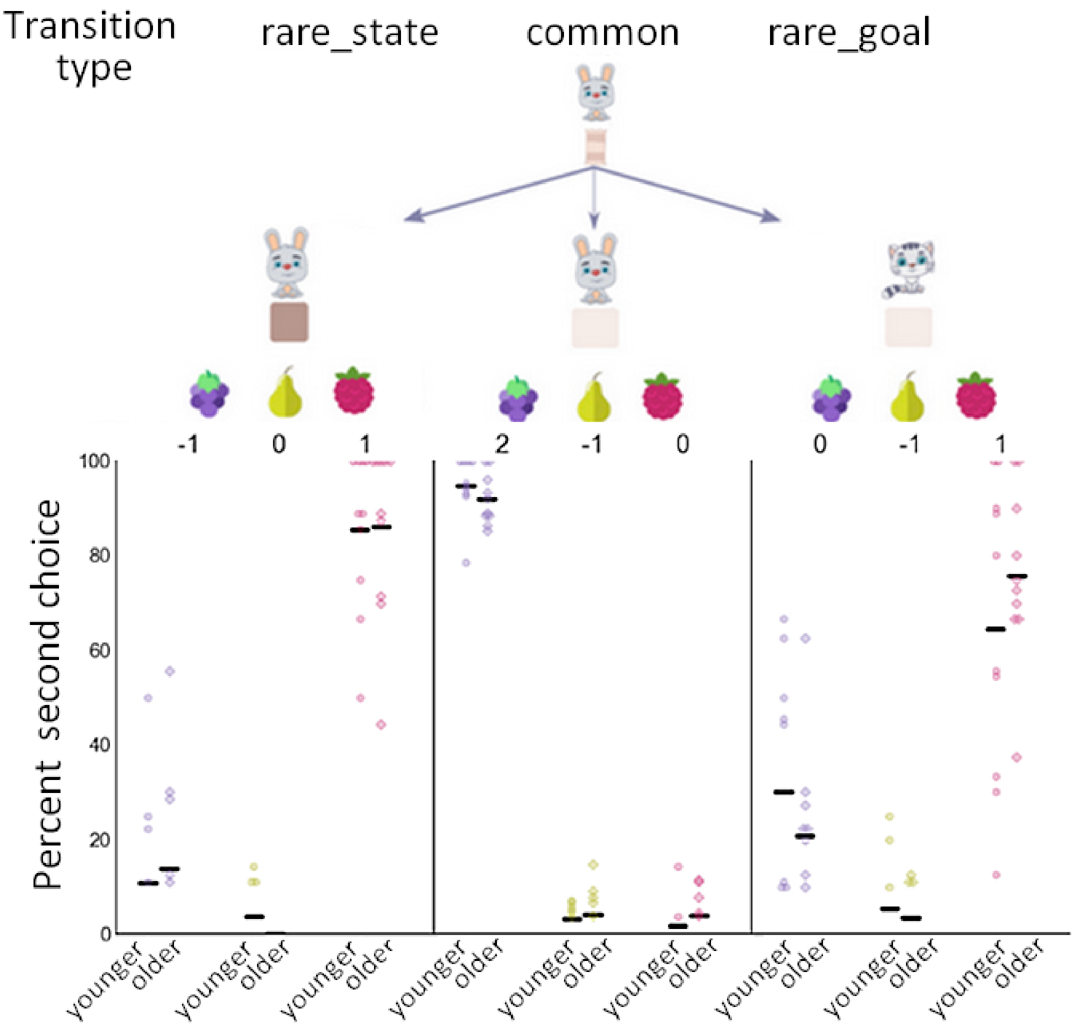


Fig. 4.13: **Percent type of choices at Step 2 following a correct choice at Step 1, as a function of age and transition type. The dots represent the data of individual participants.**

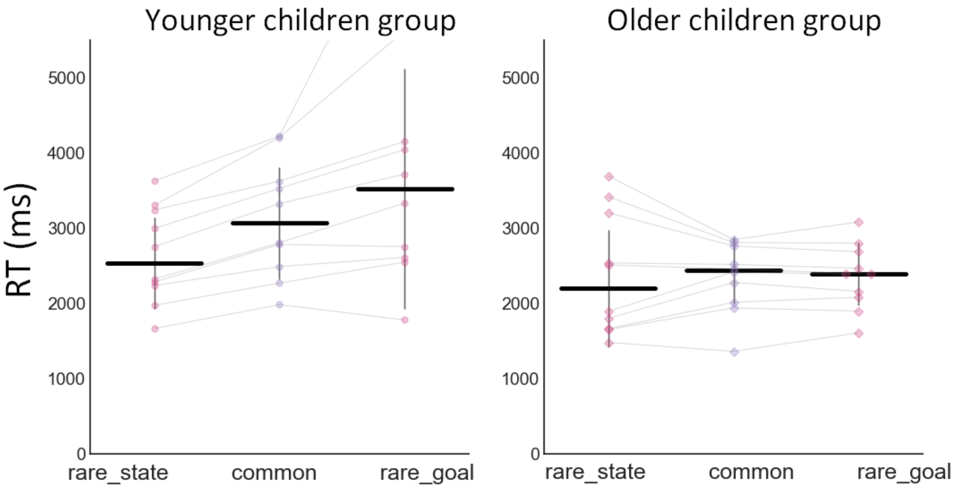


Fig. 4.14: **RT as a function of transition type (common, rare_state, rare_goal) and age group (left panel: younger children, right panel: older children). The horizontal thick bar represents the median, the vertical light bar represents the standard deviation and the dots represent individual participants.**

***Table 4.9:* Linear Mixed Model Regression on RT at test phase, with age and transition as predictors. Table (a): Entire dataset, with baseline for transition *common* transitions. Table (b): Only rare transitions are included, with transition *rare_goal* transitions as baseline.**

Predictor	β	CI (l)	CI (u)	SE	p	p (FDR)	sig
Table (a)							
Intercept	8.41	7.90	8.92	0.26	<0.001	<0.001	***
transition:rare_goal	0.39	-0.01	0.79	0.20	0.054	0.108	
transition(rare_state)	-0.23	-0.62	0.16	0.20	0.254	0.305	
age	-0.08	-0.16	-0.00	0.04	0.045	0.135	
transition(rare_goal):age	-0.04	-0.09	0.02	0.03	0.198	0.297	
transition(rare_state):age	0.02	-0.03	0.08	0.03	0.396	0.396	
Table (b)							
Intercept	8.82	8.29	9.34	0.27	<0.001	<0.001	***
transition(rare_state)	-0.64	-1.08	-0.20	0.22	0.005	0.010	*
age	-0.12	-0.20	-0.04	0.04	0.005	0.007	**
transition(rare_state):age	0.06	0.00	0.12	0.03	0.04	0.040	*

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p*: *(FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values: .: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

***Table 4.10:* Wilcoxon signed-rank test comparing the RT on transitions in the test phase across age groups (median-split). The figures in the table are: W-value, (N of pairs with non-zero differences), p-value.**

	Younger	Older	All ages
Common vs. rare_state	W(10)=9, p=0.0645	W(10)=26, p=0.9219	W(20)=75, p=0.2774
Common vs. rare_goal	W(10)=16, p=0.2754	W(10)=26, p=0.9219	W(20)=83, p=0.4304
rare_state vs. rare_goal	W(10)=7, p=0.0371	W(10)=25, p=0.8457	W(20)=59, p=0.0897

Kinematics

There were no significant effect of age, of transition or of age-by-transition interaction in avTrajAngle (Figure 4.15 and table 4.11) and maxLogRatio (Figure 4.16 and table 4.12). Thus, we will not discuss these results further.

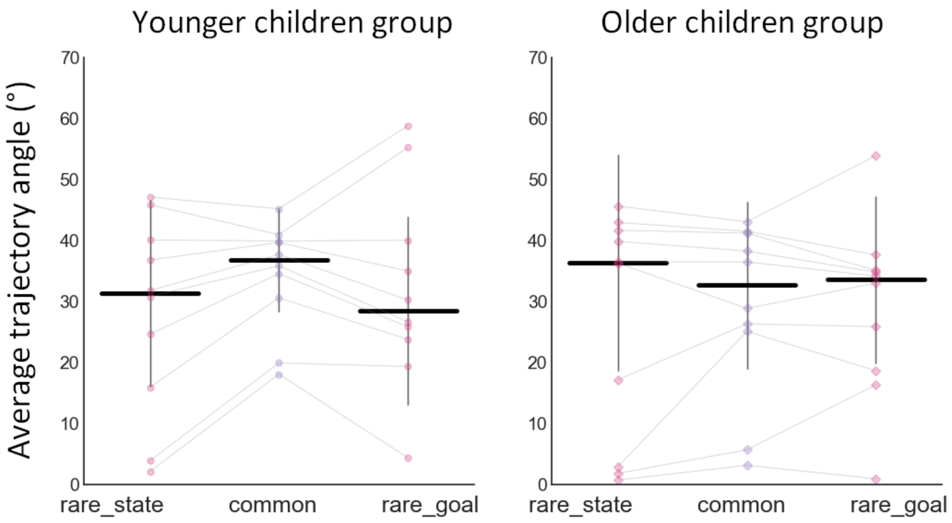


Fig. 4.15: **Trajectory angle (avTrajAngle) as a function of transition type (common, rare_state, rare_goal) and age group (left: younger children, right: older children). The median is represented with standard deviation and the dots represent individual participants.**

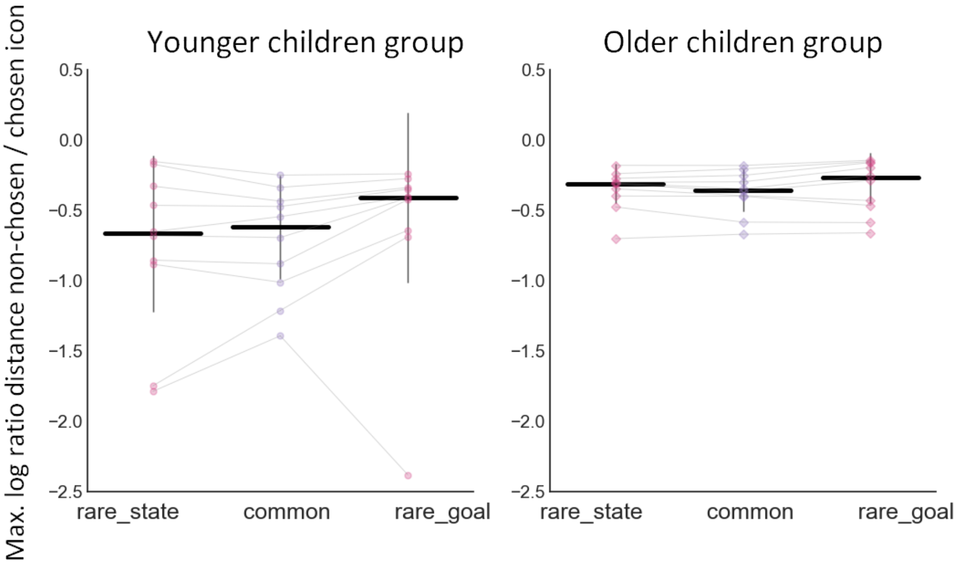


Fig. 4.16: **Maximal log ratio of distance to the non-chosen icon over the distance to the chosen icon (maxLogRatio) as a function of transition type (common, rare_state, rare_goal) and age group (left: younger children, right: older children). The median is represented with standard deviation and the dots represent individual participants.**

***Table 4.11:* Linear Mixed Model Regression on the kinematic feature *avTajAngle* (cf. table 4.1) at test phase, with age and transition as predictors. Table (a): All the data is included, with *common* transitions as baseline. Table (b): Only rare transitions' data are included, with *rare_goal* transitions as baseline.**

Predictor	β	CI (l)	CI (u)	SE	p	p (FDR)	sig
Table (a)							
Intercept	5.12	3.44	6.79	0.85	<0.001	<0.001	***
transition(rare_goal)	-0.51	-2.22	1.19	0.87	0.556	1.112	
transition(rare_state)	-0.40	-2.09	1.29	0.86	0.643	0.965	
age	-0.30	-0.56	-0.05	0.12	0.022	0.066	.
transition(rare_goal):age	0.05	-0.18	0.29	0.12	0.666	0.799	
transition(rare_state):age	0.01	-0.23	0.25	0.12	0.955	0.955	
Table (b)							
Intercept	4.64	2.40	6.89	1.14	<0.001	<0.001	***
transition(rare_state)	0.10	-1.45	1.65	0.79	0.9	0.900	
age	-0.25	-0.59	0.08	0.16	0.133	0.266	
transition(rare_state):age	-0.04	-0.25	0.17	0.11	0.711	0.948	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, p: unadjusted p-value, p: *(FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values: .: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

***Table 4.12:* Linear Mixed Model Regression on the kinematic feature *maxLogRatio* (cf. table 4.1) at test phase, with age and transition as predictors. Table (a): All the data is included, with *common* transitions as baseline. Table (b): Only rare transitions' data are included, with *rare_goal* transitions as baseline.**

Predictor	β	CI (l)	CI (u)	SE	p	p (FDR)	sig
Table (a)							
Intercept	-1.37	-2.34	-0.39	0.50	0.006	0.036	*
transition(rare_goal)	-0.06	-0.63	0.52	0.29	0.846	0.846	
transition(rare_state)	-0.32	-0.89	0.25	0.29	0.267	0.401	
age	0.10	-0.05	0.25	0.07	0.168	0.504	
transition(rare_goal):age	0.02	-0.06	0.10	0.04	0.641	0.769	
transition(rare_state):age	0.05	-0.03	0.13	0.04	0.223	0.446	
Table (b)							
Intercept	-1.38	-2.47	-0.29	0.55	0.013	0.052	.
transition(rare_state)	-0.34	-1.04	0.35	0.35	0.331	0.441	
age	0.12	-0.05	0.28	0.08	0.150	0.300	
transition(rare_state):age	0.04	-0.06	0.14	0.05	0.421	0.421	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, p: unadjusted p-value, p: *(FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values: .: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

For *varSpeed* (Figure 4.17 and table 4.13), when looking at the entire dataset, there was a significant effect of age (<0.001) and of rare_state transition (p<0.001). When looking at the rare transitions, there was also a significant effect of age (p<0.001), a non-significant but notable effect of transition (p=0.056) and a non-significant but notable age-by-transition interaction (p=0.059) on varSpeed.

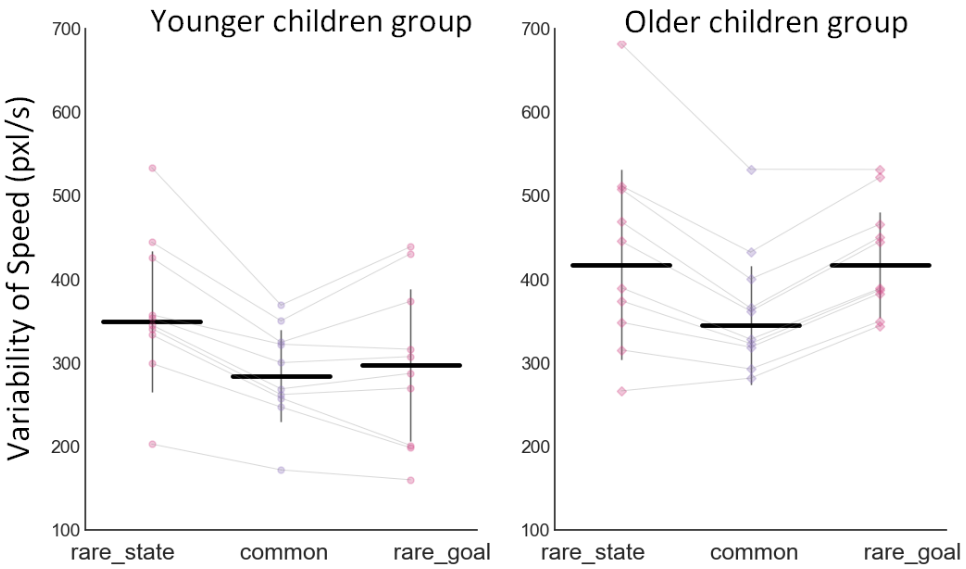


***Fig. 4.17:* Speed variability (varSpeed) as a function of transition type (common, rare_state, rare_goal) and age group. The figures shows the median, std and individual participants.**

***Table 4.13:* LMM Regression on the kinematic feature *varSpeed* (cf. table 4.1) at test phase. (a): All the data is included, with *common* transitions as baseline. (b): Only rare transitions are included, with *rare_goal* transitions as baseline.**

Predictor	β	CI (l)	CI (u)	SE	p	p (FDR)	sig
Table (a)							
Intercept	5.65	5.597	5.69	0.024	<0.0001	<0.001	***
transition(rare_goal)	0.05	-0.045	1.63	0.053	0.271	0.271	
transition(rare_state)	0.19	0.09	0.29	0.051	0.0002	<0.001	***
age	0.19	-0.29	0.39	0.03	<0.0001	<0.001	***
transition(rare_goal):age	0.08	-0.06	0.23	0.072	0.261	0.313	
transition(rare_state):age	-0.09	-0.23	0.05	0.071	0.213	0.319	
Table (b)							
Intercept	5.7	5.61	5.79	0.047	<0.0001	<0.001	***
transition(rare_state)	0.13	0.005	0.26	0.065	0.052	0.056	.
age	0.28	0.15	0.41	0.064	<0.0001	<0.001	***
transition(rare_state):age	-0.17	-0.034	0.006	0.089	0.059	0.059	.

β: Regression coefficient, *CI (l)*: Confidence Interval (lower bound), *CI (u)*: CI (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p: (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values: .: $p < 0.1$, ***: $p < 0.001$.

When looking at non-parametric tests for the age-by-transition interaction in the entire dataset (table 4.14), the older group had significantly higher varSpeed at rare_goal than common (p=0.0273), but not the younger group (p=0.4316). When looking at the non-parametric tests for the interaction between age and rare transitions, the younger children had a significantly longer varSpeed at rare_state than rare_goal (p=0.0137) but not the older (p=0.9219). Finally, the significant effect of rare state was corroborated by the non-parametric test, which showed a significant difference between common and rare_state transitions in age groups combined (p<.0001).

***Table 4.14:* Wilcoxon signed-rank test comparing the *varSpeed* on transitions in the test phase across age groups (median-split). The figures in the table are: W-value, (N of pairs with non-zero differences), p-value.**

	Younger children	Older children	All ages
Common vs. rare_state	W(10)=0, p=0.002	W(10)=7, p=0.037	W(20)=7, p<.001
Common vs. rare_goal	W(10)=19, p=0.431	W(10)=6, p=0.027	W(20)=44, p=0.021
rare_state vs. rare_goal	W(10)=4, p=0.013	W(10)=26, p=0.921	W(20)=61, p=0.105

Given the significance results in varSpeed, for better understanding of the feature we illustrate the trajectories and speed profiles of trials as a function of the value of varSpeed (Figure 4.18). Visual inspection of the bottom panels of Figure 4.18 shows that both the range of values of speed encountered throughout the trial, and the variability within this range, are characteristics of the task's trials that influence the metric. For example, on the left panel the speed profiles had both very slow and very fast portions, with about 3 notable peaks, resulting in high varSpeed. The middle panel show relatively smoother speed profiles, with a medium range of speed values and mostly two phases (a slow phase followed by a faster phase), resulting in smaller varSpeed than in the left panel, but higher than in the right panel. The right panel shows speed profiles each with a relatively small range of values, resulting in smaller varSpeed values.

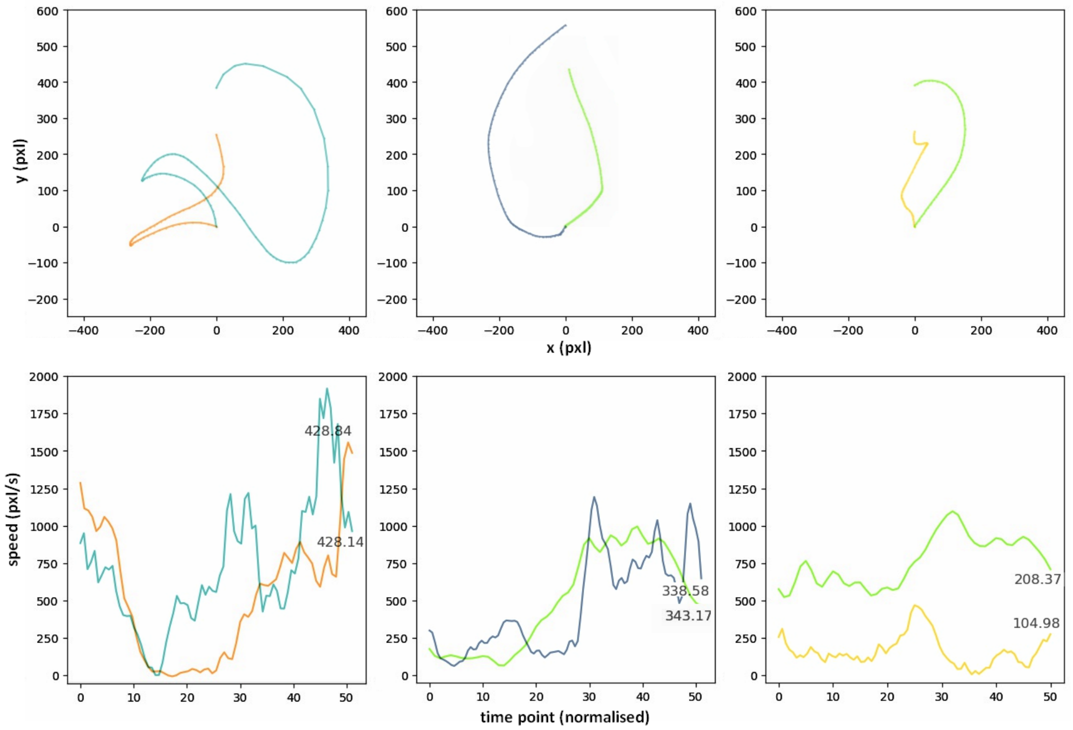


Fig. 4.18: **Trajectories (top panels) and corresponding speed profile (bottom panels) of the second-step selection portion (from the centre of the screen until the selection of a peripheral icon; trajectories were rotated so that they end at the top centre). A few trials are taken to illustrate the visual aspect of trials with relatively high (left), average (middle) and low (bottom) values of varSpeed. The figure indicated at the end of a speed curve indicates the value of varSpeed for that trial. The colours identify the corresponding trials on a pair of top and bottom panel.**

4.3.6.3 EF tasks

Switching task

The switch cost was not significantly correlated with age, but the direction of the relationship was as expected: the older the children, the smaller the switch cost ($r = -0.353$, $p = 0.107$; cf. Figure 4.19).

Inhibition task

As mentioned previously, the stimulus presentation duration in the inhibition task was 1000ms for the first 6 participants and was later adjusted to 750ms for the 14 remaining participants. If we include the entire dataset regardless of the stimulus presentation parameter (20 participants), there was a significant negative correlation between the age and False alarm rate ($r = -0.504$, $p = 0.024$; cf. Figure 4.20). If we exclude the 6 participants who had the unadjusted stimulus presentation time, the negative correlation between the age and False alarm rate remains significant ($r = -0.634$, $p = 0.015$; cf. Figure 4.20, blue dots).

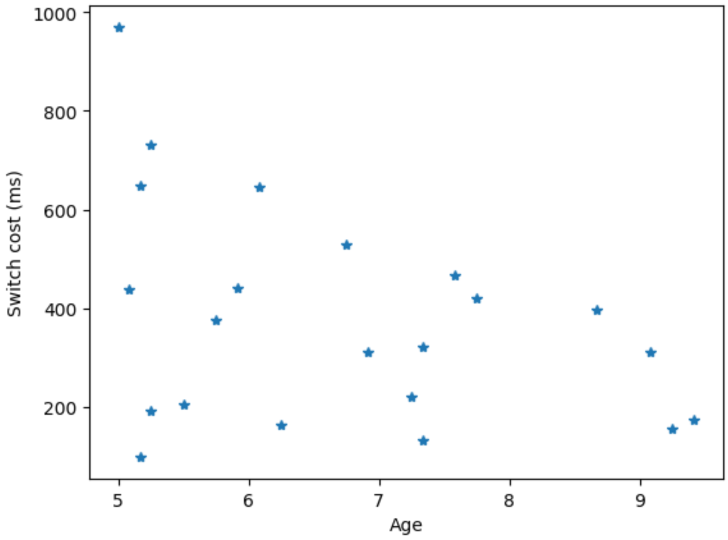


Fig. 4.19: **Scatter plot: switch cost (in the switching task) against age.**

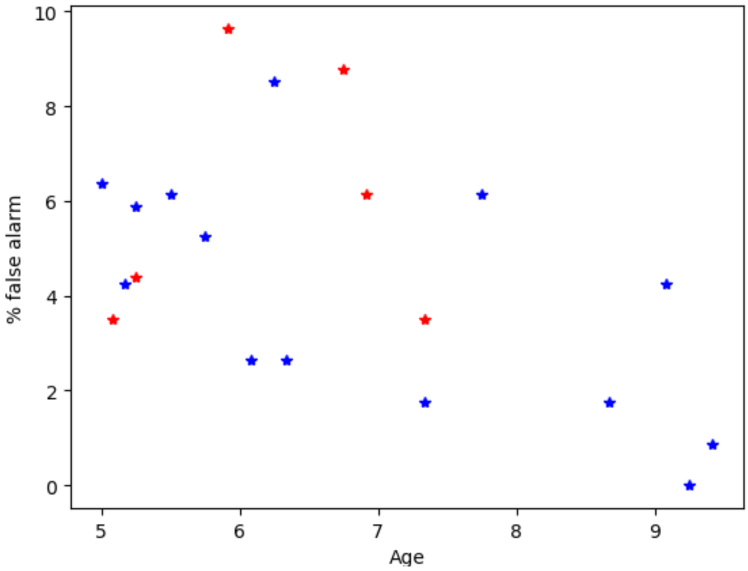


Fig. 4.20: **Scatter plot: false alarm rate (in the inhibition task) against age. The red dots indicate the participants which had a stimulus presentation time of 1000ms, the others 750ms.**

Relative importance analysis (relimp)

Table 4.15 presents the full results of the relimp regression on optimal selection, while table 4.16 presents the full results of the relimp regression on RT. We focus on the normalised relative contribution of age, EFsw and EFwh scores. Their values were **0.044** for *age*, **0.074** for *EFsw* and **0.882** for *EFwh*. For RT, they were **0.825** for *age*, **0.074** for *EFsw* and **0.101** for *EFwh*.

The results indicate an overwhelmingly higher contribution of age to the variability in RT. A drawback is that this analysis does not fully separate the typical ('baseline') changes in RT observed with age from the specific effect of transitions on RT. When looking at RT as a function of transition, we had found that rare_state

Table 4.15: **Relative importance analysis (at test phase): results of the underlying logistic Mixed Model Regression on Optimal selection with the predictors age, trans EFsw (switch score in the switching task), EFwh (false alarm score in the inhibition task) and EFwh_tStim (stimulus presentation time in the inhibition task). The baseline for transition was *common* transitions.**

Predictor	β	CI (l)	CI (u)	SE	z-value
Intercept	1.59	0.91	2.35	0.34	4.72
age	-0.12	-0.96	0.68	0.37	-0.32
EFsw	0.26	-0.49	1.08	0.36	0.74
EFwh	-0.59	-1.53	0.26	0.41	-1.45
EFwh_tStim	-0.06	-0.87	0.75	0.37	-0.17
trans(rare_goal)	-0.82	-1.28	-0.36	0.23	-3.54
trans(rare_state)	-0.55	-1.04	-0.05	0.25	-2.17

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower), *CI (u)*: Confidence Interval (upper), *SE*: standard error.

Table 4.16: **Relative importance analysis (at test phase): results of the underlying linear Mixed Model Regression on RT with the predictors age, trans EFsw (switch score in the switching task), EFwh (false alarm score in the inhibition task) and EFwh_tStim (stimulus presentation time in the inhibition task). The baseline for transition was *common* transitions.**

Predictor	β	CI (l)	CI (u)	SE	t-value
Intercept	7.85	7.71	7.99	0.08	93.44
age	-0.13	-0.29	0.03	0.10	-1.36
EFsw	0.00	-0.15	0.16	0.10	0.02
EFwh	0.08	-0.09	0.26	0.11	0.79
EFwh_tStim	0.04	-0.12	0.19	0.10	0.37
trans(rare_goal)	0.19	0.09	0.30	0.05	3.67
trans(rare_state)	-0.00	-0.11	0.10	0.05	-0.05

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower), *CI (u)*: Confidence Interval (upper), *SE*: standard error.

and/or rare_goal varied significantly from common transitions within given age groups. Thus, we corrected for the baseline difference in RT with age, by computing the scores *Delta RT Common − rare_state* and *Delta RT Common − rare_goal* and by running the same relative importance analysis (table 4.17 and table 4.18) on these difference scores (the only difference being that we computed the median per participant prior to computing the *Delta* because there were fewer trials in the rare than common transition conditions by design.)

For *Delta RT Common − rare_state*, the normalised relative contribution values were **0.715** for *age*, **0.065** for *EFsw* and **0.219** for *EFwh*.

For *DeltaCommon − rare_goal*, these were **0.767** for *age*, **0.175** for *EFsw* and **0.058** for *EFwh*.

***Table 4.17:* Relative importance analysis (at test phase): results of the underlying linear Mixed Model Regression on *Delta RT Common − rare_state* with the predictors age, trans EFsw (switch score in the switching task), EFwh (false alarm score in the inhibition task) and EFwh_tStim (stimulus presentation time in the inhibition task). The baseline for transition was *common* transitions.**

Predictor	β	CI (lower)	CI (upper)	SE	t-value
Intercept	0.06	-0.22	0.34	0.12	0.50
age	0.10	-0.25	0.45	0.15	0.67
EFsw	0.03	-0.29	0.35	0.13	0.22
EFwh	0.06	-0.32	0.44	0.16	0.38
EFwh_tStim	0.00	-0.35	0.35	0.15	0.00

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower), *CI (u)*: Confidence Interval (upper), *SE*: standard error.

***Table 4.18:* Relative importance analysis (at test phase): results of the underlying linear Mixed Model Regression on *Delta RT Common − rare_goal* with the predictors age, trans EFsw (switch score in the switching task), EFwh (false alarm score in the inhibition task) and EFwh_tStim (stimulus presentation time in the inhibition task). The baseline for transition was *common* transitions.**

Predictor	β	CI (lower)	CI (upper)	SE	t-value
Intercept	-0.05	-0.22	0.13	0.07	-0.66
age	0.10	-0.12	0.32	0.09	1.06
EFsw	0.01	-0.19	0.21	0.08	0.07
EFwh	-0.08	-0.31	0.16	0.10	-0.77
EFwh_tStim	-0.19	-0.40	0.03	0.09	-2.02

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower), *CI (u)*: Confidence Interval (upper), *SE*: standard error.

4.3.7 Discussion

Children aged 5 to 9 years old successfully understood, learnt and executed the task. The touchscreen was well-suited for this age range and the game engaging enough so that the vast majority of participants finished the task.

4.3.7.1 Sequence acquisition

There was no significant improvement in optimal selection from the first to the second half of trials. Contrary to our predictions, there was no significant improvement in response time either. It looked as if there was a reduction in response time at step 1 in the younger children's group, but this result was not significant nor notable. Thus, it seems like the task in the training phase was carried out with a good optimal selection proportion and at an even speed from the beginning. The lack of change in optimal selection is not surprising given there were practise trials verifying that the task was understood and given that this phase was easy (compared to the test phase).

4.3.7.2 Monitoring and access to a change of state or a distal cue

The optimal selection in the test phase was equally good for both age groups in common transitions (which corresponded to trials in the training phase). More errors were performed at rare transitions and the error pattern varied with age. Notably, younger children made particularly more errors at transitions with a change of goal (rare_goal) than at transitions with a change of state (rare_state). However, older children did not have significant differences across the three types of transitions.

The response time patterns partly mirror the optimal selection patterns. The younger children had a time cost when selecting actions following rare transitions where there is a change of goal compared to the rare transitions where there is a change of state. This was not the case later in the older children. This pattern was consistent with the optimal selection data in rare transitions, where younger children but not older children had a lower optimal selection at rare_goal transitions than at rare_state transition.

This indicates an increase in difficulty for the younger children to react following the change of goal compared to the change of state. The difference in difficulty may arise from different difficulties in *detection* of the change, or after unaffected detection, from different difficulties in *updating* the action following the change, or a mixture of both. While we controlled at best for the difference in salience between the goal cue (character displayed) and the state cue (bowl colour)- by keeping both in central locations, and controlling for the change in contrast-, we cannot entirely rule out the possibility that the difference is purely perceptual. However, it is plausible that differences in difficulty are attributable not to difference of salience in cues due to their perceptual properties but instead to how much children are monitoring one or the other. We propose that children were monitoring the bowl's change of colour more carefully than the goal cue, because the bowl change of state is a direct consequence of the action they are performing. This difficulty, manifested in longer RT, to monitor the state cue compared to the goal cue was found in younger children but not older children. This is consistent with prior evidence that through preschool, children become less driven to perceptual cues and more capable of relating their action sequence to a more distal goal with age (Freier et al., 2017), if we assume that this developmental trajectory continue beyond preschool.

4.3.7.3 Inhibition of an ongoing sequence and replacement of action

The types of errors made tell us more about the mechanisms behind the lower optimal selection. First we note that children in our sample made both type of errors. Each error type reflects different underlying mechanisms behind the errors. For the errors that we coined *'failing to update'*, we assumed that in the test phase when children have practised the task extensively, they form a sequence action planned ahead (that is, the appropriate fruit icon 'select action' is already activated before step 1 is completed), and that a reason for error can be that a (partially-)prepared action fails to be replaced by the accurate one.

Although our data is not clear to that respect, the younger children might have a reduction in time with practice at step 1, but not at step 2, during the training phase. If that finding was replicated, it would strongly suggest that, indeed, the planning of the two steps occurs earlier than the time of selecting action at step 2 (otherwise we would observe a decrease of the selection time at step 2 with practice). What we call 'sequence action plan' or schema bears similarities with the 'motor chunks' in Sternberg et al. (1988), who showed that for rapid sequences of motor responses, participants compile several actions into a motor chunk before performing the first action. Thus, within a motor chunk, actions are performed one after another rapidly, and we assume that here this does not always leave sufficient time for updating the action following rare transitions. However, one could argue that the strategy of chunking, or pre-planning of the two actions at the beginning of the sequence would occur only during training phase, and that the introduction of rare transitions in the test phase would 'reset' this strategy. We argue that such a 'reset' in strategy is unlikely, because it would imply a reorganisation of motor chunk after an extensive training phase, and the frequency of rare transitions may be too low for this to occur.

Second, regardless of the type of rare transition and age, *'failing to update'* errors were significantly more frequent. This suggests that most of the time, the presumed action plan was not updated. This might be due to a failure of monitoring the cue (change of state of the bowl, or change of the character representing the goal). Alternatively, this might follow the successful monitoring and detection of the change, but a failure to interrupt an ongoing sequence. If that is the case, we propose that inhibition is the key limiting factor in failing to update the sequence. The important contribution of the inhibition score of the separate inhibition task (compared to age and the switching cost) on the performance in the SuperCook task is consistent with a role of inhibition, albeit not sufficient to conclude as will be discussed in next section.

Even if they were less frequent, the other errors should also be explained. This second error type was coined *'activating the alternative prepotent plan'*, because we assume the two sequences practised in the training phase (sequence AA and BB) have reached a certain level of prepotency when later executed at the test phase. Let us take the example of the goal where the initially appropriate sequence is AA. At the time of a rare transition, there might be awareness of an update of plan being needed that leaves the place for a replacement action (in other words, action A previously planned is no longer needed), but that does not appropriately select the alternative action (which would have been action C), leaving the place to the other very-well practised action (instead of action C, action B from the very-well practised BB). In terms of competitions of schemas (Schmidt, 1975), it might be that the schema A is appropriately inhibited, but the action C has too little activation compared to the prepotent schema B (either or both linked to insufficient activation of schema C or over-activation of schema *B*).

A limitation of the present findings is that the errors were relatively few. Together with a small sample size, this means that we can not draw clear conclusion from the error data alone.

4.3.7.4 Role of executive functions

As mentioned, it is possible that the 'failure to update' type of error at the second-stage action arises from a failure to inhibit a pre-planned action in the middle of the sequence.

When looking at the relative contribution of age and executive function task scores on optimal selection in the SuperCook task, the inhibition score shows an overwhelmingly high contribution (more than 85%). This suggests that development in inhibitory control drives improvement in performing an accurate sequence. However, when predicting RT, age was the largest predictor, even when using composite scores of RT that tried to take into account individual baseline speed. We would have expected the inhibition score to predict RT to a larger extent. It might be that inhibitory control plays a role in optimal selection but is not involved at all in speed of the processes underlying change in action control following rare transitions. It might also be that differences with age in the speed of update of action plans are the largest and thus overtake other differences, consequently hiding the involvement of differences in inhibitory control. Alternatively, it might be that our sample size is too small to reliability assess and correlate inter-individual differences in EF to inter-individual differences in action control.

Regarding the switching task, we first note that our novel child-friendly version of the task was validated since we found, as expected, switch costs far from zero in all participants. Our switching score did not correlate with age nor contributed much to optimal selection or RT variables. The absence of even a small but significant correlation between switching score with age is surprising given findings on switching improvement with age in this age range (e.g., Crone et al., 2006). It is possible that switching does not contribute, or contributes much less than inhibition, to action control in the task, but this would require more evidence. However, we note that Yanaoka and Saito (2019) found very similar results where inhibition was strongly correlated with the control of script execution and not switching.

It is not possible at this stage to draw conclusions about the role of executive functions in action control, but we will rediscuss the evidence together with the study in chapter 5.

4.4 General Discussion

Experiment 1 consisted of a pilot task and it helped pre-selecting kinematic features that, based on the pilot data, would capture a large amount of variance in experiment 2. Experiment 2 studied action control in school-aged children using an improved version of the task, and used the features thanks to experiment 1. Here we discus the interpretation of that feature in the context of the development of action control and how it relates to the other findings.

4.4.1 Kinematics feature selection

Our kinematic feature selection (experiment 1, Section 4.2) allowed us to use 3 kinematic dependent variables, that were defined and selected a-priori and that were likely to capture variance in the task. By this method, experiment 1 identified three candidate kinematic features: *avTrajAngle* (Average of the instantaneous angle between the trajectory and the most direct path to the chosen icon), *varSpeed* (Variability of the speed's norm within a trial) and *maxLogRatio* (Maximal log ratio of distance to the non-chosen icon over the distance to the chosen icon). In experiment 2, *varSpeed* proved informative, as will be described in Section 4.4.3 .

4.4.2 Error patterns and developmental changes in monitoring and inhibition

In experiment 2 (section 4.3), after a learning phase of two short sequences of action, the adaptation of children to varying types of unexpected perturbation in the action sequence was assessed via the analysis of error patterns, response time and kinematic variables. Furthermore, the relative contributions of age, inhibition and switching tasks' scores to SuperCook tasks' variables (action selection patterns and response time) were tested.

The findings were broadly as expected following the results of Decker et al. (2016), that is: in the absence of perturbation children did equally good across the age range, proposedly due to an already mature model-free (routine) system, whereas with perturbation, errors decreased with age, proposedly due to an improvement of model-based (supervisory) system with age.

Furthermore, we were able to break down the role of some components of supervisory control. Younger children made less optimal action selection after rare_goal transitions than after rare_state, while older children performed equally well in both types of rare transitions, suggesting a difficulty in younger children to monitor or react to more distal change of cue (related to the goal, rather than related to the action they have just performed).

Additionally, among the two possible actions that were not optimal, the 'failure to update' seemed more frequent than the 'alternative action plan' regardless of the age and type of transition (nonetheless, the errors are too few to confirm an absence of effect of age and type of transition on the types of errors). Considered with a more important contribution of the inhibition task's score to explain optimal selection rate (as compared to age and the switching task's score), the results suggest inhibition may play a central role in developing model-based control.

4.4.3 Variability of speed and signature of model-based/supervisory control recruitment

The kinematic data showed that older children have larger variability of speed (*varSpeed*) in both types of rare transitions than in common transitions. Younger children seem to have identical patterns of varSpeed and RT in common and rare change of state transitions than older children. Both age groups have higher varSpeed in rare change of state transitions than in common transition, which was linked to the absence of significant difference in RT between state transitions and common transition. Thus absence of difference in RT, together with the identical optimal selection state and common transition for older children, indicate either that the common and rare_state transitions do not pose additional difficulty or demand, or that a mechanism is able to compensate for the increased difficulty for rare_state compared to common transitions, such that the optimal selection and RT are identical in common and rare_state transitions (despite the increased difficulty of rare compared to common transitions). Given the structure of the task alone, the second possibility is more likely. In accordance, and if we admit that kinematics data can capture unfolding cognitive process (e.g., Freeman, 2018), the significant difference in varSpeed between common and rare_goal transitions suggest that there is a mechanism compensating for the difference in difficulty between rare and common transitions in older children. Younger children's RT was longer in rare_goal than in rare_state, while such a RT cost was not present in the older children's group. varSpeed was significantly lower in rare_goal than in rare_state transitions for younger children, which suggests that the compensation mechanism found in rare_state was not at play in rare_goal.

Taking the RT and varSpeed results together with the optimal selection data, where both age groups had equally high optimal selection proportion in common and rare_state transitions, but the younger children had lower optimal selection proportion in rare_goal, we propose the following. The hypothesised compensation mechanism mentioned earlier is the recruitment of model-based control. Such recruitment is needed to adapt an ongoing sequence following any type of rare transitions. High values of varSpeed may be a kinematic marker of the recruitment of model-based control. Indeed, common transitions can be assumed to rely mainly on model-free control given the design of the task where this type of trial has been repeated extensively in the training phase, and low varSpeed would be an indicator of model-free use.

Both age groups seem to be able to use model-free control. This is supported by good optimal selection on common transitions, and may be reflected the varSpeed kinematic marker (with low values of varSpeed in common transitions for both age groups). Older children seem able to recruit model-based control in both types of rare transitions, as evidenced by their good optimal selection proportion in all types of transitions, and which we can link to high values of varSpeed. Younger children seem able to recruit model-based control in rare_state transitions but not in rare_goal, as evidenced by their drop in optimal selection at rare_goal compared to rare_state. This is reflected in their varSpeed at rare_goal transition, which looks as low as in the common transitions, suggesting that the younger children keep relying on model-free control at rare_goal transitions.

Note that the varSpeed measure is proportional to the baseline speed for a given age (the correlation between the per-age average of varSpeed and average 2D speed is significant: $r = 0.2815$, $p = 0.0392$), which explains that varSpeed is always higher for older than younger children and makes direct between-age groups comparisons difficult. We thus limit ourselves to within-participants comparisons.

We also note that RT and varSpeed give different results from each other, and varSpeed allows us to make more precise deductions about the processes responsible for the change of RT, hence the change in action control capacities with development. This strengthens the interest of using kinematic measures as a complement to RT.

In summary, our results (better optimal selection, smaller RT cost and finer varSpeed adaptations found in certain conditions with age) indicate improvements in supervisory/model-based control with age. This is consistent with the findings of Decker et al. (2016) yet goes beyond by having two types of rare transition and investigating of the types of error.

The two-stage task used by Decker et al. (2016) dissociated model-free and model-based control by using probabilistic transitions and analysing the behavioural pattern (probability of repeating a previous action) at first-stage as a function of the previous reward and the transition that occurred. The model-based signature was the integration of the transition information with reward information (i.e., the behaviour differed according to the type of transition and to the previous reward received). The model-free signature was identified by the absence of integration of this transition information, using only reward information (i.e. the behaviour differed according to the type of reward but not the type of transition). There, model-based control is essentially the capacity to form and use a probabilistic model of the task transitions to guide choice. In the SuperCook task, probabilistic transitions were used to probe whether one would successfully update a sequence following a change in the middle of its execution. We assumed that common transitions could be carried out by model-free type of control, and that rare transition would elicit the recruitment of model-based type of control. The SuperCook's data was consistent with the view. Therefore, two different tasks, SuperCook and Decker et al.'s (2016), suggest the overall same conclusion that model-based increases with age, although the time scale differed. Indeed, Decker et al. (2016) compared a 8 to 12 years old group to adults, while the present study SuperCook compared changes within the 5 to 9 years old range.

Decker et al. (2016) linked the developmental changes in *model-free/model-based* balance to the developing capacity for forming *model-based* computations. The Decker et al. (2016) does not explain how children become able to use such computations to guide their action. In our task, the knowledge and ability to perform the computations for selecting the correct actions seem present (given a non-negligible proportion of correct responses even in the rare_goal transitions). Difficulties arise possibly due to poorer ability to monitor, or poorer inhibitory abilities to override an ongoing sequence. Thus, the SuperCook task goes a step further in allowing to separate the supervisory processes at play.

4.4.4 Future work

Although the manipulation of icons on a touchscreen is more naturalistic than responding by key presses, future studies should confirm the findings with more naturalistic paradigms.

Furthermore, more studies should confirm the role of variability of speed as a marker of model-based control, notably in Decker et al.'s (2016) two-stage task. To do so, one could think of a touchscreen-based adaptation of the two-stage task, exactly identical except that an icon must be drag-and-dropped onto the options (instead of selecting the options with the keyboard). The following results would strengthen the role of variability of speed as a model-based signature: finding a high varSpeed when the stay patterns commonly associated with model-based are found, while finding low varSpeed when the stay patterns commonly associated with model-free are found.

Finally, the present results are not sufficient to draw conclusions about the relative importance of monitoring and inhibitory control, but adaptations of the task could tease them apart. It is possible that in this task children were able to detect the change of goal but sometimes too late to interrupt the ongoing sequence and/or update the action plan. Indeed, the supervisory system/model-based computations are known to be more time consuming (e.g., Hardwick et al., 2019). A variation of the SuperCook task with two conditions could test this hypothesis. A condition would encourage the children to take their time (*slow* condition). Another condition would encourage speed over accuracy by giving bonuses for faster responses (*fast* condition). If the frequency of errors at rare_goal type of transitions of younger children is smaller in the *slow* condition than the *fast* condition, this would indicate that the model-based system is too slow or unable to interrupt the ongoing sequence, presumably due to low inhibitory control capacity. However, if their performance is identical in *slow* and *fast* conditions, this would suggest that it is the monitoring process that fails in younger children.

4.5 Conclusion

Taking together the frequency of errors according to transition type, the types of errors, the RT and kinematics, the results suggest that children across the 5 to 9 years old age range use model-free type of control equally, while they increasingly use model-based type of control with increasing age. The increase in model-based use with age is consistent with and extends the findings of the study of Decker et al. (2016) (who compared 8- to 12-year-olds to adults while this study finds a change within the 5- to 9-year-old range). The task also allowed us to probe the subprocesses underlying model-based control.

Manipulating the type of rare transition (modifying the goal cue or the state following the first action) revealed that younger children struggled more with the change of goal cue, which may be linked to higher difficulty in monitoring the goal, a relatively distal cue, while having the capacity to update a sequence following a change of sate. The errors patterns indicated that the capacity to monitor and/or to inhibit an ongoing sequence are developing within this age range but were not sufficient to discriminate between the relative importance of the two mechanisms. The high contribution of inhibition score to predicting optimal selection suggests that inhibition plays an important role in updating action plans following unexpected changes of the environment.

All in all, key supervisory subcomponents driving the change in improving action control throughout the school years most likely involve both the monitoring of the environment and inhibitory control to correct an ongoing action sequence before a pre-planned, no longer appropriate action has been executed. This needs to be explored with further studies, and the variability of speed kinematic marker would be a useful tool for that purpose.

Chapter 5

Hierarchical Control of Action Sequences: Two Online Studies

Abstract

This chapter presents two related online experiments investigating children's and adults' control of action sequences, in particular action sequences that are routinised (or partially routinised). A new task was developed for this purpose, involving the repetition of 7 or 9 step action sequences over a large number of trials spread across 4 different days.

Both experiments replicated the finding that adults use hierarchically-structured task representations as presented in the dual-systems theory (Norman and Shallice, 1986, Cooper et al., 2014). Crucially, they extended this finding to school-aged children. The chapter validates the suitability of the dual-systems framework to investigate 7-12 years old children's action control and addresses qualitative and quantitative similarities between adults' and children's sequential action control.

Specifically, experiment 3 found a high concentration of errors at subtask boundaries, and a trend for higher frequencies of those errors in children. In experiment 4, children had an even more marked increase than adults in selection latencies at special points (branch points) compared to control points. Together with the effects of a secondary task varying according to points in the sequence and age, as well as investigating a kinematic marker (coefficient of variation of speed), the results highlight the role of developing supervisory functions in the emergence of efficient sequential routine performance throughout childhood.

5.1 Introduction

We begin by reviewing the key theoretical and experimental elements of sequential action control which guided the design and analysis of the two studies. Theories of adult action control use converging evidence from error and response time data. Experiment 3 focuses more on error data whereas experiment 4 focuses more on response time and kinematics data.

Experiment 3 explores the overall representation used when executing routine sequences. The dual-systems theory (Norman and Shallice, 1986, Cooper et al., 2014, presented in detail in the General introduction) posits that representations have a hierarchical structure, where actions are organised around subgoals, which themselves can be part of higher-level subgoals. This structure becomes apparent when looking at errors produced by participants: the errors are particularly more frequent at the boundaries between subtasks (that is, subsequences that achieve a subgoal).

Experiment 4 explores special points in the sequence, branch points, that are particularly difficult given their position at the beginning of a subsequence and requiring a decision between diverse possible subtasks. Let us take the daily-life example of sequence ’making a cheese sandwich’. This sequence shares a common start with the sequence ’making a ham sandwich’: in both cases, one would start with preparing the bread and then adding butter. At this point, one should either enter the subtask ’adding cheese or ’adding ham’. This constitutes a branch point. When executing a routine sequence, branch points have shown to have longer selection latencies than non branch points (as found in Ruh et al., 2008, Ruh et al., 2010). The longer selection latencies has been attributed to the need for supervisory control at branch points whereas non branch point require no supervisory influence for a very-well practised sequence.

Given the correspondence between supervisory control and executive functions, we will use one or the other term interchangeably. As mentioned, executive functions are known to improve vastly throughout childhood (e.g., Casey et al., 2005b). Our core assumptions are grounded in the gradual improvement of executive processes throughout childhood which is related to increasing efficiency of action sequence performance with age. Such interaction between executive processes and action sequence performance can be attributed to the role of executive functions in action control. We will specify predictions in the dedicated section of each study.

5.1.1 Ruh et al.’s (2010) virtual gardener task

We designed a new task to investigate the acquisition and control of action sequences and its underlying representations in school-aged children (6-12 years old) and adults, by means of a computerised game. The design of the game was inspired by Ruh et al.’s (2010) ’virtual gardener’ computer task albeit adjusted to be suited for children. As mentioned previously, Ruh et al. (2010) investigated routine action control in hierarchically structured sequences in adults. In this computer-based study, participants were tasked to make virtual fertilizers for a plant according to an instructed goal: the instructed goal was to elicit a given outcome (the plant growing larger, getting bushier or making flowers), and this was achieved by three different action sequences. The action sequences consisted of actions such as picking up items, placing items on the worktop and so on, which were carried out by clicking on and moving items with the computer mouse. The sequences comprised at least 10 basic actions. All three sequences had a common structure and an identical start, and later varied in the ingredients to be picked up or the number of ingredients to be added. Some trials were carried out under dual-task conditions (sound counting). The study was carried out over 3 different sessions and involved 200 trials in total.

5.1.2 The GoGelato task

The game developed for these experiments consisted of a succession of attempts to achieve one or the other of the two instructed overarching goals by manipulating icons on screen, and each goal could be completed by distinct but vastly overlapping action sequences. The goals were ’make a chocolate - wafer - sprinkles ice cream’ or ’make a chocolate - wafer - banana ice cream’. The sequences were designed to present common features of everyday action sequences, namely:

- the individual actions leading to the main, higher-level goal could be conceptualised in terms of several groups of actions that achieve a subgoal (just like making a lunch box can be broken down into subgoals such as making the sandwich, preparing a drink, etc.);
- the subsequences used in this game can easily be imagined to serve different overarching goals, e.g., scooping Nutella can be thought of being part of the ’making a toast’ goal (instead of making an ice cream);
- the subgoals could be used in different sequence variants in this game, and some sequences share the same beginning (just like making a cheese sandwich or a jam toast will start by picking-up bread, spreading butter, etc.). Consequently, in the section common between the two goals, actions have higher transition probabilities than at the point where they differ (at which there is a lower transition probability for one or the other possible actions at this step). This consists in a branch point, as will be elaborated on later.
- different actions have different functions: while some of them consist of adding ingredients onto the ice cream (chocolate, wafer or sprinkles/ banana), others are ’enabling’ action (e.g., opening the Nutella lid) or ’cleaning-up’ (placing the spoon back to its slot) actions.

When describing the paradigm, we will use the term ’subsequence’ to designate a section of the sequence that can be conceived as achieving a subgoal. It is left to be proven in the data whether participants actually use such hierarchically-structured representation (where actions are organised around subgoals) to guide their actions.

Besides adding features of real-life action sequences, several aspects were controlled for in order to ensure the interpretability of the results:

- For one of the two goals, the sequence includes using the same implement (spoon) in two different parts (as can be seen in Figure 5.1): within a subsequence (the Nutella subsequence) and at the start of a subsequence (the Sprinkles subsequence). Furthermore, a subsequence was inserted in-between the two usages of the spoon to justify having the participants place the spoon back to its slot (this ’discarding’ action was instructed like the other actions in the sequence). Consequently, in both cases the selection of the spoon started from the centre to go on a peripheral location. This allowed us to directly compare the two 'select spoon' actions, and investigate the differences in selection latencies between a within-sequence and a between-sequence action.

- The two sequences had the same number of actions, and the intuitiveness of the overall recipe did not strikingly differ in favour of one or the other.

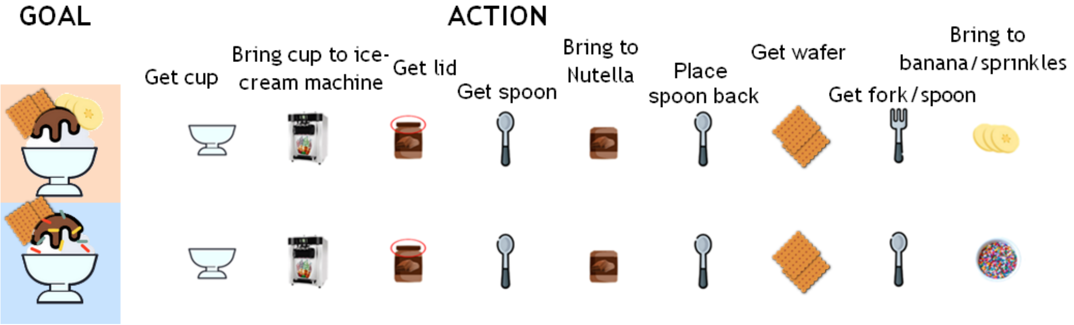


Fig. 5.1: **Action sequence: the 9 main actions of the game.**

This, along with the controlled parameters (e.g., constant distance between icons, described in the methods), were chosen to balance the naturalistic aspect of the game with analysability and interpretability, given that a child sample notoriously leads to noisier data than adults. Importantly, the complete study had a large number of trials, spread over 4 sessions on 4 different days, to allow investigating the control of a very-well practised sequence, which we assume we get to a '(partially) routinised' stage, if not fully routinised.

5.1.3 Manipulation of cognitive load through the addition of a secondary task

Additionally, a hallmark or consequence of routine sequence control is that such action sequences demand very little cognitive effort and can be very well executed under divided attention or cognitive load, such as when multi-tasking. However this is only true of sufficiently routinised sequence or subsequence, and some points can remain vulnerable to error under heightened cognitive load. Hence, studying action performance while simultaneously performing a secondary task can tell us about the nature of representation underlying performance, even more than under normal conditions. This is why we introduced a secondary task, that used a different modality (auditory), in the last two sessions where action execution can be expected to be (partially) routinised. This follows the use of a secondary task in order to compete with higher-level resources as used in, for example, Ruh et al. (2010), Dunbar and Sussman (1995) or Humphreys et al. (2000). In Humphreys et al. (2000), a secondary task successfully induced errors in healthy participants by presumably taxing executive processes resulting in loss of activation over component actions.

To ensure the secondary task was adapted for children we followed Broadbent et al. (2018) who used a sound counting dual-task, and found deteriorated performance compared to single-tasking. Thus the sound counting is an appropriate way to heighten cognitive load while allowing multi-tasking in 6- to 10-year-old children.

5.2 Experiment 3: Establishing the Emergence of Hierarchy

5.2.1 Introduction

In this experiment, the errors performed by children and adults on two 9-steps action sequences are used as a window into processes underlying action control, and the development of such processes. With this experiment, we addressed the following questions: How well do school-aged children routinise action sequences compared to adults? What representations underlie (partially) routinised sequential action control across development? How is such action control affected by a secondary, attention-demanding task? As a prerequisite to answering the latter questions, we also asked: is the newly-developed task suited to investigate routine sequential action control in school-age children?

5.2.2 Methods

5.2.2.1 Participants

Children (total N=23, 14F) between 7 and 12 years old were recruited mainly through Birkbeck's Babylab database (where they were contacted by email), but also via Babylab's social media and word-to-mouth. There were no specific inclusion or exclusion criteria beyond having access to suitable computer hardware (see below).

Adults (total N=23, 11F) were aged 20 to 35 years old. Half of the adults was recruited via word-to-mouth. The other half was recruited via the platform Prolific, with the following inclusion criteria:

- Use a desktop computer or laptop;
- Age 20-30;
- English-speaking country (among: the UK, the US, New-Zealand, Norway, Sweden, and the Netherlands);
- Had taken part in a least 10 studies on Prolific, and had had an approval rate above 97% in previous participation. Such criteria on participation were added to maximize chances to get all four sessions completed by most participants.

We included adults in this paradigm for three reasons: 1) to replicate and extend findings on adult sequential action control, with a different context (ice cream making) than previous paradigms, 2) to validate the newly-developed task (assuming that the adult sample would have less noise than the child sample and thus permit to verify if key findings on adults action control are replicated by our new task), and 3) to permit a direct qualitative comparison, and (with some caveats developed in the discussion) permit a quantitative comparison between adult and children's action control.

Adults were compensated £2.5 for session 1, £2 for session 2, £2 for session 3 and £3 for session 4 (the amount was higher at the start because of the extra time it takes to familiarise with the study, and higher in the end to encourage participants to come back). Families were offered a gift card of the following amount (for each child taking part): £4 for session 1, £3 for session 2, £3 for session 3 and £5 for session 4. Adults gave formal consent by ticking the consent form's box. Children in session 1 gave formal consent verbally and by ticking their consent form's box, the same went for their caregiver. On the other sessions, children and caregivers gave consent by ticking the boxes. The study had ethical approval from Birkbeck's College Ethics Committee and was conducted in accordance with the Helsinki declaration.

Table 5.1 shows the number of completed sessions of the recruited participants per age.

Table 5.1: **Sample sizes, broken down by group (adults and children), and children's group further broken down by age.**

N (sample sizes by group)	Adults	Children	7 y.o.	8 y.o.	9 y.o.	10 y.o.	11 y.o.	12 y.o.
Took part (started Session 1)	23	23	5	7	3	3	4	1
Included (completed at least S1)	23	22	4	7	3	3	4	1
Completed at least S1 & 2	21	19	4	6	2	3	4	0
Completed at least S1, 2 & 3	18	18	4	5	2	3	4	0
Completed S1, 2, 3 and 4	17	18	4	5	2	3	4	0

For analyses, we median-split the children into two groups (median: 8 years old). When looking at all children having completed session 1, the age distribution was: 7-8 years old: N = 11, 9-12 years old: N = 11. When looking at children having completed session 4, the age distribution was: 7-8 years old: N = 9, 9-11 years old: N = 9.

5.2.2.2 Materials

The study took place entirely online. Participants were required to have a desktop computer or laptop to complete the study. They were strongly encouraged to use a mouse over a laptop trackpad, but both were authorised. Devices with only touch-screen responses (e.g., tablets without external mouse) were not allowed. Participants also were required to have sound (via headphones or speaker) and a normal internet connection.

Apparatus

The game was programmed in JavaScript with some help of the library JsPsych, and hosted on a University server accessed through the software JATOS.

I entirely programmed the game in order to ensure reliable data acquisition, avoid unexpected biases due to variations in the participants' devices and ensure the game would be suited to measure children's action control. This allowed fine control over the behaviour of the game and over the data recorded, which can be a challenge for web-based experiments and impact the findings if not well mastered (controlling display, size, speed of the display, fluidity, sampling rate etc., see for example Tsay et al., 2021).

While developing the paradigm, unsystematic pre-piloting was carried out on small batches of participants (N<5), and data was partially analysed to develop an intuition of the usability, the intuitiveness of the game and instructions, its duration, and help select parameters such as the number of trials and number of sessions.

Action sequences

The task was gamified so that children would stay engaged throughout the entirety of trials within one session and remain eager to come back for the other sessions. The sequence (ice cream making) was realistic so that it would be easier to learn than an abstract sequence of actions and consequently so that participants could reach a level of routinisation more quickly. However, it was chosen so that it would not be a very routine sequence in participants' life (as would have been making a sandwich for example), which limits the risk of interference with personal knowledge and varying mastery of the routine with age.

The task consisted of a child-friendly cover story and stimuli, introduced to the participants in several stages. Participants were instructed that the objective was to make ice creams to sell them, and that from one time to the next customers could ask for one of two recipes (which would sometimes change from one customer to the next and sometimes not): if the background was blue, they had to make the chocolate, wafer and sprinkles ice cream, and if the background was orange, the chocolate, wafer and banana ice cream.

To make the ice creams, icons had to be dragged-and-dropped by clicking on an icon with the computer mouse to select the icon, displacing it to a target location and releasing the click to drop the icon on that location. There were 9 steps to be made strictly in the right order. Table 5.2 presents the actions to be taken at each step. The table describes the action in terms of source icon to target icon and consequently does not describe the entirety of mouse movements. The complete subsequence of mouse movements can be deduced from the location of the target icon of the preceding action. For example, Step A involved moving the cursor from the centre (where the cursor initiates) to the cup, clicking on the cup and drag-and-dropping the cup onto the ice cream machine; or step E involved clicking and drag-and-dropping the spoon from the Nutella icon to the centre.

Direct feedback was given each time an icon was dropped (by releasing the mouse click). If the action was correct (that is, the correct icon was dropped onto the correct location at the correct step), then the logical consequence was shown on screen (e.g., the cup was filled with ice cream after being placed onto the ice cream machine; or the centre was updated as shown on Figure 5.3). If the action

Table 5.2: **Description of the correct action at each step.**

Step	Correct action
A	Get cup and bring to ice cream machine machine
B	Cup filled with ice cream to centre
C	Get lid and bring to centre
D	Get spoon and bring to Nutella
E	Spoon with Nutella to centre
F	Place spoon back to original slot
G	Wafer to centre
H1/H2	Fork/spoon to banana/sprinkles (topping)
I1/I2	Topping to the middle

was incorrect, the game automatically placed the moved icon back to its original location. At the same time, a 'buzz' sound was played to indicate the action was incorrect. In other words, only the correct action enabled continuing the trial, and consequently actions could not be performed in the wrong order. This was done to facilitate learning and to get more analysable data (by constraining the range of possible behaviours). Note that it was still possible to complete incorrect subsequences at the last step (e.g., fork-banana instead of spoon-sprinkles, or even spoon-banana or fork-sprinkles) since the action still matched the step.

When an icon had to be dropped onto another icon, the action was valid if the overlap between the source and target icon was greater than 1100 squared pixels (i.e., it did not require a perfect alignment of the centre of both icons), for purpose of fluidity of the game and to avoid distortion due to less fine motor control in children.

As described in the following Section, the appearance of the ice cream changed when a correct item was added to it, which gave a subtle cue regarding where in the sequence one was (just like in real-life baking, the dough would change colour after adding chocolate). This could be used by participants as a cue to disambiguate the position in the sequence in case of forgetting.

At the end of a trial, feedback was given. The 'correctness' of goal execution refers here to doing the right recipe for the given background (e.g., sprinkles ice cream for the blue background). An upbeat sound indicated correct goal execution, or a 'buzz' sound and the display of a red cross indicated incorrect goal execution. Furthermore, the sound for positive feedback came with:

- for adults, a green thumbs up and an animation of a coin falling on a stack of coins (the coins accumulated throughout the game);
- for children, an animation of a coin falling into a piggy bank, with a cartoon pig grunt sound.

For children, every 4 trials another animation was added to encourage them to continue until the end: the piggy bank broke to reveal the coins accumulated and this led to adding a part of a new ice cream shop (which they had to build until the end; a new shop for every session).

The instructions were phrased to avoid influencing children towards one side or the other of the speed-accuracy trade-off: *'Try to work as quickly as it's comfortable for you'*.

Stimuli appearance

The display at the start of a trial is shown in Figure 5.2, left panel. The game area was delimited by the background colour (the rest of the web page being the user's default, most likely white), forming an 'arena' of dimensions 650 pixels. This represents 11.5 cm on a 24-inch screen with a 1920x1080 screen resolution. For the following dimensions, we also give the equivalent for this setup in cm, between brackets.

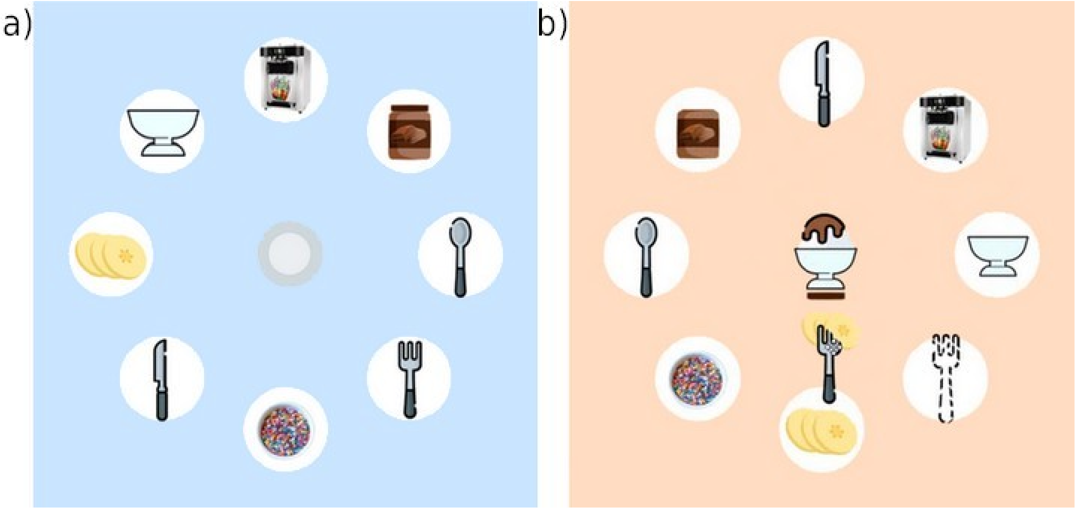


***Fig. 5.2:* Task's main display a) at the start of a trial, and b) towards the end of a trial, while the final icon is being displaced.**

The background colour was blue or orange. Both colours had the same saturation (100%) and lightness (89%) and their hue was 210°for blue, 26°for orange. In the centre of this arena was the icon of a plate of diameter 85 pixels (1.5 cm). This icon could not be moved nor picked-up by clicking (we refer hereafter to such icons as 'static'). Peripheral icons were located within 8 'slots' arranged around the centre, at a 211 pixels (3.8 cm) radius distance from the plate. The slots around icons were disks of diameter 113 pixels (2 cm). All the peripheral icons had comparable surfaces, but their shape inherently varied: in particular the implements such as spoon or fork had about 29 pixels (0.5 cm) width.

When the spoon was released onto the Nutella, the spoon was updated with Nutella on it (similarly for the sprinkles, or for the fork onto the banana, or the cup on the ice cream machine). When a correct icon was released onto the centre, the icon in the centre was updated to show the progress of the ice cream state as shown in Figure 5.3.

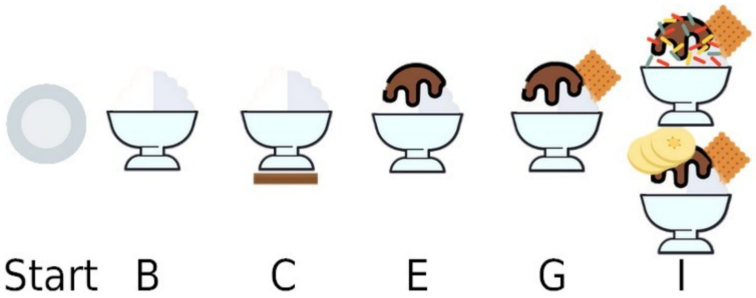


Fig. 5.3: **Evolution of the middle icon throughout the game, illustrating the progress in the task. The letters indicate the step at the end of which this central state appear.**

Because the cursor location cannot be controlled on a web browser, there was an extra page between two successive ice cream trials. The extra page required a click in the middle of the page to proceed to the next page. It showed a button with a 'play' symbol, that when clicked on triggered the start of a trial. The location of the cursor when clicking the 'play' symbol coincided with the centre of the plate in the next page, which allowed a normalised start location of the cursor across the ice cream trials. This also allowed participants to take breaks and thus potentially reduced the number of within-trial interruptions.

Randomisations

The task is primarily concerned with action selection, mediated by reaching for icons, unlike tasks such as the SRTT (Serial Response Time Task, Pascual-Leone et al., 1996) which are more geared towards motor learning. Consequently, to ensure the task was performed by selecting items and not simply remembering and going back to a sequence of locations, we randomised the position of items (among the possible slots on the peripheral circle) at every trial. Note that this introduces a visual search stage to the task, and thus prevents participants from getting response times as fast as those that can be found in motor learning tasks with extensive practice. A trial that had timed out restarted with a new random set of positions.

Each trial's goal (indicated by the background colour) was pseudo-randomised with a shuffled sequence of 8 trials of each type (the pseudo-randomised sequence was different for each participant).

Normalisation trials

Before the ice cream making trials, a different type of trial was administered, which had two aims: 1) to collect individual motor time, that is the time to go from the centre to a peripheral icon (red target) when there is no item choice to be made, and 2) to have participants familiarise themselves with the type of mouse-controlled drag-and-drop actions used in the game, thus making the instructions of the game gradual.

It consisted of an arena and white disks with dimensions and arrangements in space identical to the main trials (and a grey background with the same lightness as the coloured backgrounds), as shown in Figure 5.4. A red diamond was the icon to be selected and to be drag-and-dropped onto a blue diamond. The trial stopped at this time. Four of these trials were given at the start, and three at the end.

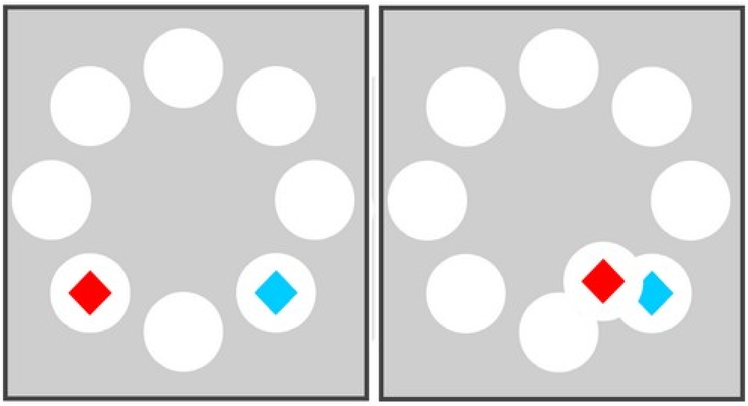


Fig. 5.4: **Normalisation trial: (left) at the start, and (right) close to completion of the trial (where the red target overlaps the blue target).**

Demonstration trials

The game was instructed by animated images (GIF) that showed the game to the participants step-by-step. The demonstration used a direct path in order not to influence the representation of the sequence that participants might build, nor influencing their cursor trajectory. The demonstration was broken down into units that correspond to the single steps shown in 5.2. As a consequence, if chunking in the sequence appeared in the data, it would not be attributable to the demonstration.

Each of the steps was introduced by an animated GIF made of three different frames with the start location, end location, and a middle point location right in the centre of a straight path (with several frames to indicate the start/end), as shown in 5.5. The study window was split into two parts: the animation was shown on the left side and looped until the participant executed that step on the right side.

Secondary task

The secondary task was added to study the role of top-down control in the sequential action task. Like in Ruh et al. (2010), the secondary task was (1) auditory, (2) required responding after the sequence execution was over, and (3) was made unpredictable. These ensured that the task would not become routinised, would not be facilitated by switching nor would be easy to guess.

The task consisted of counting bird chirp sounds which occurred at random times during the ice cream making. Participants had to report the count at the end

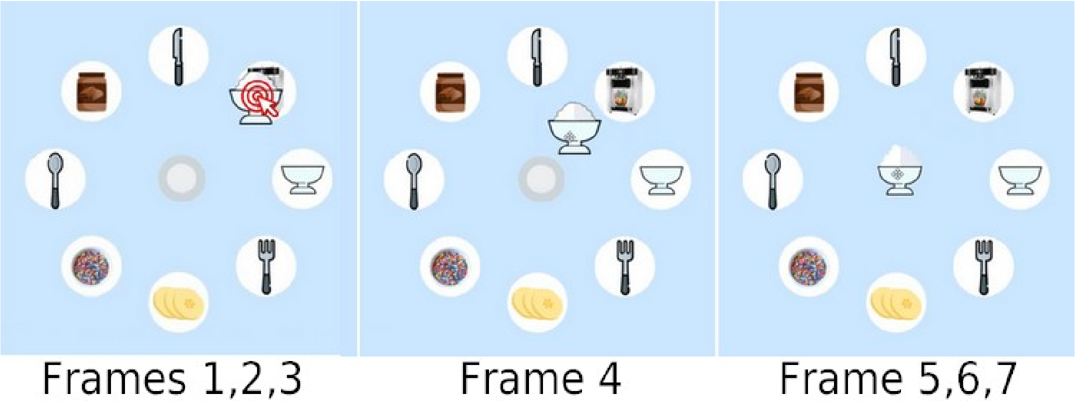


Fig. 5.5: **Frames used to make the demonstration animation (step B is taken as example here).**

of a trial. For the secondary task, the sentence in the instructions prompted participants to give equal importance to the ice cream making task and the bird-chirp counting task, and importantly not to pause when counting: *'(...) But don't stop the game to count- Keep making the ice cream while you are counting.'*

The number of sounds varied randomly between 3 and 7 for children, and between 6 and 11 for adults. The duration of a chirp sound was 170 ms. The interval between successive sounds varied randomly between 250 and 1500 ms.

Earlier in the study (cf. procedure in Section 5.2.2.3), participants had been familiarised with the bird chirps sounds in the System set-up stage prior to the main game, where they had played a recording which had 4 to 7 of these bird chirps and indicated the counts. The study only went forward if the count was correct.

5.2.2.3 Procedure

Participants completed four sessions over separate days, as shown in Figure 5.6. Day 1, 2, 3 and 4 were consecutive days, except for a few adults who had up to three days spacing between two consecutive sessions.

One online session had the following structure:

- Introduction and consent form
- Questionnaire (age, preferred hand for using the computer (left/right), preferred among mouse or trackpad in general, and the one (mouse of trackpad) used for the study)
- Set-up (verification display and sound)
- Normalisation trials (start round)
- Instructions and demonstration of the main game
- Main trials of the game
- Normalisation trials (end round)
- Debrief

The overall procedure was the same for children and adults, except that the first child session involved a video call with the family and the child. After an introduction, they were taken step by step through the study, with a screen-share to ensure the instructions were understood. They also had the opportunity to ask questions in the end and receive instructions for doing the three following sessions on their own on three different days.

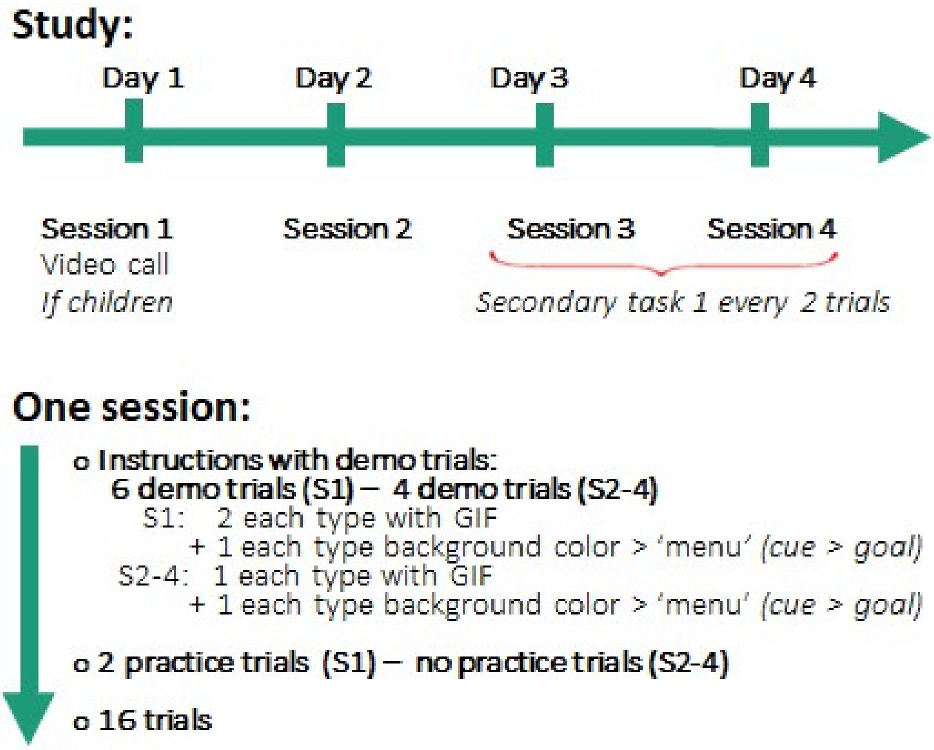


Fig. 5.6: **Overview of the multi-day experiment, and the number of trials and trial type per session.**

5.2.2.4 Analyses

In this study, we consider three age groups: the 'younger children' group **7-8**, the 'older children' group **9-11** and the 'adult' group **ADU**.

In all the following analyses, we excluded: (a) the demonstration and practice trials, (b) the trials with a time-out and (c) the steps which included a 'motor/click error' (i.e., a click right next to an icon that failed to select the icon, or a click-release (drop) right next to another icon but not sufficiently close).

Analyses are separated between (a) an overview of performance and trial durations across sessions and age groups, and (b) finer-grained analyses in the *(partially) routinised* sessions, i.e., sessions 3 and 4.

Trial duration measures started when the stimuli appeared, and stopped when the last icon was dropped onto the centre. The first analysis evaluates the assumption that sessions 3 and 4 can be considered *(partially) routinised*: the evolution of trial duration across sessions was assessed with the Wilcoxon signed-rank test on the median response times of participants in each session, within each age group. Additionally, we tested the proportion of incorrect goal execution across sessions within each age group with paired samples t-tests for with and without secondary task comparisons, and independent samples t-test for age group comparisons.

The rest of the analyses are concerned with *(partially) routinised* performance, and thus are restricted to sessions 3 and 4. They are also restricted to the trials with correct goal execution, to minimise noise which would impede the interpretations. We tested the step errors with paired or independent samples t-test, as for the incorrect goal execution. We carried out all comparisons for a given step in a given age group, or for each step in a given condition (with or without secondary task), either comparing with to without secondary task, or comparing age groups.

T-tests were chosen because of the high number of participants having an error proportion of zero. For such comparison, a Wilcoxon test is not sufficiently sensitive. The statistics were computed with the module *scipy.stats* in Python. Finally, due to the high number of comparisons (across all steps), significant results need to be interpreted with caution.

The error classification in the current task was based on the error taxonomy from Schwartz, Buxbaum, Montgomery, Fitzpatrick-Desalme, et al. (1998). We translate the type of errors, from the multi-level action test (MLAT) task of Schwartz, Buxbaum, Montgomery, Fitzpatrick-Desalme, et al. (1998) to the equivalent errors from the present task:

- Omission: failing to add ice cream in the ice cream cup;
- Sequence Anticipation/Omission: adding the wafer before adding Nutella;
- Sequence Perseveration: making two ice creams (note that this was not made possible in the present task);
- Object substitution: attempt to pick up sprinkles with the fork instead of the spoon;
- Action addition: placing the spoon onto the ice cream machine.

We distinguish two types of omissions: 'step omission' when one step within one subgoal is omitted, and 'anticipation/omission' when an entire subgoal is omitted. Indeed, it is hard to distinguish whether a participant forgot a subgoal or jumped to the one after. We present the errors that occurred at least two times (in the total dataset across participants and conditions).

When looking at response time data, we are concerned with action selection latencies, from comparable points across similar groups of actions. These latencies were either (a) from the centre of the game arena until the step's icon has been selected (by clicking) or (b) from a peripheral icon to another peripheral icon. Table 5.3 gives the step by step description of these portions of action selection whose latency we are interested in.

We conducted a series of linear mixed-effects regressions with the nlme package (Pinheiro et al., 2018) for the R software environment (Version 3.6.0; R Core Team, 2019), fitting the models by maximizing the restricted log-likelihood (REML

Table 5.3: **Description of the portion of the action whose selection latency is measured, at each step.**

Step	Action
A	Move cursor from centre to cup
B	Move cup onto ice cream machine
C	Move cursor from centre to lid
D	Move cursor from centre to spoon
E	Move spoon onto Nutella
F	Move spoon from centre to its original slot
G	Move cursor from centre to wafer
H	Move cursor from centre to fork/spoon
I	Move fork/spoon to banana/sprinkles

approach). The dependent variable (step duration) was unskewed by a log-transformation. We fitted one model per step, whereby we defined a random intercept per participant to account for the repeated measures, and included Age group (categorical variable, 3 levels) and secondary task (categorical, 2 level) as fixed effect factors; as well as all the possible interaction terms.

5.2.3 Results

5.2.3.1 Acquisition phase and overview age differences

Figure 5.7 shows the trial durations across steps and condition (with/without secondary task; for session 3 and 4) in each age group. When comparing sessions 1-2 against session 3-4 (without secondary task), session 3-4 had significantly shorter trial duration than 1-2, within each age group ($W(N = 19) = 7.0$, $p = 0.0001$ for 7- to 8-year-olds and 9- to 11-year-olds, $W(N = 19) = 7.0$, $p = 0.0001$ for adults). When comparing with secondary task against without (in sessions 3-4), trial duration was significantly larger with the secondary task than without, within each age group ($W(N = 8) = 5.0$, $p = 0.0391$ for 7- to 8-year-olds and 9- to 11-year-olds, $W(N = 19) = 5.0$, $p < 0.0001$ for adults). Results are summarised in table 5.4.

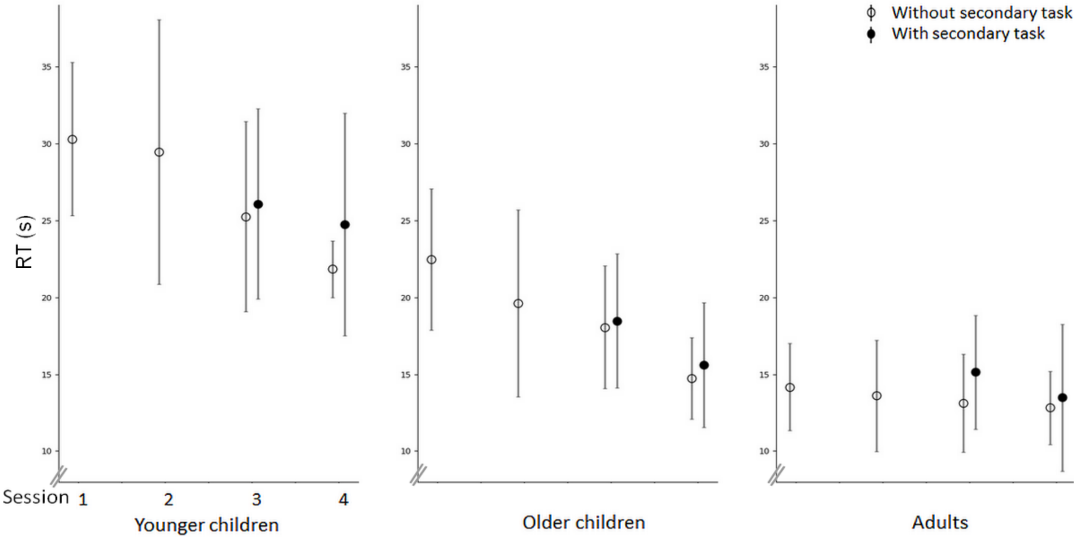


Fig. 5.7: **Median trial duration across sessions and age groups, with or without secondary task. Error bars show the inter-quartile range. Data from all participants regardless of their number of completed sessions.**

Table 5.4: **Wilcoxon signed-rank tests for Trial Duration comparing sessions (only without secondary task: ST_-) or conditions (with secondary task: ST_+ against without: ST_-; only in sessions 3&4). The table shows: W-value (Wilcoxon test statistic), N of pairs with non-zero differences, p-value.**

	ST_-, S1&2 vs S3&4	S3&4, ST_- vs ST_+
7-8	$W = 0.0$, $N = 8$, $p = 0.0078$	$W = 5.0$, $N = 8$, $p = 0.0391$
9-12	$W = 0.0$, $N = 8$, $p = 0.0078$	$W = 3.0$, $N = 8$, $p = 0.0391$
Adults	$W = 7.0$, $N = 19$, $p = 0.0001$	$W = 5.0$, $N = 19$, $p < 0.0001$

The Proportion of Incorrect Goal Executions slightly decreased between S1&2 and S3&4 in younger children ($t(DF = 7) = 2.15, p = 0.0498$). Otherwise, it did not differ significantly between sessions of between conditions (i.e., with/without secondary tasks) (Figure 5.8 and table 5.5).

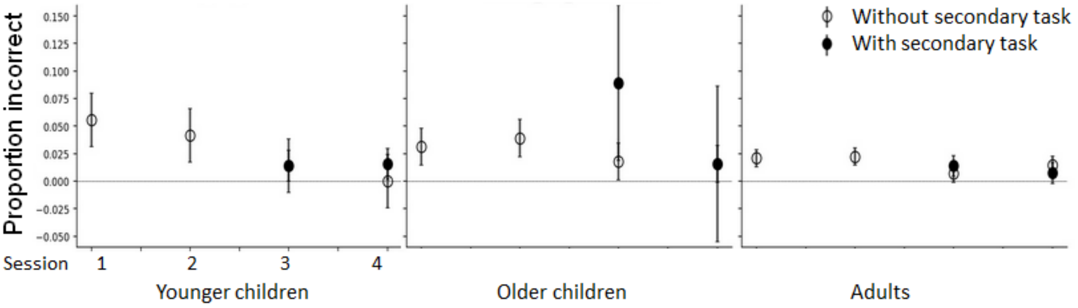


Fig. 5.8: **Average Proportion of Incorrect Goal Execution across sessions and age groups, with or without secondary task. Error bars show the s.e.m. Data from all participants regardless of their number of completed sessions.**

When collapsed across sessions and conditions (see Figure 5.9 and table 5.6), the age differences on Proportion of Incorrect goal execution were not significant either.

Table 5.5: **Independent samples t-test on Proportion of Incorrect Goal Execution comparing sessions (only without secondary task: ST_-) or comparing conditions (with secondary task: ST_+ against without: ST_-; only in sessions 3&4). The table presents t(degrees of freedom) and the p-values.**

	ST_- , S1&2 vs S3&4	S3&4, ST_- vs ST_+
7- to 8-year-olds	t(7)=2.15, p=0.0498	t(8)=-1.00, p=0.3466
9- to 11-year-olds	t(7)=-0.94, p=0.3807	t(7)=1.00, p=0.3343
Adults	t(18)=0.56, p=0.5765	t(18)=0.00, p=1.0000

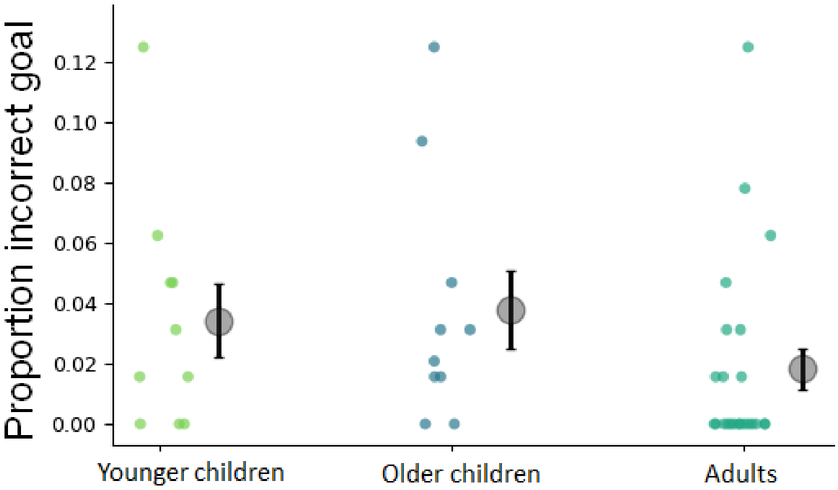


Fig. 5.9: **Average Proportion of Incorrect Goal Execution across age groups (collapsed across all sessions and conditions). Error bars indicate the s.e.m. and individual points represent individual participants.**

Table 5.6: **Independent samples t-test on Proportion of Incorrect Goal Execution comparing age groups. DF: degrees of freedom.**

	test-statistic	DF	p-value
7- to 8-year-olds vs 9- to 11-year-olds	0.27	15	0.7892
7- to 8-year-olds vs Adults	0.27	25	0.7892
9- to 11-year-olds vs Adults	0.27	26	0.7892

5.2.3.2 Step errors and latencies

Figure 5.10 presents the errors at the level of a single step but summed together to look at overall age differences. Younger children made significantly more errors at step-level (summed across steps) than older children ($t(DF = 26) = 6$, $p < 0.0001$) and adults ($t(DF = 16) = 2$, $p = 0.012$), as detailed on table 5.7.

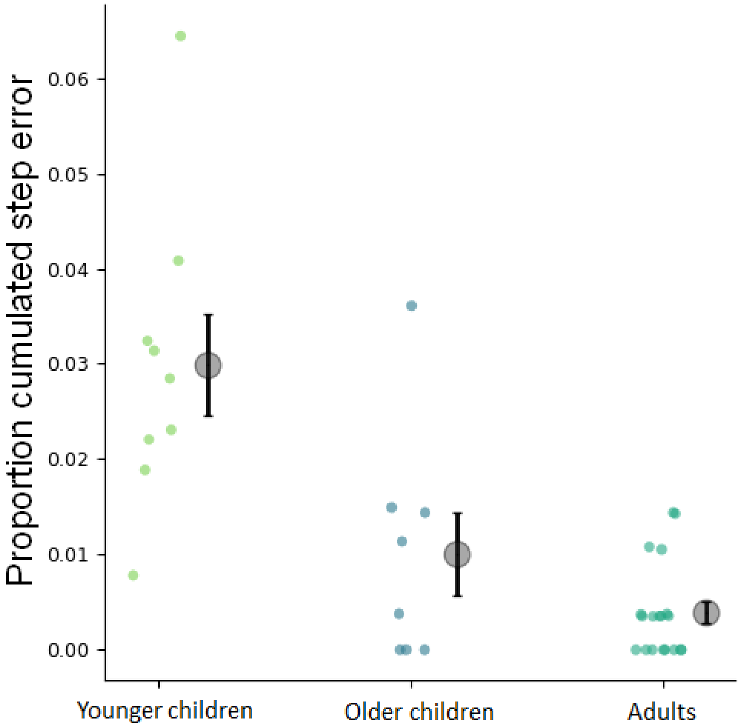


Fig. 5.10: **Cumulated step errors (proportion) in sessions 3 and 4 across age groups. Grey dots indicate the average across participants, error bars indicate the s.e.m. and individual points represent individual participants.**

Table 5.7: **Independent samples t-tests on the Proportion Cumulated step errors in sessions 3&4 across age groups. N_1: sample size of the group on the left-hand side, N_2 for the right-hand side of the comparison. DF: degrees of freedom.**

	test-statistic	DF	p-value
7- to 8-year-olds vs 9- to 11-year-olds	6	26	<0.0001
7- to 8-year-olds vs Adults	2	16	0.0124
9- to 11-year-olds vs Adults	1	27	0.0726

Figure 5.11 shows the proportion of errors across steps (for all the age groups and conditions together). The number of errors is not uniform across all steps. The steps B, E and I all show remarkably few errors (only 2 participants have non-zero proportion errors in these steps). The errors were concentrated at the other steps, and a one-sample t-test confirmed that the proportions of errors at all steps but step B, E and I significantly differed from zero, as reported in table 5.8.

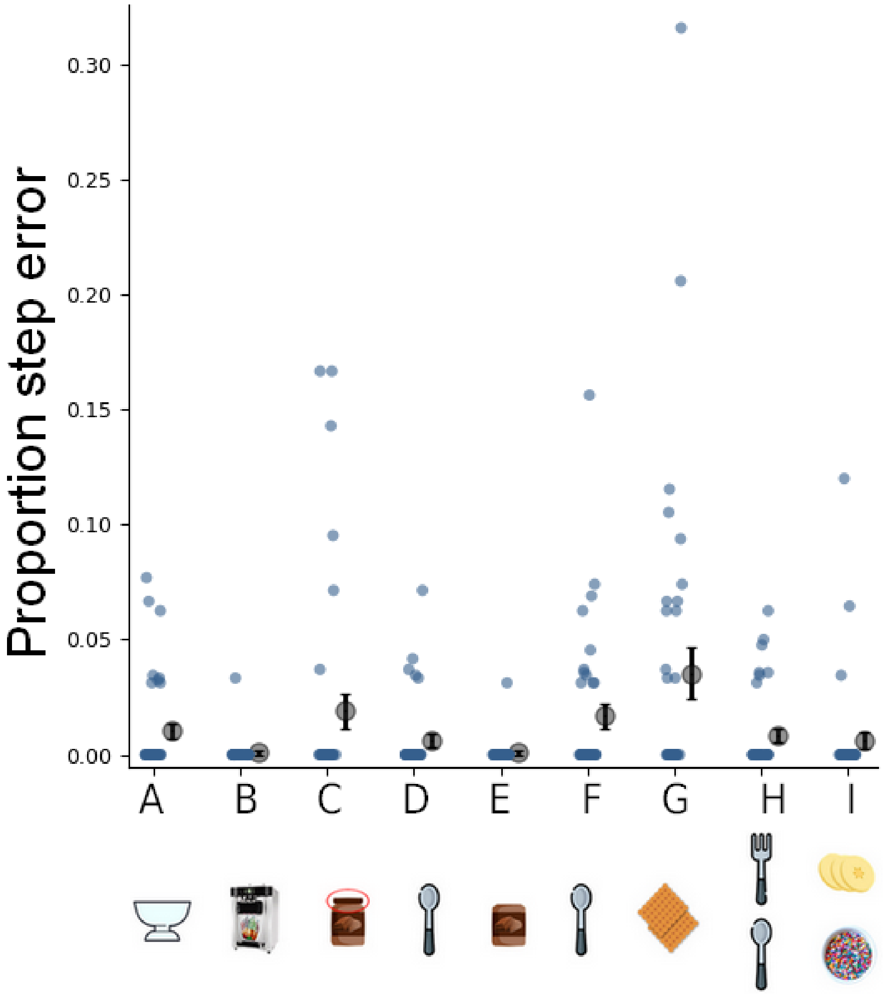


Fig. 5.11: **Proportion of step errors in sessions 3 and 4 (collapsed across age groups and conditions). Grey dots indicate the average across participants, error bars indicate the s.e.m. and individual points represent individual participants.**

Table 5.8: **One-sample t-tests comparing the errors collapsed across participants and conditions against 0, in sessions 3&4. The degrees of freedom are 71.**

Step	A	B	C	D	E	F	G	H	I
test-statistic	3.09	1.00	2.76	2.01	1.00	2.97	3.13	2.22	1.31
p-value	0.003	0.321	0.007	0.048	0.321	0.004	0.003	0.030	0.194

We now turn to errors at each step shown in Figure 5.12 and Figure 5.13. There was no significant effect of secondary task as detailed in table 5.9. Although not significant, the younger children's group had higher proportion errors at step C with secondary task than without.

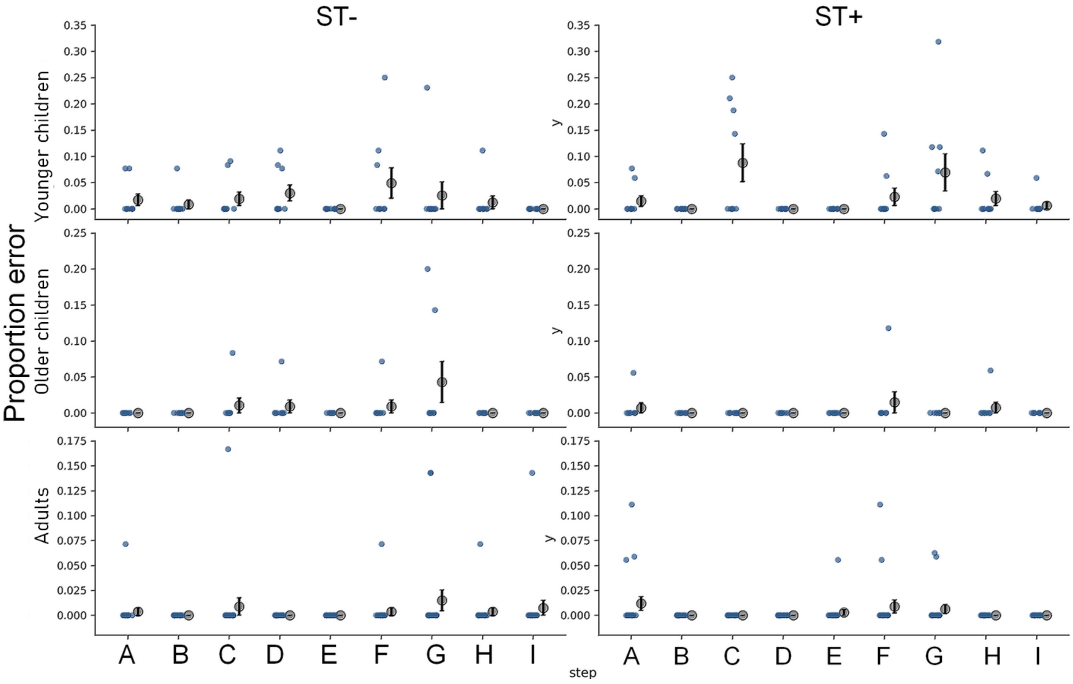


Fig. 5.12: **Proportion of step errors in sessions 3 and 4 across age groups, without secondary task (ST_{-}; left) or with (ST_{+}; right). Grey dots indicate the average across participants, error bars indicate the s.e.m. and individual points represent individual participants. N.B : the y-scales differ across rows (i.e. age groups).**

Table 5.9: **Paired-sample t-tests comparing With secondary task against errors Without secondary task, within each step and each age group, all in sessions 3&4. The degrees of freedom are 8 for each children's group (*7-8 and 9-11*), 17 for the adults' group.**

Step	A	B	C	D	E	F	G	H	I
7-8									
test-statistic	-1.000	-1.0	2.266	-1.961	Both 0	-0.861	0.901	0.371	1.0
p-value	0.347	0.347	0.053	0.086	Both 0	0.414	0.394	0.720	0.347
9-11									
test-statistic	1.000	Both 0	-1.000	-1.0	Both 0	0.319	-1.486	1.000	Both 0
p-value	0.347	Both 0	0.347	0.347	Both 0	0.758	0.176	0.347	Both 0
Adults									
test-statistic	0.994	Both 0	-1.000	Both 0	1.0	0.654	-0.739	-1.000	-1.0
p-value	0.334	Both 0	0.331	Both 0	0.331	0.522	0.470	0.331	0.331

In the without secondary task condition, the 7-8 group made significantly more errors than adults at step D ($t(DF = 25) = 2.83$, $p = 0.009$). There was no other significant difference between age groups without secondary task (table 5.10).

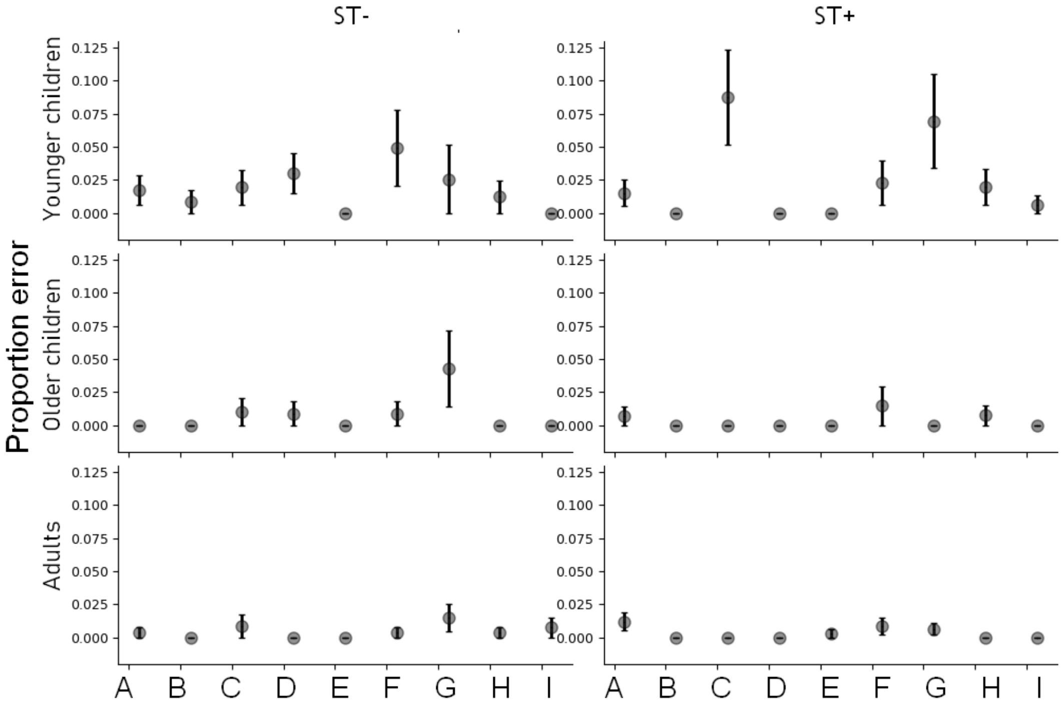


Fig. 5.13: **Proportion of step errors in sessions 3 and 4 across age groups, without secondary task (ST_-; left) or with (ST_+; right); with identical y-scales across all rows and columns. Grey dots indicate the average across participants and error bars indicate the s.e.m.**

Table 5.10: **Independent samples t-tests comparing Proportion of errors between pairs of age groups, all Without Secondary Task, in sessions 3&4, within each step. The degrees of freedom are 16 for the first row's tests, 25 for the second and 25 for the last.**

Step	A	B	C	D	E	F	G	H	I
7-8 vs 9-11									
test-statistic	1.512	1.0	0.639	1.284	Both 0	1.386	-0.343	1.0	Both 0
p-value	0.15	0.332	0.532	0.217	Both 0	0.185	0.736	0.332	Both 0
7-8 vs Adults									
test-statistic	1.358	1.443	0.634	2.83	Both 0	2.187	0.414	0.816	-0.7
p-value	0.187	0.161	0.532	0.009	Both 0	0.038	0.683	0.422	0.49
9-11 vs Adults									
test-statistic	-0.7	Both 0	0.0	1.443	Both 0	0.503	0.941	-0.7	-0.7
p-value	0.49	Both 0	1.0	0.161	Both 0	0.62	0.356	0.49	0.49

When comparing errors between age groups this time with secondary task (table 5.11), there were differences at Step C, G and H that were not present without secondary task. At step C, the 7-8 group made significantly more errors than the 9-11 group ($t(DF = 16) = 2.45, p = 0.026$) and than adults ($t(DF = 25) = 3.53, p = 0.002$). At step G and H, the 7-8 group made significantly more errors than adults (step G: $t(DF = 25) = 2.46, p = 0.021$, step H: $t(DF = 25) = 2.10, p = 0.046$).

Table 5.11: **Independent samples t-tests Proportion of errors between pairs of age groups, all With Secondary Task, in sessions 3&4, within each step. The degrees of freedom are 16 for the first row's tests, 25 for the second and 25 for the last.**

Step	A	B	C	D	E	F	G	H	I
7-8 vs 9-11									
test-statistic	0.753	Both 0	2.447	Both 0	Both 0	0.463	1.957	0.877	1.0
p-value	0.462	Both 0	0.026	Both 0	Both 0	0.65	0.068	0.393	0.332
7-8 vs Adults									
test-statistic	0.205	Both 0	3.531	Both 0	-0.7	0.907	2.463	2.099	1.443
p-value	0.84	Both 0	0.002	Both 0	0.49	0.373	0.021	0.046	0.161
9-11 vs Adults									
test-statistic	-0.567	Both 0	Both 0	Both 0	-0.7	0.289	-1.02	1.443	Both 0
p-value	0.575	Both 0	Both 0	Both 0	0.49	0.775	0.317	0.161	Both 0

Table 5.12 presents the details of the actions made by participants in the erroneous steps, and their classification among anticipation/omission (of a subgoal), omission (of a single step), substitution, and perseveration. This table presents the raw number of errors rather than proportion, given the small numbers.

Figure A.2 in the appendix shows the step-level action selection latencies (with or without individual participants data points). The corresponding linear mixed model regressions at each step are presented in the appendix (table A.2 until table A.9 in the appendix). Figure A.3 in the appendix shows the step-level action selection latencies in the trials with the 'spoon-sprinkles' goal only.

Table 5.12: **Types of erroneous actions.**

				N 7-8		9-11		Adults	
	Correct action at this step	Erroneous action taken	Classification	ST_{-}	+	−	+	−	+
A	cup to ice cream machine	cup directly to center, before going to ice cream machine	step omission	1	1		1	1	1
		click lid then realize mistake and release		1	1				1
C	lid to centre	spoon to Nutella while Nutella lid still on	step omission		2				
		spoon to sprinkles while Nutella lid still on	step omission + substitution	1	2				
		wafer to centre	anticipation/omission		6	1			
D	spoon to nutella	spoon to sprinkles (wrong content, from a later subgoal)	substitution content	1	1				
F	place spoon back to its slot	bring spoon back on Nutella	perseveration		1			1	3
		wafer to centre (before discarding spoon)	anticipation/omission	5	3	1	3		
G	wafer to centre	spoon to wafer (should be picked up without implement)	substitution				2		
		next subgoal (before wafer)	anticipation/omission	3	2 18	5		3 1	2
		next subgoal AND substitution (content of alternative goal)	anticipation/omission + substitution		2				
H	fork/spoon to banana/sprinkles ('get topping')	fork to sprinkles	substitution	1	1			1	
	placing topping to the middle	putting spoon back to its slot on the side (instead of middle)	action addition			1	1	2	

5.2.4 Discussion

5.2.4.1 Task validity

First, the newly-developed task appears to have the right level of difficulty and duration (number of trials per session and number of sessions) for our age range, and to be engaging enough so that a good number of participants completed all four sessions of the study. The following sections will discuss 1) the task validity, by verifying behaviour that can be expected based on previous studies on adults routine action, and 2) what the task reveals about changes in routine action across development.

5.2.4.2 Task acquisition

Our study is concerned with routinised action selection, and we start by overviewing trial duration and goal-level performance across sessions. Although age groups show different baseline trial duration, there is a significant drop in response time from the first two sessions to the last two, and this in each group. While it is not possible to rule out that additional sessions would see a further decline in response time, response times in session 4 are not qualitatively shorter than in session 3. This evolution of trial duration across sessions supports the proposal that learning occurs during the first two sessions, and that the last two can be taken to study *(partially) routinised* action control in this experiment.

The Proportion of Incorrect Goal Execution shows a small improvement between the first two and the last two sessions in 7- to 8-year-olds. However, there is no significant difference in the proportions between groups, although there seems to be a slight trend for children to make more Proportion of Incorrect Goal Execution overall than adults. Moreover, Incorrect Goal Execution errors are infrequent overall. This suggests that completing the main goal of the task (either keeping track of the trial's goal cue - the background colour, and/or retrieving the correct mapping between the goal's cue and the appropriate sequence) is not significantly more challenging for children than adults. Thus, the task seems appropriate to study children's control of the sequences. No further pattern appears on this variable and we focus on the errors at step level in the following section.

5.2.4.3 Routinised action control

Error distribution across steps

When looking at all age groups and conditions together, it is apparent that the errors are not randomly distributed across the sequence. A group of three steps show remarkably few errors: *placing ice cream to the centre, placing Nutella onto the ice cream, and placing the final topping onto the ice cream*. These three steps are all within-subtask actions. However, the rarity of errors at such steps may not simply be due to their within-subtask location in the task. Indeed, the *get spoon*, that is also located within-subtask in the Nutella subtask, shows relatively more errors. Given the small number of errors overall, these comparisons should be taken with caution.

These steps (*placing ice cream to the centre, placing Nutella onto the ice cream, and placing the final topping onto the ice cream*) however all have the specificity of achieving the subgoal they belong to (by updating the state of the centre). They constitute, according to Schwartz et al. (1991), *crux* actions. This indicates that participants use a goal-directed, hierarchical representation of the sequence when executing the task. That is, instead of a chain of single actions, actions are grouped in terms of subgoals (e.g., the subgoal *prepare ice cream* with the sub-actions: *get cup, fill the cup with ice cream and place to the centre*; or the subgoal *add Nutella*, with the sub-actions: *remove the lid, get the spoon, get the Nutella, drop onto the ice cream, discard the spoon*). We must note that the *drop the Nutella onto the ice cream* is not the last action of that subgoal, but is still the one that visibly achieves the goal, since *discarding the spoon* can be seen as peripheral to that goal.

When broken down by age, the same structure appears in the data of the three age groups (fewer errors at the *crux* actions). This suggests that even the youngest children of the sample use a hierarchical representation of the task like adults do. This is consistent with Schröer et al.'s (2021) results suggesting that from 5 years of age, children are able to follow hierarchical goal structures.

Age and secondary task effects

When looking at step-level errors summed across steps, younger children made more errors overall than the two older age groups. The errors that were more common when looking at all groups combined seem to be the errors most common in the younger children's group.

The effects of the secondary task were not significant in any age group per condition case. In all age groups, the lack of significant difference with and without secondary task suggests that actions may have been well-learnt, and thus could be executed without disruption even under cognitive load. It might also be that errors are so rare in general that the study is unable to detect significant changes, while response time data might be able to capture more differences. Further studies might therefore confirm the trend for the secondary task to interfere with the Get lid step, by measuring errors or by being designed to make controlled response time comparisons. If the trend is confirmed, it would indicate that *Getting the lid* being an 'enabling' action (that prepares the subgoal of *getting the Nutella onto the ice cream*), it is less central thus more susceptible to errors.

When comparing age groups together under cognitive load, younger children struggled more than one or the other age group in three steps: *Getting the lid*, *Getting the wafer* and *Getting the spoon or fork*. The step *Getting the spoon or fork* constitutes a branch point. The more frequent errors at this branch point in younger children compared to adults, together with the higher proportion of incorrect goal execution in younger children compared to older children, suggests that younger children may struggle to keep the overarching goal sufficiently activated until the selection between the spoon or fork. This would be consistent with the theories of *competitive queuing* (e.g., Houghton, 1990), according to which the actions in a sequence receive top-down activation from a *goal state unit*, initialised at the start of a sequence, and such activations decrease across later steps. Humphreys et al. (2000) found that errors in normal subjects under secondary task conditions could be explained by the decrease of activation of actions resulting from the decreasing gradient of activations from the goal state unit across steps. Therefore, the *spoon or fork* step may be more susceptible of errors in younger children because: (1) the step constitutes a branch point where there might be competition between two compatible actions, that requires more intense top-down activation (compared to other steps) from the goal unit to link the ongoing goal's cue to the appropriate action, and (2) due to the position of the action late in the sequence, and combined with the presumably lower supervisory control abilities in younger children (compared to older children or adults), this top-down activation may not always be sufficient to guide the appropriate action selection at this step.

The *Getting the lid* and *Getting the wafer* errors were also relatively highly frequent although they do not constitute a branch point like the *Getting the spoon or fork* step. As mentioned earlier, the enabling nature of the *Getting the lid* action may make it less central thus less susceptible to errors even without secondary task, and more clearly under secondary task in younger children. Likewise, the subgoal *Getting the wafer* may be considered less central to the overarching goal of making an ice cream and consequently may be forgotten particularly often. The two types of errors may be due to a tendency, under cognitive load, to use memory scripts based on real-life experience (where adding wafers on ice creams is uncommon). It might also be linked to the temporal proximity of the *Wafer* subgoal with the final subgoal, that constitutes a branch point. Indeed, during the steps preceding the branch point, participants may anticipate the choice between two alternatives (fork-banana or spoon-sprinkles). They may prioritize, at the expense of lower-level goals, the higher-level goal which is differentiated by the last subgoal. Consequently, participants may over-activate the last subgoal representation, leading to the anticipation/omission of the Wafer subtask. This explanation of the high errors in younger children at the Wafter subtask is incompatible with previous explanation, where errors at the subsequent branch point were attributed to the under-activation of the *goal state unit*.

Another explanation may be that the Wafer step belongs to the Wafer subgoal and differs from the other subgoals in that it is a single-step subgoal. Being a single action, it does not benefit from associative support from other actions (that would belong to the same subgoal). For that reason, it can be harder to remember. Either way, the cognitive load or divided attention produced by the secondary task reveals that some of the subgoals' representation may be more fragile in the 7- to 8-year-olds group than in adults.

Methodological considerations

Selection latencies were too ambiguous to be interpreted. The succession of certain actions and the placement and boundaries had sometimes different starting points (centre, or side). Consequently, the travel distance on screen limits the possible comparisons. A pair of actions could have constituted a branch point (BP) and a comparable non branch point: selecting between the spoon and the fork for the last subgoal (BP) and selecting the spoon for the Nutella subgoal (nBP). However, a non-negligible number of errors occurred in the step preceding the branch point (wafer), adding noise that may impede the direct comparison of BP and nBP latencies. We therefore adapted the task in the subsequent experiment to allow such action latency comparisons and interpretations.

5.2.5 Interim conclusion

In conclusion, all age groups appear to use a hierarchically-structured representation of the task sequences. Instead, it might have been that the younger group used a flat task representation, i.e. with all actions equal. This would have led to a uniform distribution of errors across steps. But the opposite was observed. Even the youngest children structure their acquired task representation, presumably by bringing to bear knowledge either from other experience (e.g., making other foods) or by identifying enabling, crux and clean-up actions and chunking actions into super-ordinate mini-sequences.

Within this non-flat task representation, 7- to 8-year-olds seemed to have certain action representations weaker than others, resulting in errors at certain steps under cognitive load. This is consistent with the view that supervisory control is still developing throughout the school age years and thus may bring insufficient top-down control at certain challenging steps. This will be investigated further in the next experiment by making use of response times and kinematic measures.

5.3 Experiment 4: Studying Branch Points to Probe the Supervisory System

5.3.1 Introduction

This experiment explores further the role of the supervisory system in interaction with the routine system. It does so by circumventing a limitation of experiment 3, which was that the response time difference between points would be difficult to attribute either to the pure organisation of the task (placements of icons and succession of actions) or to the cognitive processes of interest.

In experiment 3, the distribution and types of errors informed us about the hierarchical nature of the representation used for sequence control, in particular at subtask boundaries. To be able to dissociate and measure clearly the involvement of the supervisory system and routine system, we focus here on the comparison of latencies at two types of points. The first type, branch points (BP), is hypothesised to require maximal modulation from the supervisory system. The second type, non branch points (nBP) is hypothesised to require less modulation from the supervisory system (or in case of strong routinisation, no modulation at all), presumably due to the absence of alternative compatible actions.

5.3.1.1 The GoGelato task version 2

The only conceptual change between the present task and the previous one is the composition of the action sequences. In experiment 3, it seemed that the frequent forgetting of the step ′discard the spoon′ and/or the step ′adding the wafer′ resulted in large variability of behaviour at the action ′select the spoon′ or ′select the fork′. Indeed, these steps were placed earlier in the sequence than the spoon or fork step, possibly adding noise which was not there when selecting the spoon in the Nutella subgoal. In the current task, we adjusted the action sequences in order to reduce the noise and make possible a cleaner comparison of the branch point′s and the control, non branch point′s, latencies. Specifically, we deleted the ′discarding the spoon′ and the ′adding the wafer′ steps.

Following these changes, the differences that characterised the BP and nBP steps were that the nBP were within-sequence and had a 100% transition probability from their preceding action, while the BP were between-sequence and had a 50% transition probability.

5.3.1.2 Stop-signal task

To link the supervisory control in action sequences with executive functions (EF) measured by standard tasks in the developmental literature, we also administered a classic EF task. The task was the stop-signal task (Logan, 1994), held to measure specifically the inhibition component of EF. This was done for the adult sample only, since including it would make the study too long for children. Nonetheless, the results in adults would be informative for development: if inter-individual differences in EF measured in adults correlate with action sequence performance, then this would give indirect evidence for the changes with age in the link between executive functions and action sequence performance in children.

Performance on the stop-signal task has been found to indicate age-related changes in inhibitory control (Carver et al., 2001, Williams et al., 1999). Given that inhibitory control is a component of EF, this task was aimed to be a partial indicator of the role of executive functions as measured by a standard test. We are interested both in the inhibition component and general executive functioning measured by the stop-signal; indeed performance on various EF tasks is known to be strongly intercorrelated (Miyake et al., 2000b), and separating components would have required a battery of other tasks which would have made the study too long even from adults.

The inhibition component was chosen because we expected it to be important in this specific task (although, as just mentioned, we believe other EF components may play an important role as well). There is sparse evidence demonstrating which role inhibition plays in action sequence control. In Sebold et al. (2016), adults with lower inhibitory control were argued to rely less on the hypothesised model-based system in the two-stage task (similar to the task of Decker et al. (2016) that we used for Chapter 3). In contrast, within the theoretical perspective adopted here, we assume that at BPs the action from the inappropriate subtask must be inhibited to allow selection of the appropriate action. Thus, stronger inhibitory control should correlate with a reduced cost at the BPs, as will be developed in the next Section.

5.3.2 Hypotheses and predictions

5.3.2.1 Hypothesis 1 (H1): Action selection at BPs is slower than at nBPs

Our main hypothesis, and a pre-requisite for the subsequent hypotheses, is that action selection at branch points (BPs) takes longer than at non branch points (nBP). This is because branch points are thought to rely more on supervisory resources than non branch points do. This has been shown in adults (e.g., Ruh et al., 2008, Ruh et al., 2010), and we expect it to be the same in children. Alternative accounts would be that children's action control does not follow the same principles as for adults. One could imagine, for example, that the representation of the primary sequence production is flat (i.e., non-hierarchical) and that all actions are conceptually equivalent. This would imply that selection latencies should be identical at BPs and nBPs.

5.3.2.2 Hypothesis 2 (H2): The secondary task affects latencies

We expect that introducing a secondary task will lengthen the performance of the sequence production task (assuming participants are indeed performing both tasks). Indeed, a secondary task adds demands to higher-level resources that are normally allocated to the sequence execution. Specifically, the overall trial duration with secondary task should be longer than without secondary task. How the secondary task affects specific points or not is the topic of the next hypothesis.

5.3.2.3 Hypothesis 3 (H3): The secondary task affects BP latencies more than nBP latencies

We hypothesised that in both adults and children, action selection would be more slowed by a secondary task at BPs than at nBPs. As mentioned, this is because the secondary task demands supervisory resources which interfere with the supervisory resources needed at BPs, resulting in longer latencies at BPs than at nBPs. We will refer to the difference in latencies between BPs and nBPs as *BP cost*.

5.3.2.4 Hypothesis 4 (H4): The BP cost changes with age

Furthermore, given that executive functions are known to improve with age, the longer latencies at BPs than nBPs, and the extent to which the secondary task would affect action selection at different points, in particular the BP cost (hypothesis 1, 2, and 3) was expected to be enhanced in children compared to adults, and in younger compared to older children.

5.3.2.5 Exploratory investigations

There are a number of ways in which the availability and maturation of executive functions, as well as the degree of routinisation, could affect the kinematics of action selection. Based on the previous chapters, measures of deviation such as maximal deviation, and the measure of speed variability, seemed to be the most informative and least correlated to response time (which is our primary measure). We did not make precise predictions regarding kinematics but expected, for example, that the maximal deviation would increase with supervisory demands (at branch points and/or under cognitive load), and that this would be occur more intensely for children than adults and for younger children than older children. We also expected the speed variability to exhibit different patterns than response time, as was the case in the SuperCook study of Chapter 4, without a-priori hypotheses on this pattern.

5.3.2.6 Relationship between stop-signal measure and action sequence control variables

Finally, our predictions regarding the SSRT measured by the stop-signal task, and some key measures of action sequence control were the following. Our main prediction was regarding the BP cost. We expected a positive correlation between the SSRT and the BP cost: having a shorter SSRT would reflect better executive function skills, and this would correlate positively with the BP cost. Indeed, a smaller BP cost would reflect less difficulty in action execution at BP. On the contrary, the BP latency were expected to show a negative correlation with the SSRT. For the Coefficient of Variability of speed CoV, we expected a positive correlation. That's because we expected lower CoV to reflect a better routinisation degree in the task and to be associated with lower SSRT. We also explored the correlation between SSRT and the Proportion of incorrect goals, for which we expected a positive correlation. Indeed, we expected fewer errors in the task to be associated with lower SSRT.

We can think of the SSRT broadly as a measure of EF. Alternatively, we can consider the particular role of the inhibitory control component of EF, as follows. Both for the *proportion of incorrect goals* and $DeltaBP - nBP$, the assumed role of inhibitory control (measured by the stop-signal task) is that one must inhibit the representation of the incorrect subgoal in order to select the proper action of the subgoal. Higher inhibitory control abilities would result in fewer errors (performing the wrong subtask hence achieving an incorrect goal), and/or (at least) reduced slow down at the time of selecting the action at a BP.

5.3.3 Methods

We will outline the aspects of the present task's version that differ from the task described in Section 5.2- all the rest were identical. The key conceptual changes in the present task's version are the nature and number of actions that make up the action sequences. On top of this, minor changes were introduced to improve the reliability and interpretability of the results.

5.3.3.1 Participants

All adults in this study were recruited via Prolific, and all the children via Birkbeck's Babylab database. 50 adults (N females = 25) and 69 children (N females = 33) took part in the experiment. Not all participants completed all sessions, and the total numbers of participants per age, as well as the number of sessions they completed, are presented in table 5.13.

Some sessions completed by child participants were lost due to server communication bugs. We were however confident, thanks to timestamps, and despite the missing data, about when the sessions had been performed. We were able to include session $i + 1$, knowing that session i had been performed, even if the data from session i was missing. The data of 12 sessions, performed by 9 different participants, are missing due to this bug (8 sessions number 1, 2 sessions number 2 and 2 sessions number 3). Importantly, this loss of data does not appear systematically more in a given age group (three 7-year-olds, one 8-year-old, three 9-year-olds and two 10-year-olds) which meant we could keep the other sessions from these participants. Another bug happened for some participants where, instead of doing a given session, they were brought to the next one. In the cases where, at the time of doing session 2, participants reached the study for session 3 (meaning they got the secondary task too soon), that session and all the following ones were be removed from the data. Because of this bug, 8 sessions from 4 different participants were discarded (from one 7-year-old, one 8-year-old and two 10-year-olds).

The two aforementioned bugs never occurred in the adult sample. The bugs were probably caused by a wrong key combination, by refreshing or closing the web browser at the wrong times. We assumed that the reason why the bugs did not happen in the adult's sample is because the adults were all tested on Prolific, where they are redirected to the Prolific platform after the study (and are understandably more experienced with online studies and presumably careful with proceeding until the final page).

Additionally, 3 sessions each from different participants (aged 7, 8 and 12) were deleted for having too few trials. However, the sessions that followed them were kept due to participants' having completed half of the preceding session (if we consider the demonstration trial) and the concerned sessions and ages not seeming to have a pattern that would bias the results (session 3 of one 8 y.o, session 2 of one 8 y.o. and session 1 of one 11 y.o.). Participants whose last session had fewer than 6 trials, and who did not attempt to resume the study afterwards, had their last session not included; their count is merged with the counts of participants who stopped at the prior session. The number of sessions per age is presented in Table 5.13.

The analyses ensure that equal weight is given to each participant regardless of the imbalance in the number of sessions available.

Table 5.13: **Sample sizes, broken down by group (adults and children), and children's group further broken down by age (the integer number indicate age in years).**

N (samples sizes by group)	Adults	Children	7	8	9	10	11	12	13
Took part (started Session 1)	50	69	15	12	14	14	4	8	2
Included (completed S1)	50	69	15	12	14	14	4	8	2
Completed at least S1 & 2	44	51	11	9	11	8	4	7	1
Completed at least S1, 2 & 3	37	47	11	9	11	6	3	7	0
Completed S1, 2, 3& 4	27	42	10	7	10	6	2	7	0

To simplify the analyses, we aimed to create two groups of children via a split around the median age. Given the large discrepancy in the participants in session 1 ($N = 69$) and in session 4 ($N = 4$), and given that our main hypotheses will be tested in data from sessions 3 & 4, we carried out the median-split based on the participants who did at least session 3. The resulting median was 9 years old, with $N = 20$ children between 7 and 8 years old (the group will be denoted as *7-8*), and $N = 27$ children between 9 and 12 years old (the group will be denoted at *9-12*). The adults' group will be denoted as *ADU*.

As explained in Section 5.3.1.2, only adults took part in the stop-signal task. Adults were compensated the same amount as in the previous study, except at session 4 where they received £4 because they performed the additional stop-signal task.

5.3.3.2 Materials

Novel action sequence

The present version of the game had a different sequence with 7 steps from that used in Experiment 3, instead of 9. Apart from introducing a knife icon, the other icons, their dimension and start locations were identical to the previous version icons as described in Section 5.2.2.2. The number of slots also remained the same since the wafer icon was no longer present. The sequence of actions to be carried out are listed and illustrated in Figure 5.14.

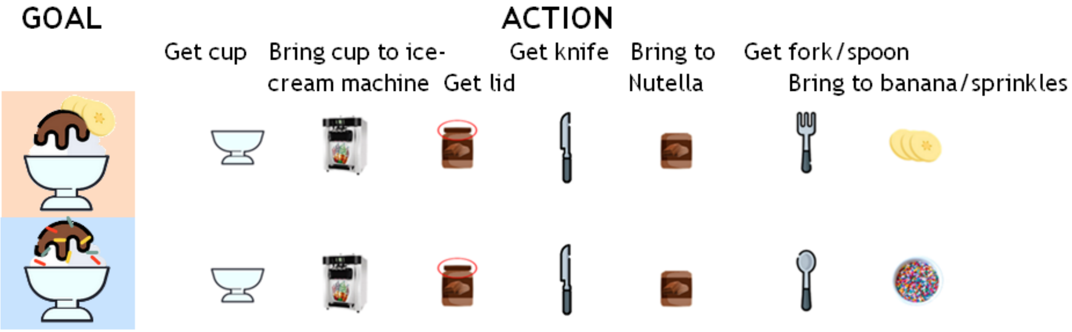


Fig. 5.14: **Action sequence: the 7 main actions of the game.**

The actions 'Get cup', 'Get lid', 'Get knife' and 'Get fork or spoon' all started from the centre of the screen. The actions 'Bring cup to ice-cream machine', 'Get lid', 'Bring to Nutella' were all followed by bringing the corresponding object (that is, the cup filled with ice cream, the lid or the knife with Nutella) to the central location. The 'Get knife' action constituted the 'reference' non branch point against which the *branch point (BP)* could be compared. We call it *non branch point (nBP)* hereafter. The BP was the 'Get fork' or 'Get spoon' (according to the sequence) action.

Within this novel sequence, our BP (branch point) and nBP (non branch point) were comparable in terms of the distance to be travelled for selection (from the centre to a peripheral slot, the slots being all equidistant to the centre) and type of object (implement: fork or spoon at BP, knife at nBP) with comparable colour, broad shape, and salience, while being distinguishable enough from each other. The first version of the study had intentionally the exact same object (spoon) at BP and nBP. This was done to ensure that the differences observed would be uniquely due to the role of the action in the sequence (BP or nBP) and not to the object itself. A disadvantage was that the same item was consequently used twice (while all the others were used once), which could lead to biased latencies (either by facilitating the visual search for the object, hence the selection latencies, or on the contrary by triggering avoidance of a previously visited location). Controlling for this would have required designing a version where the relative order of BP and nBP are reversed in some participants, thus requiring a larger sample size which was not achievable in the project time frame.

Minor changes from the previous task version

In the version of the task used in Experiment 3, the sequence of trial goals (indicated by the background colour) was pseudo-randomised by shuffling a sequence of 8 trials of each type. We added a constraint to this sequence randomisation: the resulting sequence had to have fewer than 3 successive identical trials in a row. This was done to ensure the background colour cue would not risk becoming irrelevant and be ignored.

At the end of the study, we added a question page asking the participants if they had technical problems (e.g., game getting blocked or extremely slow, which could occasionally happen depending on the internet connection). We did not have to exclude participants for that reason and thus will not mention it further.

The 'buzz' sound which accompanied incorrect actions in the first version was reported by some to be frustrating. Given that the feedback for the incorrect nature of the action was clear visually (with the icons moving back to their original location), we removed the sound.

In the version of Experiment 3, the bird chirp count sequence in a given trial was pre-programmed (randomly) regardless of the time taken by participants. This had the risk of occasionally having the sound sequence end relatively early in a participant's performance (in particular for a slow participant). To prevent this, in this version we programmed the sounds to be contingent on the participant's progress (i.e., programmed it relative to the actions completed). The sounds' occurrence was still random (randomly spread among the actions, following a given action by a random interval of between 100 and 900ms). We verified via informal piloting that the occurrence of sounds was unpredictable. The pilot participants reported not having noticed that the sounds were contingent on their performing the actions.

The Stop-signal task

We used the stop-signal task developed and made available by Verbruggen et al. (2019), online with JavaScript and jsPsych.

The task, whose sequence of events is presented on Figure 5.15, included Go and Stop trials. *Go trials* required participants to respond to an arrow by pressing the key with the arrow pointing in the same direction. A minority of trials were *Stop trials*, starting the same as Go trials but where the arrow turned red after a variable Stop Signal Delay (SSD), in which case participants were required to refrain from responding.

A trial started with a fixation sign, replaced by the stimulus after 250ms. The stimulus disappeared after a participant's response or after 1,250ms. The interstimulus interval was 750 ms. The proportion of Stop trials was one in four.

On Stop trials, the stop signal was presented after a variable Stop Signal Delay (SSD). The SSD started at 300 ms and was adjusted continuously with a staircase procedure: after a correct Stop trial (i.e. response inhibited), the SSD increased by 50 ms; if incorrect, the SSD decreased by 50 ms.

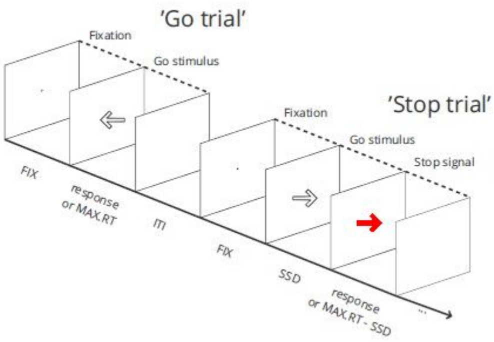


***Fig. 5.15:* Sequence of events in the stop-signal task. FIX = fixation duration; SSD = stop signal delay; MAX.RT = maximum reaction time; ITI = inter-trial interval. Participants are instructed to press the arrow key corresponding to the displayed arrow. However one fourth of the trials the arrow turns red after a certain delay (the SSD), in which case participants are instructed to suppress their response. Figure reproduced with permission from Verbruggen et al. (2019), licensed under CC BY 4.0.**

5.3.3.3 Procedure

The procedure was almost the same as for Experiment 3, described in Section 5.2.2.2. The only difference was the addition of the stop-signal task, that participants completed at the end of the last (fourth) session. This task included a practice phase (one block of 32 trials), with feedback, and an experimental phase (4 blocks of 64 trials). More details can be found in the task description provided by Verbruggen et al. (2019).

5.3.3.4 Analyses

For trial-level variables, we analysed all four sessions. For step-level variables, we analysed the last two sessions, unless otherwise specified.

For trial-level variables, we performed simple comparisons. For the step latencies at BP and nBP, as well as for the kinematic variables, we performed more in-depth analyses, with both non-parametric tests and linear mixed models. Consistently with the last experiment, Wilcoxon signed-rank tests were carried out on response time variables for within-participant comparisons and Mann-Whitney U tests for between-participants comparisons. For proportion variables, paired and independent samples t-tests were carried out. The four types of test were carried out with the Python module *scipy.stats*.

For the main variables of interest (step latencies and kinematics), we carried out both non-parametric and Linear Mixed Models (LMM) as described further below. We did so because both methods have some limitations with this dataset, hence using both and noting where they agree with each other would provide more reliable evidence than one or the other alone. A limitation of non-parametric tests is that they require summarising the data and lose some information. For example, the medians at participant-level were used in the non-parametric tests, because the number of trials varied by participants depending on whether they stopped at the third or fourth sessions. This ensured that participants who completed more sessions were not weighted more strongly than the ones who did not; however, this could lead to other distortions. A limitation of the Linear Mixed Models is that although log-transformation can significantly reduce skew in the data, the response time and kinematics data can still deviate from normality. Nevertheless, they are a standard way to test the effects and interactions of all independent variables together. The stance taken is that combining both approaches compensates for the shortcomings of the alternative approach, and that converging evidence from both will be reliable.

The figures represent the raw (non log-transformed) variables' median, with the error bars representing the inter-quartile range (IQR), unless otherwise specified. The Linear Mixed Models were however run on log-transformed data to remove skewness. When a Linear Mixed Model (LMM) was run, the plots of the estimated marginal means (predicted values) with confidence intervals are available in the appendix.

Definition of latency variables

Regarding latencies, our variables of interest were the portion of the trajectory that started from the central location (from the moment when the previous icon was dropped onto the centre) and that ended when a peripheral icon was selected (clicked on). To further understand how much the amplitude of effects (Secondary Task: ST, or Branch Point: BP) varied with age and condition, we computed the difference in latencies with and without Secondary Task, $Delta\ Latencies\ (ST_+ - ST_-)$ and the difference in latencies at BP compared to nBP, $Delta\ Latencies\ (BP - nBP)$. The latter quantifies the *BP cost*.

Multiple comparison corrections

In this chapter, we are interested in several dependent variables and their change across many independent variables. Our use of hypothesis-testing is motivated by several goals: it helps us confirm or disconfirm if the data is in line with our theory (when it comes to pre-defined predictions), but also to decide whether to treat the data as noise or if the data is worthy of further investigation (when it comes to exploratory hypotheses). For exploratory analyses, such hypothesis testing is helpful to identify the strong patterns and ignore the weak ones, given the large number of DVs and IVs.

We use a conservative approach by correcting for multiple comparisons in the LMMs. We use the False Discovery Rate (FDR) correction. We use an alpha level of 0.05 to interpret pre-defined hypotheses with sufficient confidence (i.e., significant results mean the data is in line with our hypothesis). Furthermore, to fully exploit the data and make sure that no interesting insights are left unconsidered, FDR-corrected p-values that fall below 0.1 will be mentioned as 'non significant but notable'.

Mixed models

The Linear Mixed Models (also known as multilevel models or linear mixed-effect regressions) were conducted with the R package nlme (Pinheiro et al., 2020), using the restricted maximum likelihood (REML) approach. For the step latencies in all four sessions, without secondary task, we included the following fixed effect factors: branch point (categorical, levels: BP, nBP), session (categorical, levels: $S1\&2$, $S3\&4$) and age group (categorical, levels: 7-8, 9-12, adults). For step latencies in sessions 3&4, we included: branch point (categorical, levels: BP, nBP), secondary task (categorical, levels: ST_{-}, ST_{+}) and age group (categorical, levels: 7-8, 9-12, adults).

For $DeltaBP - nBP$, we included secondary task (categorical, levels: ST_{-}, ST_{+}) and age group (categorical, levels: 7-8, 9-12, adults). For $DeltaST_{-} - ST_{+}$, we included Branch point (categorical, levels: BP, nBP) and age group (categorical, levels: 7-8, 9-12, adults). In all models, we also included all the possible interaction terms, and defined a random intercept per participant to account for the repeated measures (Steele, 2014). Because the age group is a categorical predictor with three levels, the LMMs ran on the entire dataset performed by default the comparison of the two children's group against the adults' group (baseline group). To compare the two children's groups to each other, we additionally ran each LMM on a subset of data containing the children's data only.

Kinematic variables

We used the kinematic variables that have been introduced in Chapter 4 (Average absolute orthogonal distance to straight path, Log-ratio maximal relative distance to end point, and Variability of speed within a trial) and that had proven to be informative there. An early investigation of Variability of speed revealed that the large baseline differences between the speed in children and adults greatly affected the scale of the variability of speed. To tease out this difference in speed (which is already captured by the latency variables) from the variability component itself, we normalised the variability of speed by the average speed. We will call this variable Coefficient of Variation in speed (or simply Coefficient of Variation, CoV). In other words, we defined: $CoV = std_{speed}/average_{speed}$.

We decided to focus on variables that had limited correlation with Step latencies, so that the kinematic analyses would bring complementary information. All three kinematic variables were significantly correlated with Step latency: Spearman correlation coefficient $R = 0.469$ and $p < 0.0001$ for Average absolute orthogonal distance; $R = 0.498$ and $p < 0.0001$ for Log-ratio maximal relative distance; $R = 0.288$ and $p = 0.0004$ for CoV. Although CoV was significantly correlated with Step latency, the correlation was noticeably smaller than for the other variables (0.29 against 0.47 or 0.50). It is not surprising that CoV brings different information than latencies, given that two trials with the same latency could have, for example, one a fairly uniform speed, and the other sharp changes in speed, where the slow portions would be compensated by fast portions. Consequently, we retained the CoV variable for more in-depth analyses. We performed the same LMM analysis on CoV as on step latency (with a random intercept per participant, and BP, ST and age group as fixed effects).

Stop-signal task

To analyse the stop-signal task, we used the analysis script made available by Verbruggen et al. (2019). The variable of interest was the Stop Signal Response Time (SSRT), as defined and calculated by Verbruggen et al. (2019). This variable is taken to indicate the covert latency of the stop process, according to the Independent Race model introduced by Logan et al. (2014).

We explain here briefly the idea behind the model's estimation of the SSRT. The stop-signal task is modelled as an independent race between a 'go runner' and a 'stop runner'. A go stimulus triggers the 'go runner', and a stop-signal triggers the 'stop runner'. If the 'go runner' ends the race before the 'stop runner', a failed stop trial is obtained (which is conceptualised as unsuccessful inhibition). If the 'stop runner' ends the race first, a successful stop trial is obtained. The Independent race model links the latencies of response on failed stop trials, the latencies on go trials and the probability of response on stop trials to the different stop-signal delays that are used in the task. It provides an estimate of the SSRT, as illustrated in Figure 5.16.

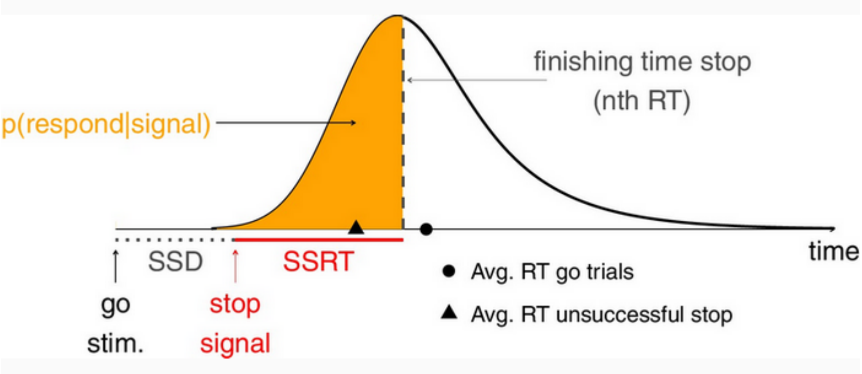


Fig. 5.16: **SSRT (Stop Signal Response Time) in the Independent race model. SSD: Stop Signal Delay. Figure reproduced with permission from Verbruggen et al. (2019), licensed under CC BY 4.0.**

5.3.4 Results

5.3.4.1 Secondary task

As a preliminary analysis, we verified that participants were not ignoring the secondary task. To do so, we looked at the average error count per participant (across sessions 3 & 4). The participant with the highest average error had an average of 1.8 errors. It was a child, hence had been presented on average with 5 sounds. We deemed the maximal error was low enough not to require excluding this participant or any participant. The participants' average error in the secondary task across age groups are presented in table 5.14.

Table 5.14: **Average error per age group in 'bird chirp' counts in the secondary task.**

	average	std
7- to 8-year-olds	0.25	0.39
9- to 12-year-olds	0.18	0.37
Adults	0.30	0.29

5.3.4.2 Task acquisition and overview of age differences

Figure 5.17 shows the trial duration across steps and conditions (with or without secondary task; for session 3 and 4) per age group. The overall trial duration was significantly longer for 7-8 year-olds than 9-12 year-olds ($Median_{7-8} = 23.4$, $Median_{9-12} = 16.1$, $U(N_1 = 24, N_2 = 37) = 734, , p < 0.0001$), which was in turn significantly longer for 9-12 year-olds than adults ($Median_{9-12} = 16.1$, $Median_{adults} = 9.5$, $U(N_1 = 37, N_2 = 50) = 1737$, $p < 0.0001$).

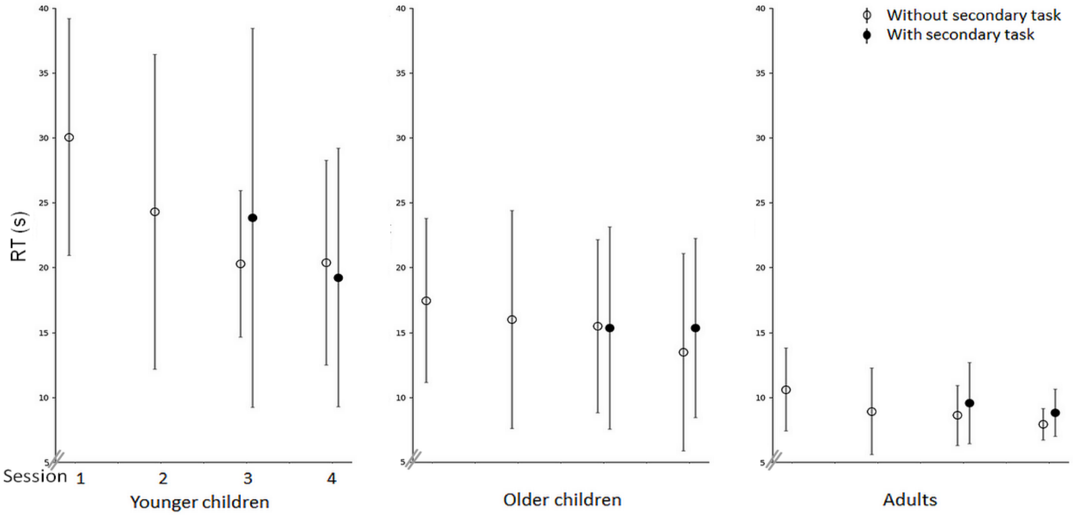


Fig. 5.17: **Median trial duration across sessions and age groups, with and without secondary task. Error bars represent the inter-quartile range. Data from all participants are included in this figure, regardless of their total number of completed sessions.**

When looking at the trials without secondary task, the sessions 3&4 grouped together were significantly shorter than 1&2, and this for all three age groups (7-8: $W(N = 20) = 18, p = 0.0005$, 9-12: $W(N = 26) = 70, p = 0.0074$, adults: $W(N = 37) = 133, p = 0.0010$; cf. table 5.15).

Table 5.15: **Wilcoxon signed-rank tests for Trial Duration comparisons between sessions (only without secondary task: ST_-) or between conditions (with secondary task: ST_+ against without: ST_-; only in sessions 3&4). The table shows: W-value (Wilcoxon test statistic), N of pairs with non-zero differences, p-value.**

	ST_-, S1&2 vs S3&4	S3&4, ST_- vs ST_+
7-8	W=18, N=20, p=0.0005	W=69, N=20, p=0.1893
9-12	W=70, N=26, p=0.0074	W=138, N=26, p=0.2205
Adults	W=133, N=37, p=0.0010	W=68, N=37, p<0.0001

To see to what extent the sessions 1&2 could be considered as a group, and sessions 3&4 as another group, we compared session 1 against 2, and session 3 against 4 (table 5.16). Session 1 differed significantly from session 2 in all age groups (7-8: $W(N = 16) = 12, p = 0.0021$, 9-12: $W(N = 24) = 45, p = 0.0018$ and adults: $W(N = 44) = 147, p = 0.0000$). The difference between sessions 3 and 4 was still significant in adults ($W(N = 27) = 53, p = 0.0011$).

Table 5.16: **Wilcoxon signed-rank tests for Trial Duration comparison of session 1 against 2, and of session 3 against 4; only in trials without secondary task (ST_-). The table shows: W-value (Wilcoxon test statistic), N of pairs with non-zero differences, p-value.**

	S1 vs S2	S3 vs S4, ST_-
7-8	W=12, N=16, p=0.0021	W=61, N=16, p=0.7436
9-12	W=45, N=24, p=0.0018	W=114, N=23, p=0.6849
Adults	W=147, N=44, p<0.0001	W=53, N=27, p=0.0011

We then looked at the difference of amplitude between the drop from session 1 to 2 and the drop from 3 to 4 in each age group (table 5.17): it was significant for both children's groups (7-8: $W(N = 13) = 14, p = 0.0266$, 9-12: $W(N = 19) = 34, p = 0.0124$), unlike for adults ($W(N = 27) = 165, p = 0.564$).

Table 5.17: **Wilcoxon signed-rank tests comparing the drop in Trial Duration from session 1 to session 2, to the drop from session 3 to session 4; only in trials without secondary task (ST_-).**

	Median drop 1-2	Median drop 3-4	W-value	N	p-value
7-8	9.11	1.32	14	13	0.0266
9-12	2.21	0.00	34	19	0.0124
Adults	0.91	0.71	165	27	0.5642

When looking at the sessions 3&4 only, the trials with secondary task were significantly longer than those without for adults ($W(N = 37) = 68, p < 0.0001$), but for the child groups the comparison was not significant. Complete results are presented in table 5.15. Nonetheless, children had visibly larger error bars than adults, and 7-8 larger errors bar than 9-12, suggesting that the ST adds considerably variability in children (and even more so for younger children) compared to adults. Furthermore, the lack of statistical significance does not imply the absence of effect of the secondary task.

There was no significant effect of session or secondary task on the Proportion of Incorrect Goal Execution (Figure 5.18 and table 5.18), except in the 7-8 group where the Proportion of Incorrect Goals decreased from S1&2 to S3&4 ($t(\text{DF=19}) = 2.33, p = 0.0309$).

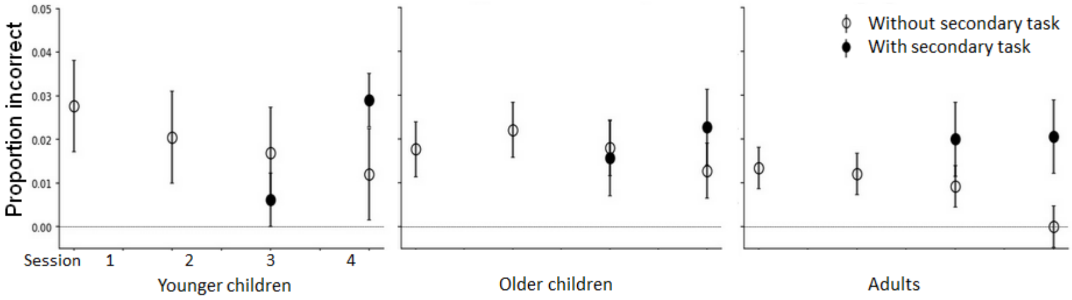


Fig. 5.18: **Average Proportion of Incorrect Goal Execution across sessions and age groups, with and without secondary task. Data from all participants are included in this figure, regardless of their total number of completed sessions.**

Table 5.18: **Paired samples t-tests on Incorrect Goal Execution comparisons between sessions (only without secondary task: ST_-) or between conditions (with secondary task: ST_+ against without: ST_-; only in sessions 3&4). The table shows: statistic (degrees of freedom) and p-value.**

	ST_-, S1&2 vs S3&4	S3&4, ST_- vs ST_+
7-8	t(19)=2.33, p=0.0309	t(19)=-0.57, p=0.5733
9-12	t(25)=1.01, p=0.3230	t(25)=1.01, p=0.3230
Adults	t(36)=1.32, p=0.1953	t(36)=-1.85, p=0.0719

When looking at the Proportion Incorrect Goals per age groups collapsed across sessions and secondary task (Figure 5.19 and table 5.19), there was no significant differences between age groups.

Table 5.19: **Independent samples t-tests on Proportion of Incorrect Goal Execution comparisons between age groups. DF: degrees of freedom.**

	test-statistic	DF	p-value
7-8 vs 9-12	0.19	59	0.8467
7-8 vs Adults	1.76	72	0.0835
9-12 vs Adults	1.39	85	0.1669

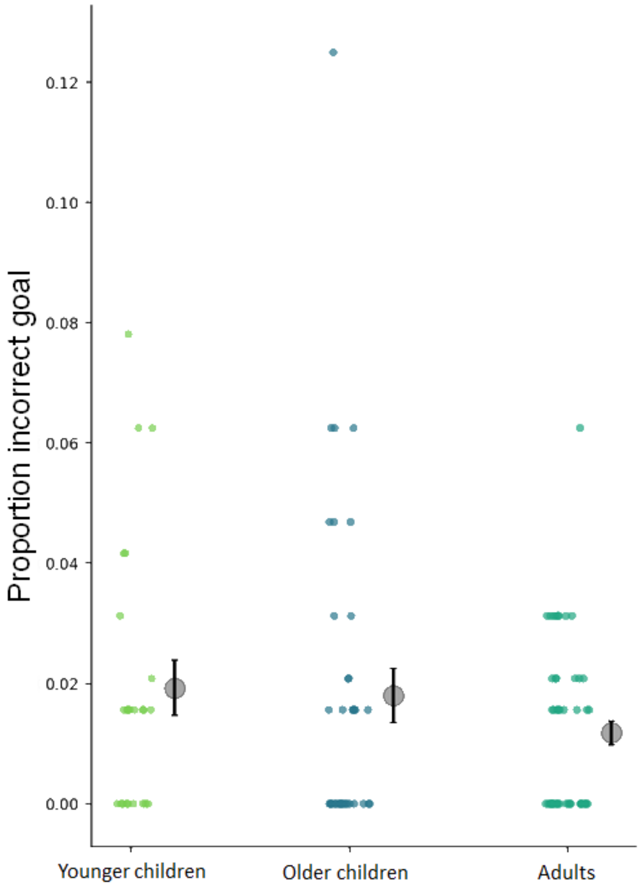


Fig. 5.19: **Average proportion of incorrect goal execution age groups (collapsed across all sessions and conditions). Error bars represent the s.e.m. and individual points represent individual participants. Data from all participants are included in this figure, regardless of their total number of completed sessions.**

We also noted that the participant with the highest Proportion Incorrect Goals performed incorrect goals less than 13% of the time. This higher bound is far from chance level, which indicates that participants were all trying to perform the task correctly by taking into account the sequence to be performed according to the background colour cue.

To give a comprehensive overview of the sequence's data, we present the durations of all the steps which had the same distance travelled (from the centre to a peripheral slot, namely the first, third, fourth and sixth steps) in Figure 5.20, averaged across the four sessions without secondary task. In the next Section, we focus more particularly on the branch point and the 'reference' non-branch point, introduced earlier, on which statistics will be performed.

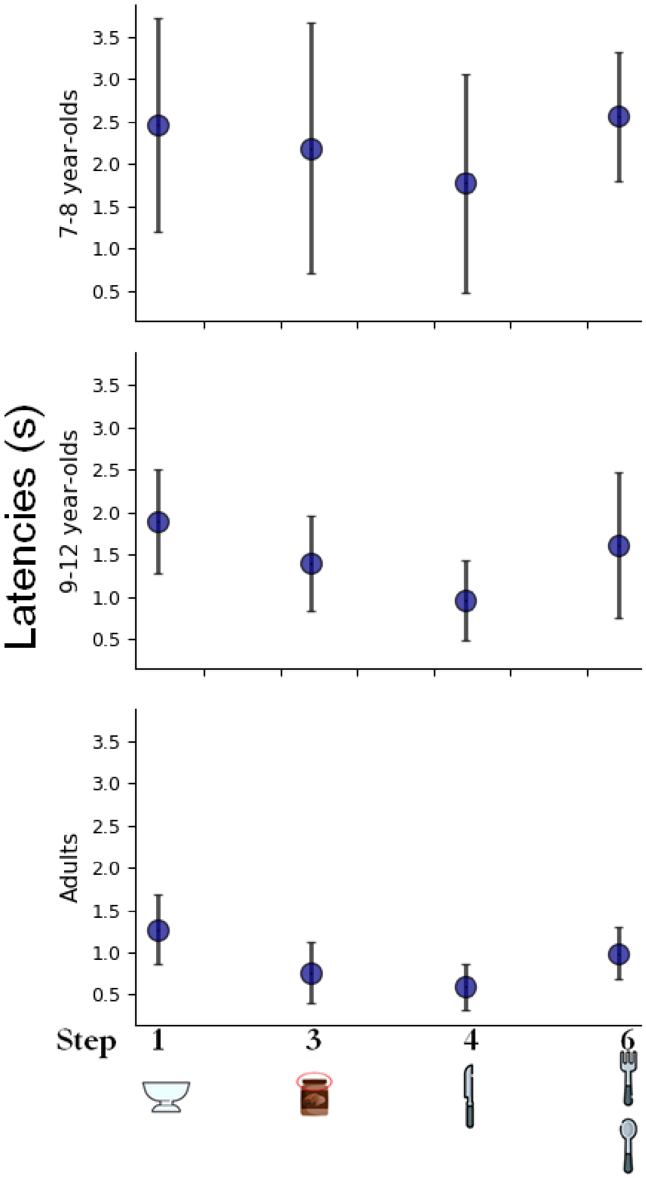


Fig. 5.20: **Selection latencies of the steps that start from the centre. Median collapsed across 4 sessions, without secondary task (ST_{-}). The data shown is the median across participants. Error bars represent the IQR (interquartile range).**

5.3.4.3 Practise effects at branch points and non-branch points

In-text, we report the FDR-corrected p-values for the Linear Mixed Models.

Sessions 1&2 vs 3&4 - Latencies

The response latencies as a function of BP and sessions (1&2 vs 3&4) and age groups (without secondary task) are shown in Figure 5.21.

The LMM (table 5.20) indicated that BP responses were significantly slower than nBP responses ($t(\text{DF}= 4278) = 13.09, p < 0.001$), 7-8 year-olds were significantly slower than adults ($t(\text{DF}= 81) = 12.85, p < 0.001$), 9-12 year-olds were significantly

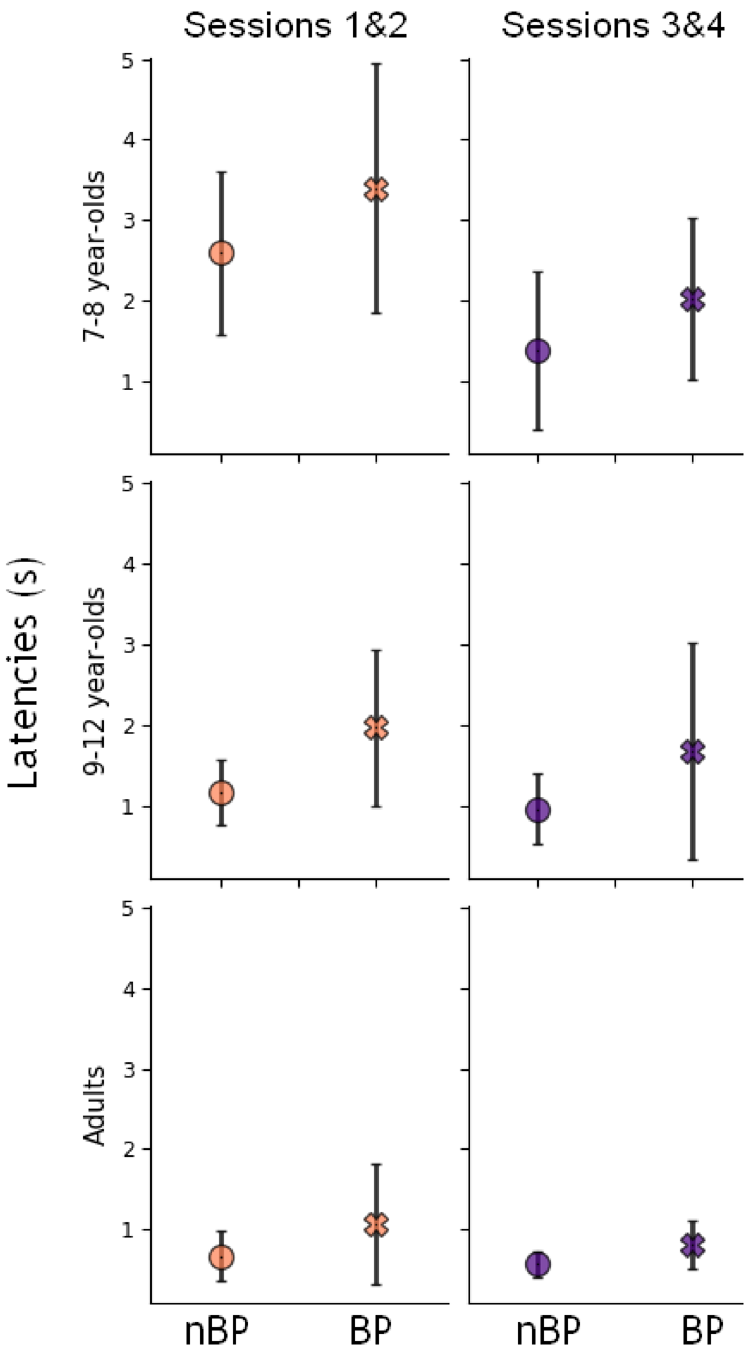


Fig. 5.21: **Step latencies at BP (branch points) and nBP (non branch points) in the first half ($S1\&2$) or second half ($S3\&4$) of sessions, without secondary task ($ST\_$): median values of participants' trials' median. Error bars represent the IQR.**

slower than adults ($t(\text{DF}= 81) = 6.47, p < 0.001$) and sessions 3&4 were 'non significantly but notably' faster than sessions 1&2 ($t(\text{DF}= 4278) = -2.01, p < 0.077$). Additionally, there was an interaction of session x 7-8 group ($t(\text{DF}= 4278) = -5.88, p < 0.001$) and of session x BP x 7-8 group ($t(\text{DF}= 4278) = 3.74, p < 0.001$).

Table 5.20: **Results of the LMM on: Step latencies regressed against session (levels: s1-2, s3-4), BP and age group. Baseline group (against which comparisons are presented): ADU.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	-0.46	-0.56	-0.36	0.05	4278	-8.66	<0.001	<0.001	***
s1-2or3-4	-0.07	-0.13	-0.00	0.03	4278	-2.01	0.045	0.077	.
BP	0.40	0.34	0.46	0.03	4278	13.09	<0.001	<0.001	***
ageGp9-12	0.54	0.37	0.71	0.08	81	6.47	<0.001	<0.001	***
ageGp7-8	1.18	1.00	1.36	0.09	81	12.85	<0.001	<0.001	***
s1-2or3-4:ageGp9-12	-0.04	-0.15	0.06	0.05	4278	-0.78	0.438	0.526	
s1-2or3-4:ageGp7-8	-0.36	-0.47	-0.24	0.06	4278	-5.88	<0.001	<0.001	***
BP:ageGp9-12	0.03	-0.07	0.13	0.05	4278	0.54	0.587	0.587	
BP:ageGp7-8	-0.10	-0.21	0.02	0.06	4278	-1.68	0.093	0.140	
s1-2or3-4:BP	-0.06	-0.15	0.03	0.05	4278	-1.30	0.195	0.260	
s1-2or3-4:BP:ageGp9-12	0.06	-0.09	0.20	0.07	4278	0.76	0.448	0.489	
s1-2or3-4:BP:ageGp7-8	0.32	0.15	0.48	0.08	4278	3.74	<0.001	<0.001	***

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:
.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Post-hoc tests revealed all pairwise comparisons were significant (as shown in tables 5.23, 5.22 and 5.21). However, when looking at the contrasts for the two-way interaction (session x 7-8 group), the regression estimates suggest a larger decrease of latencies (from session 1&2 to 3&4) for 7-8 year-olds ($\beta = 0.292$) than for 9-12 year-olds ($\beta = 0.108$) or adults ($\beta = 0.095$). When looking at the contrasts for the three-way interaction, the larger decrease of latency with session number for 7-8 year-olds is remarkably greater at nBP ($\beta = 0.421$ for 7-8 year-olds compared to $\beta = 0.107$ and 0.065 for 9-12 year-olds and adults) than it is at BP ($\beta = 0.1636$ for 7-8 year-olds compared to $\beta = 0.125$ and 0.110 for 9-12 year-olds and adults).

Table 5.21: **Contrasts following the 3-way interaction in Table 5.20's LMM: effect of sessions (1-2 vs 3-4) on Step latencies.**

contrast	estimate	SE	df	t-ratio	p	p (FDR)	sig
ageGp = ADU, BP = nBP							
s1-2 vs 3-4	0.0654	0.0326	4278	2.007	0.0448	0.045	*
ageGp = 9-12, BP = nBP							
s1-2 vs 3-4	0.1069	0.0423	4278	2.524	0.0117	0.014	*
ageGp = 7-8, BP = nBP							
s1-2 vs 3-4	0.4209	0.0510	4278	8.261	<.0001	<0.001	***
ageGp = ADU, BP = BP							
s1-2 vs 3-4	0.1248	0.0325	4278	3.834	0.0001	<0.001	***
ageGp = 9-12, BP = BP							
s1-2 vs 3-4	0.1096	0.0421	4278	2.605	0.0092	0.014	*
ageGp = 7-8, BP = BP							
s1-2 vs 3-4	0.1636	0.0507	4278	3.228	0.0013	0.003	**

Degrees-of-freedom method: containment

Table 5.22: **Contrasts following the 2-way interaction (sessions:ageGp) in Table 5.20's LMM: effect of age group on Step latencies.**

contrast	estimate	SE	DF	t-ratio	p	p (FDR)	sig
sessions 1-2							
ADU - 9-12	-0.5538	0.0792	81	-6.995	<.0001	<0.001	***
ADU - 7-8	-1.1297	0.0870	81	-12.991	<.0001	<0.001	***
9-12 - 7-8	-0.5759	0.0930	81	-6.191	<.0001	<0.001	***
sessions 3-4							
ADU - 9-12	-0.5407	0.0795	81	-6.803	<.0001	<0.001	***
ADU - 7-8	-0.9325	0.0877	81	-10.634	<.0001	<0.001	***
9-12 - 7-8	-0.3919	0.0933	81	-4.199	0.0002	<0.001	***

Results are averaged over the levels of: BP
Degrees-of-freedom method: containment
P value adjustment: tukey method for comparing a family of 3 estimates

Table 5.23: **Contrasts following the 2-way interaction (sessions:ageGp) in Table 5.20's LMM: effect of sessions (s1-2 vs s3-4) on Step latencies.**

contrast	estimate	SE	DF	t-ratio	p	p (FDR)	sig
ageGp = ADU							
s1-2 vs 3-4	0.0951	0.0232	4278	4.103	<.0001	<0.001	***
ageGp = 9-12							
s1-2 vs 3-4	0.1082	0.0302	4278	3.583	0.0003	<0.001	***
ageGp = 7-8							
s1-2 vs 3-4	0.2923	0.0362	4278	8.074	<.0001	<0.001	***

Results are averaged over the levels of: BP
Degrees-of-freedom method: containment

5.3.4.4 Branch points and non-branch points in the routinised phase

Sessions 3&4 - Latencies

The latencies at BP and nBP as a function of ST and age groups in sessions 3&4 are shown in Figure 5.22 (and their estimated marginal means in Figure A.6 in the appendix).

The LMM indicated that BP had significantly longer latencies than nBP (t(DF= 3851)=-8.29, $p < 0.001$) and that there was no interaction. Full results are presented in table 5.24. The effect of BP was also present in each group and condition with the non-parametric tests (see table 5.25).

The LMM indicated that ST_{+} had significantly longer latencies than ST_{-} (t(DF= 3851)=-6.27, $p < 0.001$), without an interaction between ST and age. With the non-parametric tests, the effect of ST was significant in adults and 9-12 year-olds (table 5.26).

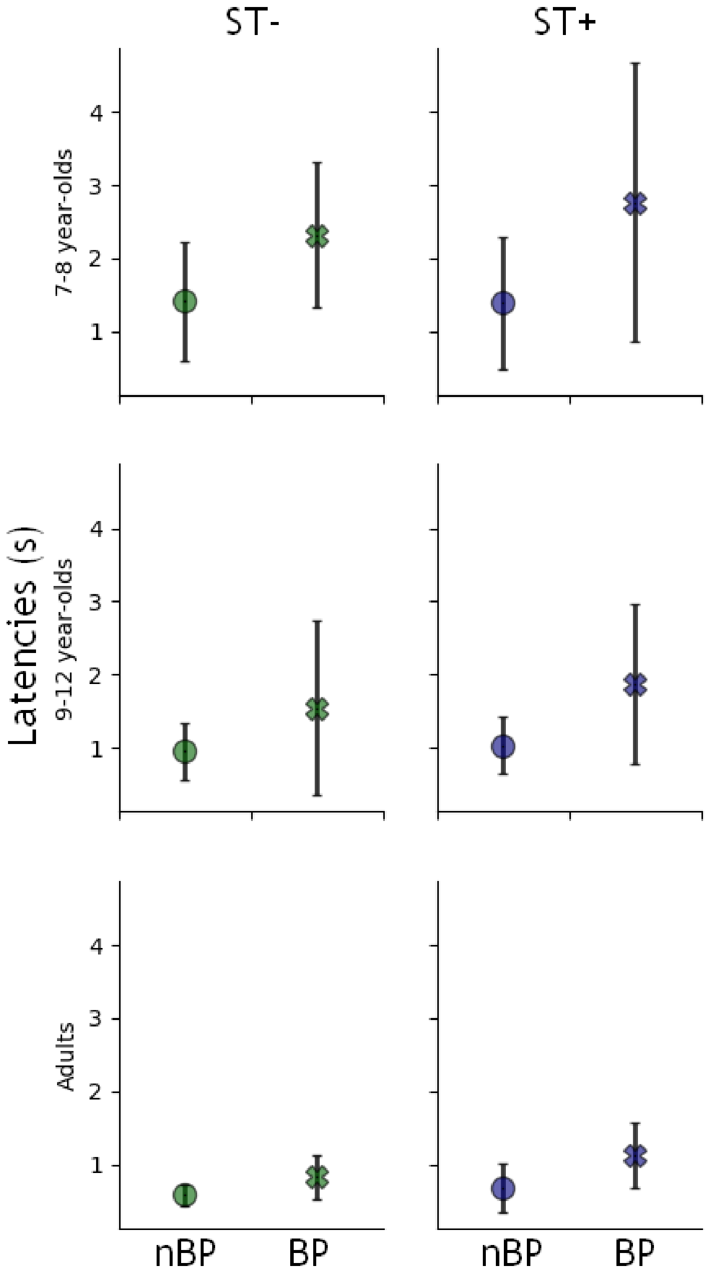


Fig. 5.22: **Step latencies at BP (branch points) and nBP (non branch points) with secondary task (ST_+) or without (ST_-): median values of participants' trials' median. Error bars represent the IQR (Inter-Quartile Range).**

Table 5.24: **Results of the Linear Mixed Model regression (LMM): Step selection latency regressed against BP, ST and age groups. Baseline group (against which comparisons are presented): ADU.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	0.05	-0.07	0.17	0.06	3851	0.87	0.386	0.515	
ST	-0.27	-0.36	-0.19	0.04	3851	-6.27	<0.001	<0.001	***
BP	-0.39	-0.48	-0.29	0.05	3851	-8.29	<0.001	<0.001	***
ageGp9-12	0.40	0.22	0.59	0.09	81	4.34	<0.001	<0.001	***
ageGp7-8	0.73	0.52	0.94	0.10	81	7.00	<0.001	<0.001	***
ST:9-12	0.12	-0.01	0.25	0.07	3851	1.75	0.081	0.194	
ST:7-8	0.13	-0.03	0.28	0.08	3851	1.62	0.105	0.180	
BP:9-12	-0.06	-0.20	0.08	0.07	3851	-0.80	0.421	0.505	
BP:7-8	-0.14	-0.31	0.02	0.08	3851	-1.70	0.088	0.176	
ST:BP	0.06	-0.06	0.18	0.06	3851	0.95	0.341	0.512	
ST:BP:9-12	-0.03	-0.22	0.15	0.09	3851	-0.35	0.724	0.790	
ST:BP:7-8	-0.01	-0.22	0.21	0.11	3851	-0.06	0.948	0.948	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p*: *(FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Table 5.25: **Wilcoxon signed rank tests comparing Branch points (BP) to non branch points (nBP) latencies, with (ST_+) and without (ST_-) secondary task. The figures in the table are: statistic(N pairs with non-zero difference), p-value.**

	ST_+	ST_-
7-8	W(20)=0, p<.0001	W(20)=19, p=0.0006
9-12	W(27)=3, p<.0001	W(27)=11, p<.0001
ADU	W(37)=29, p<.0001	W(37)=40, p<.0001

Table 5.26: **Wilcoxon signed rank tests comparing looking at Secondary task effects on latencies, within nBP and within BP. The figures in the table are: statistic(N pairs with non-zero difference), p-value).**

	nBP	BP
7-8	W(20)=99, p=0.8408	W(20)=71, p=0.2162
9-12	W(27)=99, p=0.0306	W(27)=96, p=0.0255
ADU	W(37)=131, p=0.0009	W(37)=120, p=0.0005

Consistent with the non-parametric test results on overall trial duration, the latencies were overall longer in 7-8 year-olds than adults ($t(\text{DF}= 81) = 7, p < 0.001$), and in 9-12 year-olds than adults ($t(\text{DF}= 81) = 4.34, p < 0.001$). When comparing the two children's groups together (table 5.27, the results (main effect of BP and ST) matched the LMM's results (BP: $t(\text{DF}= 2074) = -7.55, p < 0.001$, ST: $t(\text{DF}= 2074) = -2.78, p < 0.013$); and similarly to the trial-level durations, 7-8 year-olds were significantly slower than 9-12 year-olds ($t(\text{DF}= 45) = 2.58, p = 0.026$).

***Table 5.27:* Results of the LMM in the children's groups only: Step selection latency regressed against BP, ST and age continuous. Baseline group: 9-12 years old.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	0.46	0.30	0.62	0.08	2074	5.63	<0.001	<0.001	***
ST	-0.15	-0.26	-0.05	0.06	2074	-2.78	0.005	0.013	*
BP	-0.44	-0.56	-0.33	0.06	2074	-7.55	<0.001	<0.001	***
ageGp7-8	0.33	0.07	0.58	0.13	45	2.58	0.013	0.026	*
ST:7-8	0.01	-0.17	0.18	0.09	2074	0.09	0.925	0.925	
BP:7-8	-0.08	-0.27	0.10	0.10	2074	-0.89	0.376	0.602	
ST:BP	0.02	-0.13	0.18	0.08	2074	0.31	0.753	1.004	
ST:BP:7-8	0.03	-0.22	0.27	0.13	2074	0.21	0.834	0.953	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p: (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Sessions 3&4 - Latencies $Delta(ST_+ - ST_-)$

The $Delta\ Latencies\ (ST_+ - ST_-)$ as a function of BP and age groups in sessions 3&4 are shown in Figure 5.23 (and their estimated marginal means in Figure A.8 in the appendix).

The LMM indicated a 'non significant but notable' effect of BP on $Delta\ Latencies\ (ST_+ - ST_-)$: $t(\text{DF}= 249) = -2.24, p = 0.078$, with larger $Delta$ at BP than at nBP (see table 5.28). In the children's group alone the effect of BP $Delta\ Latencies\ (ST_+ - ST_-)$ was also significant: $t(\text{DF}= 139) = -2.77, p = 0.012$ (see table 5.29). In other words, for children, the effect of secondary task was significantly greater at branch points than non-branch points. For adults, the effect was notable, but not significant. Hypothesis 3 is therefore supported, at least for children.

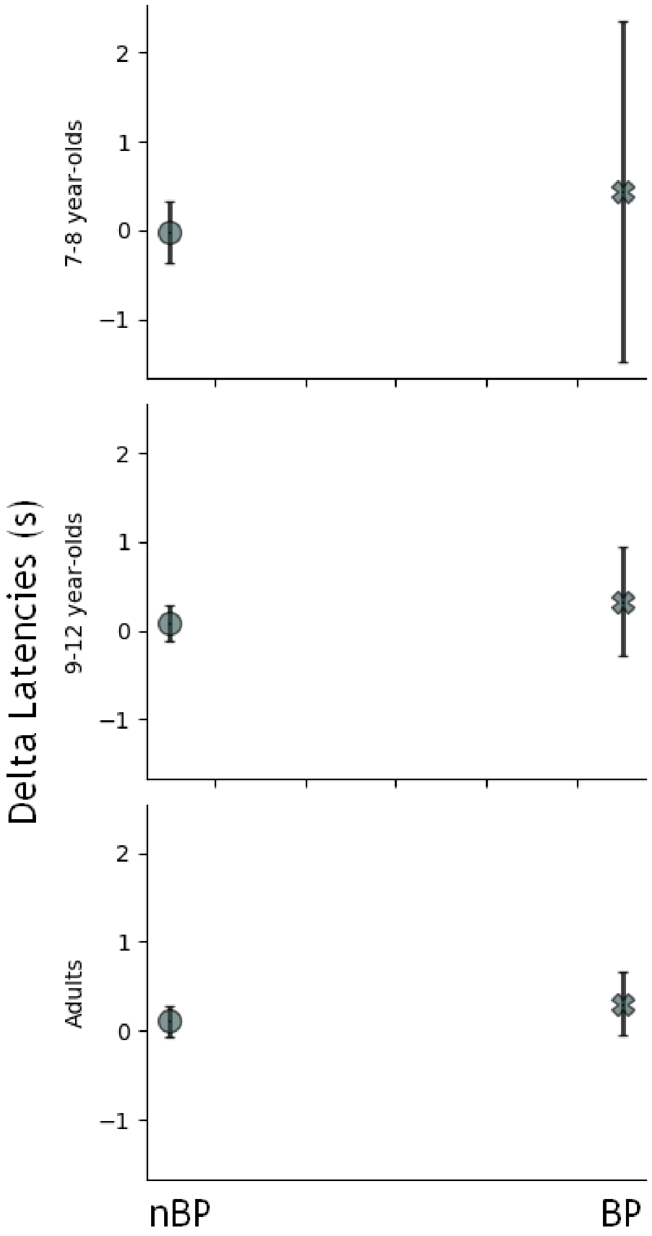


Fig. 5.23: **Differences in latencies with and without secondary task** $DeltaLatencies\ (ST_{+} - ST_{-})$**: median values of participants' trials' median. Error bars represent the IQR.**

***Table 5.28:* Results of the LMM on: *Delta Latencies* $(ST_+ - ST_-)$ regressed against BP and age group. Baseline group: ADU.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	0.25	0.16	0.35	0.05	249	5.20	<0.001	<0.001	***
BP	-0.09	-0.16	-0.01	0.04	249	-2.24	0.026	0.078	.
ageGp9-12	-0.05	-0.20	0.10	0.07	81	-0.68	0.499	0.748	
ageGp7-8	-0.12	-0.28	0.04	0.08	81	-1.46	0.148	0.296	
BP:9-12	-0.03	-0.15	0.09	0.06	249	-0.53	0.596	0.596	
BP:7-8	-0.04	-0.16	0.09	0.07	249	-0.55	0.582	0.698	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

***Table 5.29:* Results of the LMM in the children's groups only: *Delta Latencies* $(ST_+ - ST_-)$ regressed against BP and age group. Baseline group: 9-12 years old.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	0.20	0.10	0.31	0.05	139	3.78	<0.001	<0.001	***
BP	-0.12	-0.20	-0.03	0.04	139	-2.77	0.006	0.012	*
ageGp7-8	-0.07	-0.23	0.10	0.08	45	-0.84	0.403	0.537	
BP:7-8	-0.00	-0.13	0.12	0.07	139	-0.07	0.947	0.947	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

The non-parametric results (table 5.30) indicate a 'non significant but notable' effect of BP on *Delta Latencies* $(ST_+ - ST_-)$ for 9-12 year-olds ($W(N = 27) = 112, p = 0.064$) and adults ($W(N = 37) = 241, p = 0.096$). Taken together, it seems there could be an effect of BP on the amplitude of the ST effect, but the data is too noisy to draw a definitive conclusion..

***Table 5.30:* Wilcoxon signed rank tests comparing $Delta(ST_+ - ST_-)$ at BP to $Delta(ST_+ - ST_-)$ at nBP. N is the number of pairs with non-zero differences.**

	7-8	9-12	adults
W-value	67	112	241
N	20	27	37
p-value	0.1650	0.0643	0.0955

An interesting effect of BP on *Delta Latencies* $(ST_+ - ST_-)$ is that latencies are much more variable (as indicated by the larger IQR error bar) at BP than nBP, and this difference looks increasingly pronounced with decreasing age.

There was no significant effect of age group on *Delta Latencies* $(ST_+ - ST_-)$ shown by both the LMM (table 5.28) and the non-parametric tests (table 5.31).

***Table 5.31:* Mann-Whitney U tests comparing $Delta\ Latencies(ST_{+} - ST_{-})$ between groups. N_1: sample size of the group on the left-hand side, N_2 for the right-hand side of the comparison.**

	nBP	BP
7-8 vs 9-12	U=246, N_1=20, N_2=27, p=0.3066	U=252, N_1=20, N_2=27, p=0.3533
7-8 vs ADU	U=333, N_1=20, N_2=37, p=0.2708	U=335, N_1=20, N_2=37, p=0.2820
9-12 vs ADU	U=471, N_1=27, N_2=37, p=0.3517	U=496, N_1=27, N_2=37, p=0.4837

Sessions 3&4 - Latencies $Delta\ (BP - nBP)$

The $Delta\ Latencies\ (BP - nBP)$, or BP cost, as a function of ST and age groups in sessions 3&4 are shown in Figure 5.24 (and their estimated marginal means in Figure A.7 in the appendix).

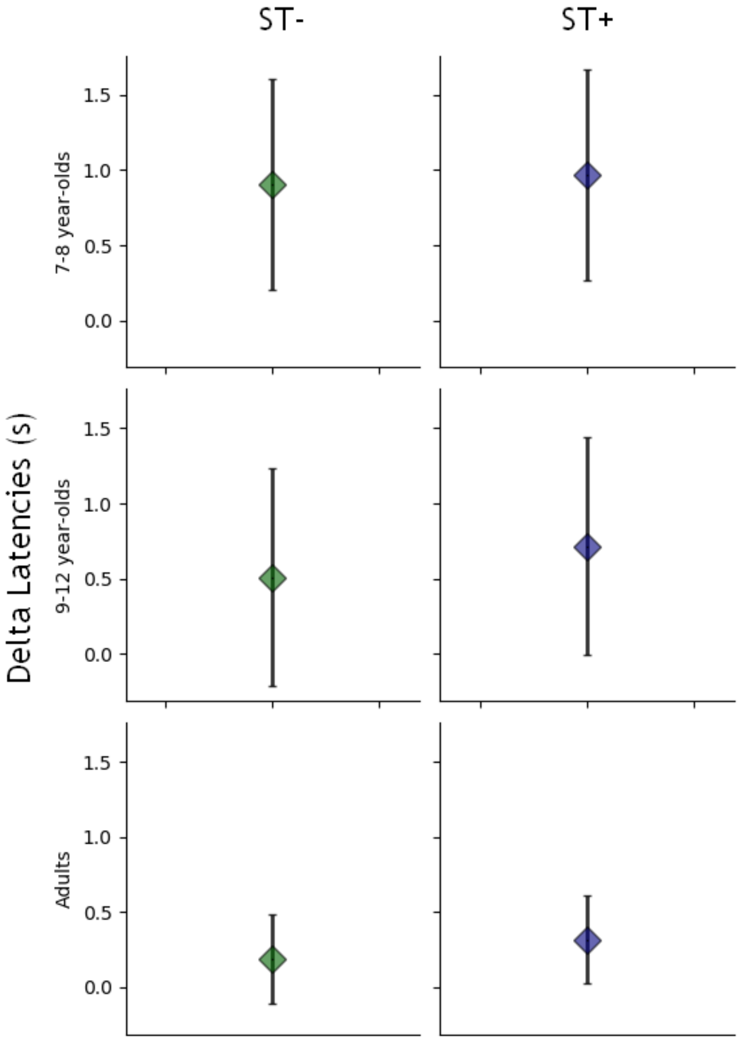


***Fig. 5.24:* Differences in latencies at BP (branch points) and nBP (non branch points) $Delta\ Latencies\ (BP - nBP)$: median values of participants' trials' median. Error bars represent the IQR.**

In the LMM (tables 5.32 and 5.33), the was a significant effect of age group: $Delta\ Latencies\ (BP - nBP)$ was significantly larger for 7-8 year-olds than adults ($t(DF = 81) = 2.38, p < 0.001$), and for 9-12 year-olds than adults ($t(DF = 81) = 4.76, p = 0.040$). $Delta\ Latencies\ (BP - nBP)$ was 'non significant but notably' larger for 7-8 than 9-12 year-olds ($t(DF = 45) = 2.02, p = 0.098$). This was almost entirely corroborated by the non-parametric tests (table 5.34): the $Delta\ Latencies\ (BP - nBP)$ differed amongst all age groups in all conditions, except for the 7-8 against the 9-12 year olds in the ST_+ condition (U(N_1=20, N_2=27)=166, p=0.0130 in ST_-; 7-8 vs adults: U(N_1=20, N_2=37)=96, p=<.0001 in ST_+ and U=96, N_1=20, N_2=37, p=<.0001 in ST_-; 9-12 vs adults: U(N_1=27, N_2=37)=289, p=0.0022 in ST_+ and U(N_1=27, N_2=37)=369, p=0.0386 in ST_-).

***Table 5.32:* Results of the LMM on: $Delta\ Latencies\ (BP - nBP)$ regressed against ST and age group. Baseline group: ADU.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	0.44	0.23	0.64	0.11	211	4.16	<0.001	<0.001	***
ST	-0.21	-0.44	0.03	0.12	211	-1.73	0.084	0.126	
ageGp9-12	0.38	0.06	0.70	0.16	81	2.38	0.02	0.040	*
ageGp7-8	0.84	0.49	1.19	0.18	81	4.76	<0.001	<0.001	***
ST:9-12	-0.09	-0.45	0.27	0.18	211	-0.49	0.624	0.624	
ST:7-8	-0.16	-0.56	0.23	0.20	211	-0.81	0.419	0.503	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u):* Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

***Table 5.33:* Results of the LMM in the children's groups only: $Delta\ Latencies\ (BP - nBP)$ regressed against ST and age group. Baseline group: 9-12 years old.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	0.82	0.53	1.11	0.15	121	5.59	<0.001	<0.001	***
ST	-0.30	-0.63	0.03	0.17	121	-1.79	0.076	0.101	
ageGp7-8	0.46	0.00	0.91	0.22	45	2.02	0.049	0.098	.
ST:7-8	-0.07	-0.58	0.43	0.26	121	-0.29	0.774	0.774	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u):* Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

***Table 5.34:* Mann-Whitney U tests comparing $Delta\ Latencies\ (BP - nBP)$ between groups. N_1: sample size of the group on the left-hand side, N_2 for the right-hand side of the comparison.**

	ST_+	ST_-
7-8 vs 9-12	U=254, N_1=20, N_2=27, p=0.3694	U=166, N_1=20, N_2=27, p=0.0130
7-8 vs ADU	U=227, N_1=20, N_2=37, p=0.0086	U=96, N_1=20, N_2=37, p=<.0001
9-12 vs ADU	U=289, N_1=27, N_2=37, p=0.0022	U=369, N_1=27, N_2=37, p=0.0386

The LMM also found no significant effect of ST (tables 5.32 and 5.33) on $Delta\ Latencies\ (BP - nBP)$. This was also the case for the non-parametric tests (table 5.35).

Table 5.35: **Wilcoxon signed rank tests comparing** $Delta\ Latencies\ (BP - nBP)$ **at** ST_{+} **to** $Delta\ Latencies\ (BP - nBP)$ **at** ST_{-}**.**

	7-8	9-12	adults
W-value	104	123	293
N	20	27	37
p-value	0.9854	0.1128	0.3775

Sessions 3&4 - Trajectory deviation

Although not analysed in-depth due to the high correlation of the deviation variables with latencies (cf. Methods' Section), we present in the appendix the median value of: the Absolute average distance of each trajectory point to the final point, i.e. the target icon (cf. appendix's Figure A.4) and the Log maximal distance relative to the final point (cf. appendix's Figure A.5). Unsurprisingly given their correlations with latencies, the two variables' patterns seem to mirror the latencies' pattern: the figure shows larger deviation at BP than nBP, and a slightly larger deviation with than without the secondary task.

Sessions 3&4 - Coefficient of Variation (CoV)

The latencies at BP and nBP as a function of ST and age groups in sessions 3&4 are shown in Figure 5.25.

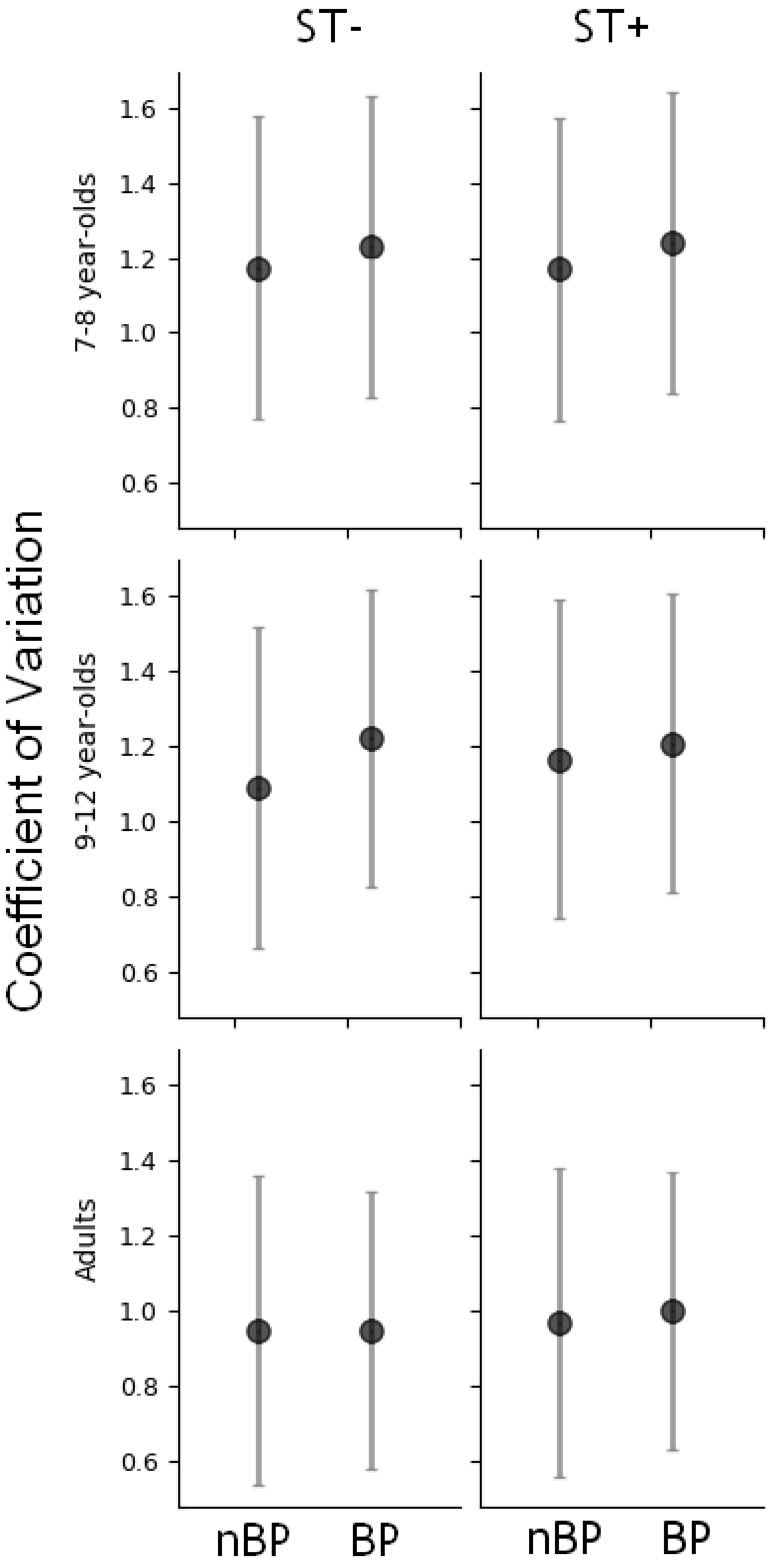


Fig. 5.25: **Speed coefficient of variation ($CoV = std_{speed}/average_{speed}$: median values of participants' trials' median. The error bars represent the IQR.**

Despite the sizable error bars, the LMM (table 5.36) revealed a significant three-way (BP x ST x 9-12 group) interaction ($t(DF = 2626) = 2.25, p = 0.056$). Post-hoc comparisons (tables 5.38 and 5.37) indicated the following. In 9-12 year-olds, CoV was significantly larger at BP than at nBP ($t(DF = 2626) = 4.204, p < 0.001$) in ST_{-}, and 'non significantly but notably' larger at ST_{+} than at ST_{-} ($t(DF = 2626) = 2.36, p = 0.055$) at nBP. Larger BP than nBP in 9-12 year-olds in ST_{-} was also found with a non-parametric test ($W(N = 27) = 55, p = 0.0013$, cf. table 5.39).

Table 5.36: **Results of the LMM on: Speed coefficient of variation (CoV) regressed against BP, ST and age group. Baseline group: ADU.**

Predictor	β	CI (l)	CI (u)	SE	DF	t	p	p (FDR)	sig
Intercept	-0.01	-0.07	0.04	0.03	2626	-0.49	0.625	0.833	
BP	-0.04	-0.08	-0.01	0.02	2626	-2.38	0.017	0.051	.
ST	-0.07	-0.10	-0.03	0.02	2626	-3.56	<0.001	<0.001	***
ageGp9-12	0.19	0.10	0.28	0.05	81	4.14	<0.001	<0.001	***
ageGp7-8	0.19	0.08	0.29	0.05	81	3.54	0.001	0.004	**
ST:9-12	0.07	0.01	0.13	0.03	2626	2.25	0.025	0.060	.
ST:7-8	0.06	-0.02	0.14	0.04	2626	1.48	0.139	0.238	
BP:9-12	0.00	-0.06	0.06	0.03	2626	0.09	0.929	1.013	
BP:7-8	0.01	-0.07	0.09	0.04	2626	0.28	0.783	0.940	
BP:ST	0.03	-0.02	0.09	0.03	2626	1.27	0.204	0.306	
BP:ST:ageGp9-12	-0.10	-0.18	-0.01	0.04	2626	-2.20	0.028	0.056	.
BP:ST:ageGp7-8	-0.00	-0.11	0.11	0.06	2626	-0.04	0.969	0.969	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u)*: Confidence Interval (upper bound), *SE*: standard error, p: unadjusted p-value, p *(FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values: .: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Table 5.37: **Contrasts following the 3-way interaction in Table 5.36's LMM: effect of ST on Speed coefficient of variation ($CoV = std_{speed}/average_{speed}$).**

contrast	estimate	SE	DF	t-ratio	p	p (FDR)	sig
ageGp = ADU, BP = BP							
ST+ - ST-	0.0664	0.0187	2626	3.557	0.0004	0.002	**
ageGp = 9-12, BP = BP							
ST+ - ST-	-0.0040	0.0251	2626	-0.159	0.8736	0.874	
ageGp = 7-8, BP = BP							
ST+ - ST-	0.0083	0.0345	2626	0.241	0.8098	0.972	
ageGp = ADU, BP = nBP							
ST+ - ST-	0.0330	0.0187	2626	1.762	0.0783	0.157	
ageGp = 9-12, BP = nBP							
ST+ - ST-	0.0598	0.0253	2626	2.363	0.0182	0.055	.
ageGp = 7-8, BP = nBP							
ST+ - ST-	-0.0230	0.0347	2626	-0.663	0.5074	0.761	

Degrees-of-freedom method: containment

Table 5.38: Contrasts following the 3-way interaction in Table 5.36's LMM: effect of BP on Speed coefficient of variation ($CoV = std_{speed}/average_{speed}$).

contrast	estimate	SE	DF	t-ratio	p	p (FDR)	sig
ST_+, ADU							
BP - nBP	0.0448	0.0188	2626	2.384	0.0172	0.052	.
ST_-, ADU							
BP - nBP	0.0113	0.0185	2626	0.612	0.5403	0.648	
ST_+, 9-12							
BP - nBP	0.0420	0.0250	2626	1.678	0.0935	0.187	
ST_-, 9-12							
BP - nBP	0.1058	0.0252	2626	4.204	<.0001	<0.001	***
ST_+, 7-8							
BP - nBP	0.0337	0.0353	2626	0.954	0.3401	0.510	
ST_-, 7-8							
BP - nBP	0.0024	0.0329	2626	0.073	0.9415	0.942	

Degrees-of-freedom method: containment

Table 5.39: Wilcoxon signed rank tests for the effect of BP on CoV. The figures in the table are: statistic(N pairs with non-zero difference), p-value.

	ST_+	ST_-
7-8	W(19)=68, p=0.2935	W(19)=81, p=0.5949
9-12	W(27)=144, p=0.2796	W(27)=55, p=0.0013
ADU	W(36)=285, p=0.4508	W(36)=317, p=0.8015

In adults, the LMM indicated that CoV was 'non significantly but notably' larger at BP than at nBP ($t(DF = 2626) = 3.55, p = 0.02$) in ST_+, and significantly larger at ST_+ than at ST_- ($t(DF = 2626) = 2.384, p = 0.052$) at BP. Larger CoV at ST_+ than at ST_- in adults at BP was also found with a non-parametric test ($W(N = 26) = 205, p = 0.044$, cf. table 5.40).

Table 5.40: Wilcoxon signed rank tests for the effect of Secondary task on CoV. The figures in the table are: statistic(N pairs with non-zero difference), p-value.

	nBP	BP
7-8	W(19)=88, p=0.7983	W(19)=95, p=1.0000
9-12	W(27)=122, p=0.1075	W(27)=161, p=0.5011
ADU	W(26)=287, p=0.4699	W(26)=205, p=0.0443

To guide the interpretations of the CoV measures, we illustrated the profile of speed, that is the variation of speed across time-normalised sample number (Figure 5.26), with a few trials as examples, separated into 'low CoV' (when below the threshold of 1) and 'high CoV' (when above the threshold of 1). We can visualise there how the low CoV speed profiles tend to be smoother than the high CoV ones. Additionally, we show the trajectories on screen that correspond to high and low CoV trials (Figure 5.27). Although the CoV is not a direct measure of the smoothness of the trajectories themselves, there seems to be an overall trend for smoother speed profiles to have smoother trajectories, as can be seen from the higher number of sharp turns or swirls on the top than on the bottom panel.

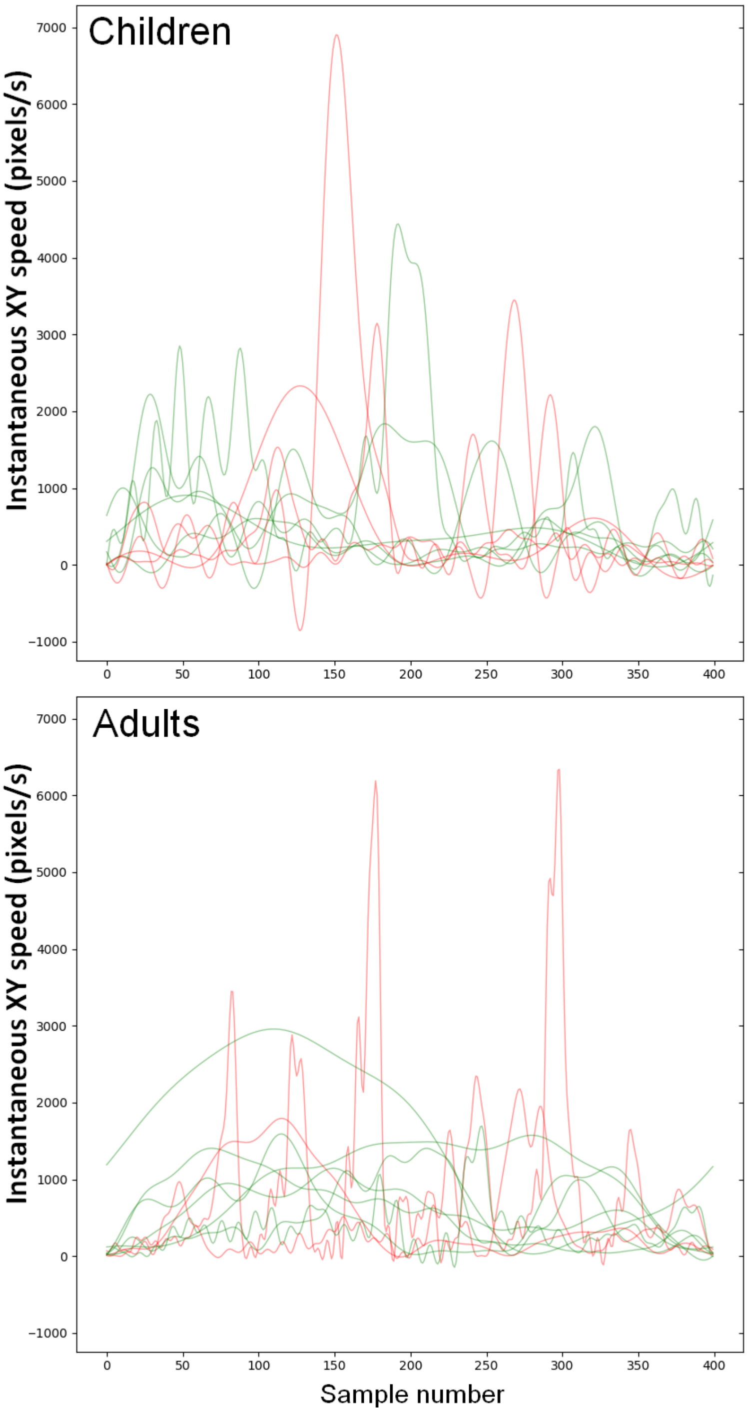


Fig. 5.26: **Instantaneous speed profiles to illustrate low and high CoV trials. Each curve represents a different trial (in children's and adults' data). The green curves have a CoV below 1, the red curves above 1. The time in abscissa is normalised in order to overlay the curves.**

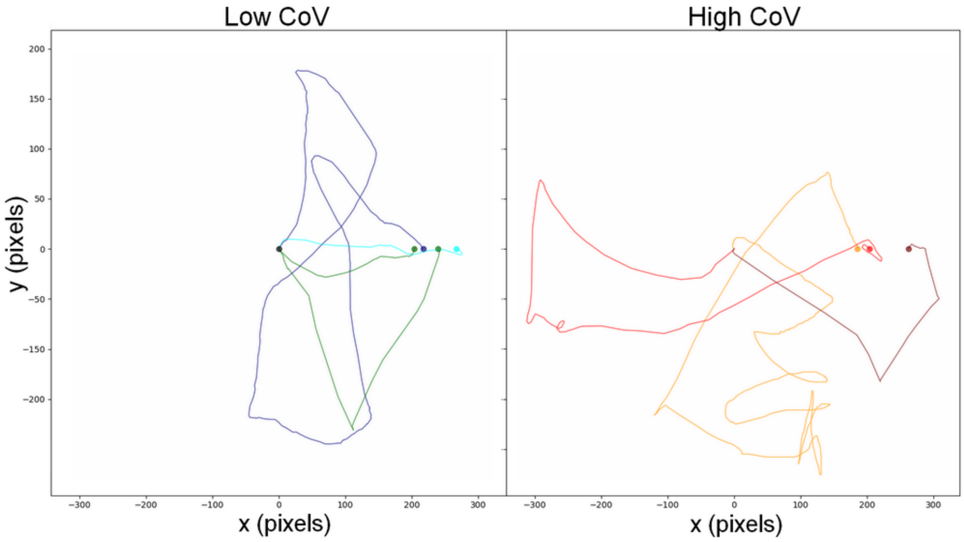


Fig. 5.27: **Mouse trajectories indirectly illustrating what low and high low and high CoV trials look like in trajectories. Each curve represents a different trial (adults' data only). The 'low CoV' panel shows trajectories with a CoV below 1; the 'high CoV' panel shows trajectories with a CoV above 1.**

5.3.4.5 Stop-signal task

The adult participants' Stop Signal Response Times (SSRT) ranged between 144 and 305 ms. The SSRT distribution is shown in Figure 5.28.

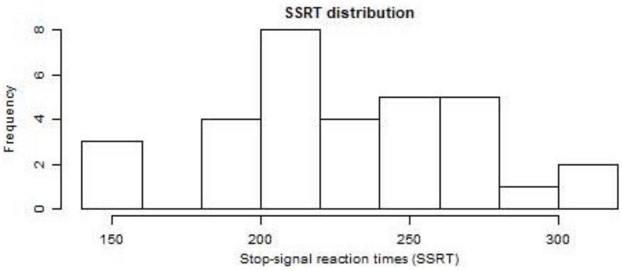


Fig. 5.28: **Distribution of the adults' SSRT in the Stop-signal task.**

We looked at the correlation between the SSRT and the sequence-making variables that could be expected to relate to inter-individual differences in supervisory process involvement in the task. Figure 5.29 represents the SSRT plotted against the proportion of Incorrect goals executed (top left panel), the latencies at BP (top right panel), the difference between latencies at BP and nBP $Delta\ Latencies\ (BP - nBP)$ (bottom left panel), and the coefficient of variation in speed CoV (bottom right panel).

As mentioned in introduction to experiment 4, we expected a positive relationship between SSRT and the four presented variables. Spearman correlation were computed, as shown in table 5.41. In contrast to this expectation, none of the correlations were significant, nor even notable.

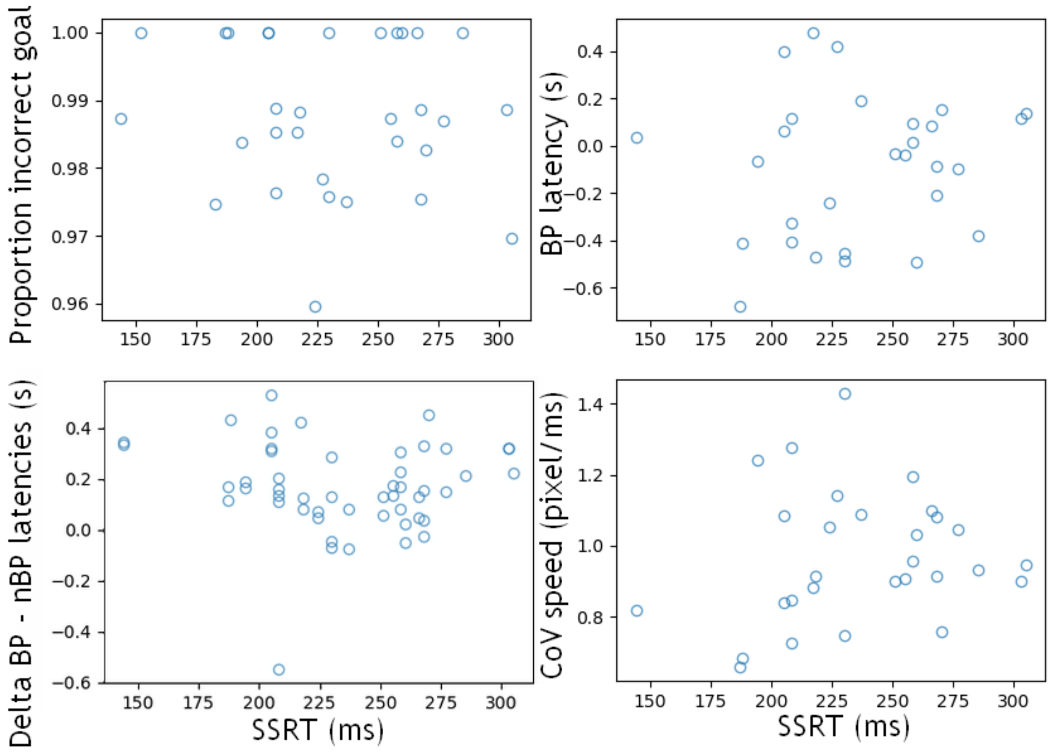


Fig. 5.29: Correlations between the SSRT and the Proportion incorrect goals, BP latency, $Delta\ Latencies\ (BP - nBP)$ and Coefficient of Variation (CoV).

Table 5.41: Spearman correlations between the Stop Signal Response Time SSRT (Stop-signal task) and the indicated variables of the main sequence-making task (correlation coefficient R, p-values p).

Variable correlated to SSRT	R	p
Proportion incorrect goals	-0.138	0.475
BP latency	0.147	0.446
DELTA latencies (BP - nBP)	-0.217	0.259
Coefficient of Variation (CoV)	0.2216	0.247

5.3.5 Discussion

Overall, the study was entertaining enough to successfully collect a large number of trials per participant over up to four sessions, allowing us to investigate routinised action control and its process of routinisation. To our knowledge, this is the first study of this kind to look at long action sequences in children with such extensive practice. We discuss the acquisition or routinisation process by looking at all four sessions, before focusing on the last two sessions and in particular on BPs and nBPs.

5.3.5.1 Task acquisition

The findings at trial-level across sessions replicate the findings of experiment 3. First, there is a drop in trial duration from the first two sessions to the last two. This strengthens the point that the last two sessions are the best suited to study (partially) routinised performance. Additionally, we replicated the finding that only younger children show a decrease of incorrect goal execution across sessions. This is most likely because the other age group were already performing close to floor level early in training. Unsurprisingly, we also replicate the finding that the younger the age group, the longer the trials.

When looking at finer changes, the trial-level duration dropped for all age groups between sessions one and two. The trial-level duration indicated that in adults, session 3 and 4 significantly differs. In other words, the gains in speed have not plateaued yet by the end of the third session, so the routinisation is probably not complete yet. Albeit not significant in adults, the drop in duration from sessions 3 to 4 seemed smaller than the drop from sessions 1 to 2, which suggested it was still sensible to group sessions 3 and 4 together and consider them as a more advanced routinisation phase. The fact that the difference in drop from session 3 to 4 is smaller than from sessions 1 to 2 in children but not adults must not be ignored. Indeed, the routinisation level may vary across age group, which should be kept in mind when doing age group comparisons.

Importantly, the fact that the trial duration further decreases between session 3 and 4 for some participants, suggests that the findings will inform us on *partially routinised* action sequence performance rather than routine action performance per se. In routine, or *fully routinised* performance, the supervisory system is held to intervene at BPs but not at nBPs (Norman, 1981). However, in *partially routinised* action sequence performance, the supervisory system is likely to intervene at all points in the sequence, yet still at a stronger degree at BPs than at nBPs.

When looking at the task acquisition at step-level , 7- to 8-year-olds had a larger decrease in latencies across sessions than 9- to 12-year-olds and adults, and even more so at nBPs than at BPs. The data suggest that younger children start at significantly poorer baseline performance (in the early learning phase) and thus have more room for improvement (speed-up) throughout the task.

An increase in trial duration with the secondary task is clear only in adults. However, the average error in this secondary task strongly suggests that all participants are attempting to do the secondary task in parallel to the sequence-making. The fact that the secondary task significantly slowed down adults and not children may be due to the increased difficulty of the secondary task in adults (6 to 11 sounds to count in the task for adults, against 3 to 7 in the task for children). The increase in difficulty may have been too large compared to the ease with which dual-tasking can be expected to increase with age. Alternatively, the variability in the children's data may have masked an increase in trial duration with the secondary task.

5.3.5.2 Characteristics of routinised action control

H1: Action selection at BPs is slower than at nBPs

A central hypothesis we aimed to build upon is the fact that action selection takes longer at BPs than nBPs: we aimed to replicate the finding in adults, and demonstrate it in children. The BP and nBP selected for analyses differed only in their subsequence position (within- or between- subsequence) and in the transition probability to the next item (and not in the distance to follow, nor their particular location in the sequence such as the start or the end). They also did not differ much semantically and perceptually as they were all implements. The finding that action selection at BPs took longer than at nBPs is clearly replicated, regardless of the age group. To our knowledge, this is the first study that demonstrated this BP effect of latency in children. This reinforces the idea that children control action in a qualitatively similar fashion to adults.

Further evidence for the qualitative similarity of routine action in children and adults comes from Yanaoka and Saito (2019). The latter found that 5-year-olds presented similar patterns to adults in Botvinick and Plaut (2004), where adults performed more errors when interrupted mid-subsequence than when interrupted at the end of a subsequence. Collectively with experiment 3, this confirms that our framework is suited to study children's routine action control, and that it is reasonable to perform cautious quantitative comparison (e.g., amplitude of BP cost shown in latencies) between adults and children.

H2: The secondary task affects latencies

The secondary task significantly slowed task performance of specific steps but did not significantly slowed the performance of the entire task in children. It could be that when looking at an entire trial, the data gets too noisy to detect an effect. It is also possible that other steps compensate the BP's and nBP's slowing down with a speeding up, for example at the end of a trial, when one realises they have been slowed down by the secondary task. This would explain that the effect of secondary task is not apparent at the level of a trial while it is apparent at the level of BP and nBP.

A slow down not only at BPs but also at nBPs is not surprising given that the task is not at its highest state of possible routinisation. Still, the extent of the slow down was expected to vary according to the BP or nBP nature of the step, which we will examine next.

H3: The secondary task affects BP latencies more than nBP latencies

We hypothesised that the supervisory control needed to pay attention to and accurately perform the secondary task (ST) would greatly interfere with the supervisory control needed at BPs, resulting in significantly larger secondary task delay at BPs than at nBPs.

The data is consistent with this view: for all but the youngest age group, there was a trend for the secondary task to affect BP more strongly than nBP across groups, but the statistics were unclear about the presence of this effect (the LMM indicated that such an effect was 'non significant but notable', while the non-parametric test indicated that such an effect was 'non significant but notable' in older children and adults but not in younger children, probably due to the large variability found in younger children).

The fact that the difference between the ST effect on BPs and the ST effect on nBPs remains small or unclear may be due to the degree of routinisation not being sufficiently advanced, and that supervisory control might still be required (although to a small extent) at nBPs. In Ruh et al. (2010), there was a clear-cut interaction with latencies at BP being prolonged by the secondary task, but not latencies at nBPs. This may be explained by the fact that the number of trials was far greater in Ruh et al. (2010) than our study (200 trials spread over 3 sessions in Ruh et al. (2010) and 64 trials spread over 4 sessions in the present study).

5.3.5.3 Development of routinised action control

H4: The BP cost changes with age

Children were overall slower than adults both at BPs and nBPs, and the younger the slower, which is unsurprising. More interestingly, if we consider that subtracting latencies at nBPs from latencies at BPs performs a normalisation for age groups' baseline speed, we can view this variable as a *BP cost*, that is, how much more difficult are the BPs (compared to the nBPs) for a given age group. This BP cost was significantly higher for both children's groups than for the adult group, and to some extent, higher for younger children than older children (bearing in mind that for the latter the non-matching statistical results do not allow to draw a clear conclusion). This is in line with our hypothesis which states that children, having less good executive function than adults, will struggle more in selecting actions at branch points, because this part of the sequence requires additional supervisory control.

There is a small discrepancy between the results of the LMM and the nonparametric test. While the former suggested that $Delta\ Latencies\ (BP - nBP)$ was significantly larger in 7- to 8- than 9- to 12-year-olds, the latter suggested that this was true only when not performing the secondary task. If this interaction was confirmed, it would imply that 9- to 12-year-olds have a smaller BP cost, hence better routinisation than 7- to 8-year-olds under normal conditions. However, their BP cost would become as high as the 7- to 8-year-olds when under cognitive load.

Thus, the higher BP cost for children than adults, and the trend of a higher BP cost for younger than older children in the present age range, suggest that the ability to recruit supervisory control at branch points, or the effectiveness of supervisory control at branch points, improves with age.

Exploratory Hypothesis: The CoV reveals developmental changes in supervisory capacity

We explored the kinematics features of the participants' mouse movements on screen. Following the preceding chapter's 'Variability of speed' feature, and following the high correlation between trajectory features and the response time in this study, we concentrated our analysis on the CoV (Coefficient of variability of speed, which is a normalised version of the 'Variability of speed' feature used in the previous chapter).

With practice, and consequently across development, actions are performed with less and less variability. The variability can be considered across repetitions of action over large developmental scales (Bruner, 1973), across trials of a task (Chu et al., 2016), or as we approach it here within-trials (Benedetti et al., 2020). Bruner (1973) points out that actions are performed with less and less variability across development. The variability within trial can also be seen as the degree of smoothness (with high variability indicating low smoothness). Benedetti et al. (2020) analysed the smoothness of velocity profiles by classifying them into 'multi-peaks' velocity profiles and 'one-shot' (steep slope without speak) velocity profiles. Less smooth movements were linked to proactive control. Although the task and context differed from ours, it goes together with the idea that the smoothness of speed profiles is informative of the type of control engaged in a task. Dotan et al. (2018) also found within-trial changes in speed informative, with momentary speed indexing the participant's degree of confidence during a touchscreen task.

The CoV as a function of BP and ST highlighted two particular points. First, in 9-12 year-olds, the nBP without secondary task had significantly lower CoV than the BP without secondary task, or than the nBP with the secondary task. This could be explained by the nBP needing minimal supervisory control at nBP without secondary task. Second, in adults, the BP with the secondary task had significantly higher CoV than the nBP with the secondary task or than the BP without secondary task. It is harder to draw conclusions about younger children since the absence of effects may be due to their noisier data.

Here, we will distinguish *partially routinised action sequence performance* (at the level of an entire action sequence) from the control of a single action autonomously by the routine system. *Partially routinised action sequence performance* involves the efficient use of both systems throughout a sequence. As a reminder, the comparison of latencies at BP compared to nBP in children and adults suggested that action selection at nBP could be carried out mainly by the routine system, albeit perhaps with some remaining supervisory influence, while selection at BP benefited from considerably more biasing of the supervisory system relative to nBP. We further propose that the CoV reflects when the resources of the supervisory system are saturated.

Under no cognitive load, it seems that both children and adults have good *partially routinised action sequence performance*, but the CoV patterns may reveal differences in how close to saturation their systems are. As mentioned earlier, unlike in fully routinised sequential performance, the non branch points in this study still require a small degree of supervisory control. The CoV patterns suggest that children under no cognitive load are able to recruit their supervisory system at the points that are less routinised within the context of partially routinised performance, but do so by using the maximal capacity of the supervisory system. Therefore, the entire sequence (both at BP and at nBP) is susceptible to interference from a secondary task.

For adults, additional cognitive load conflicts only with the recruitment of supervisory control at BP, where larger supervisory resources are needed; thus, the supervisory resources are only saturated both at BP and under cognitive load.

Thus, it seems that children, like adults, are able to recruit both routine and supervisory systems when needed in the task. Furthermore, it seems that the supervisory system capacity in children is sufficient for good routine performance under normal conditions, but saturates quickly under heightened cognitive load. This is consistent with the view that the supervisory system has an extended period of development and is not yet fully developed at the end of school age (Casey et al., 2005a, Lenroot and Giedd, 2006, Mehnert, Akhrif, Telkemeyer, Rossi, Schmitz, et al., 2013).

Role of inhibition: results on the Stop-signal task are inconclusive

We expected that some executive function scores measured by a standard task (the stop-signal task) would correlate with measures of supervisory control involvement in sequence execution, such as the BP cost or the CoV. We found no such correlations. A number of reasons could explain this absence of correlation. First, it could be that response inhibition may not be relevant to sequential task performance. It could be that components of executive functions other than inhibitory control are important. Although all executive functions components measures obtained by standard task on their own (as opposed to a large number of tasks such as Miyake et al., 2000b) correlate together, the specific inhibition component may not be central to this task. Alternatively, it could be that participants showed different levels of engagement with the sequence-making task and the stop-signal task: the sequence-making task was gamified and thus perhaps more engaging than the stop-signal task. This is all the more plausible that the game was significantly shorter for a participant who tried to execute it more rapidly, thus presumably paid more attention, than for a participant who did not. On the contrary, the duration of stop-signal task could not be significantly shortened by fast responses. Third, it could be that the online version of the stop-signal task does not compare well with inhibition measured in the lab.

Other studies have linked executive function and action sequence execution or planning. Perone et al. (2020) found that the scores in executive functions as measured by the Minnesota Executive Function Scale predicted 6- to 7- and 8- to 10-year-olds' performance in carrying out a series of different naturalistic daily-life tasks.

The task did not attempt to separate executive functions into distinct components. Yanaoka and Saito (2020) found that shifting abilities (as assessed by the Advanced Dimensional Card Change Sort task of Chevalier and Blaye, 2009) were related to how children's sequence control is affected by mid- or end-sequence interruptions. Schröer et al. (2021) found that in 3- to 5- years olds, better inhibition skill made sequence execution less vulnerable to distractors, while better updating skill was related to better following the subgoal order of a sequence. Apart from the age range of participants in this study compared to ours, a notable difference is that Schröer et al.'s (2021)'s experiment involved several distractors, which may increase the need for inhibitory control as compared to our experiment.

Consequently, although the evidence in our study does not rule out that inhibitory control plays an important role in routine sequence control across development, other studies highlight that shifting/updating abilities may play a more important role, at least when distractor objects are absent.

5.4 General Discussion

5.4.1 Task acquisition and routinisation

In both experiments, it appeared that participants learned the task without difficulty, even the youngest. Thus, errors do not arise from placing too much demand on memory. It seemed that participants reached a certain degree of routinisation, as evidence for example by a bigger drop in duration between the first two sessions compared to between the last two sessions for children. However, evidence also suggests that routinisation may not be fully achieved. Therefore, the view that we propose differs slightly from Norman and Shallice's (1986) account. In the latter account, the routine system carries out sequential action performance autonomously, and is biased by the non-routine system at branch points only. However, the data from experiments 3 and 4 suggests that participants' performance is only *partially* routinised. Thus, the supervisory system is most likely involved at all points in the sequence, but the routine system plays the dominant role at non branch points while the supervisory system plays the dominant role at branch points.

Furthermore, it is possible that different age groups have different learning rates and thus reach different degrees of routinisation. Indeed, learning rates in a variety of tasks have been found to differ between adults and children (e.g., Master, Eckstein, Gotlieb, Dahl, Wilbrecht, Anne, et al., 2020). Thus, one could argue that the differences between age groups that we may find at sessions 3&4 may not be related to different maturation levels of their routine and/or supervisory system but only to the different degree of routinisation they reach. However, we propose that the developmental differences in the supervisory system (beyond learning rate) is the more plausible explanation for the effects found at branch points and under cognitive load, given the large number of trials involved and given the theoretical knowledge of the involvement of supervisory functions at branch points in adults, and of the role of changing executive functions in children's performance on many tasks (discussed in introduction). Regarding the routine system, it is difficult to separate the efficiency of routinisation from the amount of practice needed to get to an advanced level of routinisation. Thus, we will consider them as one phenomenon (efficiency of the routine system).

In the present type of task, there is a trade-off between setting a number of trials that is high enough to reach some degree of routinisation, and ensuring that children do not get bored by the game. This trade-off might be overcome by using data from commercial video games (which are, undeniably, engaging enough to many players), if such data was available and usable for the purpose of understanding the process of routinisation. Another possibility, that we expand on later, is to study routines at a larger time scale and/or in real-life.

5.4.2 Higher-level goal execution

In both experiments 3 and 4, incorrect goal execution was rare for all age groups. The rarity of such errors indicates that participants did not fail to understand the rules and were trying to follow them. The errors resemble the occasional attention or goal monitoring failure that occurs in real-life actions, where two similar sequences used in resembling contexts can be swapped (see, for example, Norman, 1981, Reason, 1979).

While incorrect goal execution was rare in both experiments reported in this chapter, the younger the participants, the more frequent were these errors. This suggests that children ability to keep track of a goal cue (the trial's background colour), to retrieve the mapping between the goal's cue and the appropriate sequence and/or to maintain the goal in mind throughout the sequence is still developing throughout school years.

5.4.3 Hierarchical control of action

In experiment 3, the errors were distributed with a certain concentration near what can be conceived as subtask boundaries (near actions that visibly achieve a subogal), both in children and adults. This suggests an underlying hierarchical representation used already early in school years.

The difference in response time between branch points (BP) versus control non branch points (nBP), in all age groups, was consistent with the formation of hierarchically-structured representations. As a reminder, the branch points were characterised by 50% transition probability between the preceding action and the branch point action (while non branch points had 100% transition probability), and could also be seen as a boundary between two different subgoals (*adding the Nutella* and *adding the topping*).

5.4.4 Cognitive load and the role of supervisory functions

In both experiments 3 and 4, higher error rates or longer response time highlighted points in the sequence that were more challenging to participants. The fact that the secondary tasks tended to affect the same points more substantially than other points, often boundaries between subgoals, together with the results of previous findings such as in Ruh et al. (2010), further suggests that these more challenging points require higher influence of the supervisory system if they are to be performed without error.

5.4.5 Developmental changes within the dual-systems framework

In Experiment 3, age mainly seemed to increase the error rate particularly at the points where the baseline error rate was already high. Such points are assumed to recruit higher supervisory control.

In Experiment 4, inter-action latencies at branch points were longer than at non branch points and this cost was bigger in children than adults (and perhaps but not clearly evidenced in the data, bigger for 7- to 8-year-olds than 9- to 12-year-olds), suggesting a qualitatively similar functioning of partially routinised action sequence performance across developmental stages, with a larger intervention of the supervisory system at branch points, and whose efficiency improves with age.

The results confirmed that the dual-systems framework can explain children's routine sequential action control, and that it is sensible to perform comparisons between adults and children under the same framework.

A kinematic marker, Coefficient of variation (CoV) revealed further developmental differences in the availability of supervisory resources. In Experiment 4, a higher CoV was found at branch points under cognitive load in adults, and at all points except for the non branch points under no cognitive load in 9- to 12-year-olds, and we proposed the CoV to index how close to saturation the supervisory resources are being used.

The response time together with the CoV data suggest that school-age children already have an harmonious interaction of two systems to support *partially routinised action sequence performance*, where both systems intervene to various extents according to the point in the sequence. But children's supervisory system has less capacity than adults, which is reflected in the CoV measure. The CoV measure captures more subtle changes than RT because the CoV captures increased saturation even at non branch points under cognitive load.

Thus, even though school-age children can execute routine sequential actions well under normal conditions, their supervisory system is not yet at its fully matured capacity. This is consistent with the view that general executive function, as measured by standard tasks in the developmental literature and well-known to improve throughout childhood including school years, play a key part in action sequence control.

5.4.6 Limitations and methodological considerations

The GoGelato task aimed to strike a balance between encompassing key features of real-life action sequences and being gamified, easy to learn and engaging enough for children, while allowing us to measure errors, selection latencies and kinematics.

A fair criticism is that we did not counterbalance the order of nBPs and BPs. If anything, the fact that nBP appeared earlier in the sequence might have made the nBP slower and could have masked the difference in latency between BP and nBP. However, we still found much larger latencies at BP than nBP.

Another concern may be that children at both ends of the age range have different amounts of experience with using a computer mouse. From informal exchange with participants, this sounds likely. However, our key analyses either compared raw measures within participants, or compared composite scores (e.g., difference with and without secondary task) across age groups. This limits any possible interference of individual participants' experience with computer and computer mouse with our findings. The same is true for the experience or knowledge of participants with more or less similar real-life action sequences. On top of that, our game involved a large number of trials, and the analyses were focused on the second half of the trials; if any interference with personal experience or knowledge existed, it would have likely been diminished by such extensive practice by the second half of sessions. Nonetheless, the fact that some of the youngest children in our sample had very limited experience using the computer mouse before taking part in the study (even if they had the opportunity to progress throughout the four sessions) very probably added noise to the data, which limits the interpretations within this age range (for example for CoV measure). Studies with naturalistic action sequences are needed to circumvent this problem. On a related note, we have differences in variability (for error, latency and kinematic variables) across age groups, which is a common drawback of developmental data. The task also has artificial aspects and limitations that are inherent to any computerised experiment; or to some lab-based experiments, such as repeating many trials in a row, and thus does not reproduce acquisition of routines in real-life which would be more likely be repeated only a few times per day but over many months and years.

One may also be worried about data quality arising from the online nature of the task. Concerns over the quality and reliability of data collected online are discussed by Anwyl-Irvine et al. (2020). One source of concern may be that participants behaviour may differ online compared to a lab experiment (e.g., participants not concentrating on the task or multi-tasking, participants lying about who they say they are, etc.). Following recommendations on online testing with children and adults from Kochari (2019), we reduced this concern by the design of our game which did not allow cheating. Indeed, the game would move on only if the sequence were properly carried out, the only possible error being not changing the recipe according to the goal cue; but we found all participants took the goal cue into account. Furthermore, adults were selected among reliable Prolific users, that is, whose previous participation in other experiments had not been excluded by other researchers. For children, families were met via video call on the first session and were in touch via email throughout the experiment, which reinforces the confidence in the same child having taken part throughout all four sessions. Additionally, during the initial data quality check we scanned for any suspicious data pattern (e.g., surprising response time variations across trials of all sessions).

Relatedly, one might worry about whether sampling of online participants may not be generalisable. Compared to the accessibility of research laboratories, if anything the study being online might have made it more accessible (e.g. participants who live far away from a research lab and would not have travelled for in-person study were able to take part) and potentially have increased the variety of our sample (we cannot verify this because we did not collect personal information beyond age and gender for data protection matters). Another source of concern may be the variability in hardware and software across participants. We ensured that the order of magnitude of the expected response times were large enough to not be affected by online measurement imprecision (following Bridges et al. (2020) who found 10 ms as the worst precision achieved in the most commonly used software combinations). In practice, many researchers have found online studies to replicate findings from lab-based experiments (e.g., Tsay et al., 2021).

Nonetheless, a remaining limitation is that the task was a mouse computer game, which is a step further away from naturalistic action control (compared to touch-screen use in the previous Chapter's study). Motor control for mouse usage is undeniably different from real-life 3D actions. However, we believe that the task did recruit the same key processes of interest as real-life routine action control. Indeed, findings from the computerised task of Ruh et al. (2010) supports the dual-systems view, which itself is vastly supported by findings from real-life error studies in neurological patients (reviewed by Shallice, 2006) and the action slips and lapses of neurologically healthy participants (Norman, 1981, Reason, 1979). This indicates agreement between computer-based tasks and real-life behaviour.

The convergence of findings from computer-based tasks and real-life behaviour exists in adults, but it would be interesting to find evidence of such an agreement in children. To our knowledge, there has been no past study that looked at children performing long action sequences with repeated practice in a naturalistic environment. We can cite related work which, although involving shorter sequences or fewer trials, are indicative that our results may extend to real-life settings. Yanaoka and Saito (2019) found evidence for a similar use of task representation in adults and 5-6 year old children. Additionally, Yanaoka and Saito (2020) used knowledge on the central role of branch points, by inserting reminders prior to branch points while 3-6 year olds were learning routines. Compared to when inserting no reminders, this promoted routine acquisition. This emphasises the importance of goal maintenance in real-life action sequences, especially before a context-dependent action, and links to the particular difficulty at branch points in our computerised action sequences.

5.5 Conclusions

The newly-developed GoGelato task replicated key findings on sequential action control in adults consistent with the dual-systems framework (Norman and Shallice, 1986, Cooper et al., 2014), and importantly was suitable to study routine action control in children.

The patterns of errors in experiment 3 suggested that sequences of action are controlled using underlying hierarchically-structured representations both in children and adults. Children's control seemed particularly weaker at certain boundaries between subtasks. In experiment 4, the response time differences at branch points compared to non branch points, or BP cost, replicated the findings of Ruh et al. (2010). We furthermore found some evidence that this BP cost decreases throughout the school-age years.

In both experiments, the points that were more susceptible to errors or that took longer in the absence of cognitive load also tended to be the points more affected by cognitive load, further confirming that supervisory control is more strongly recruited at those particular points, and giving insights into the combined use of routine and non-routine systems to control an action sequence. Furthermore, a kinematic feature (coefficient of variation of speed) was proposed to reflect supervisory resource use, and suggested lower supervisory capacity in children compared to adults.

Overall, changes in both routine and supervisory functions may drive improvements in children's sequence execution accuracy and efficiency. Nonetheless, changes in supervisory functions are likely to play a greater role, and the lesser maturation of such supervisory function impact the sequential action performance in early school-age years.

Chapter 6

General Discussion

Abstract

This chapter reviews the key findings of the thesis in connection with each other and their context. Firstly, it summarises the key findings in each chapter. Secondly, it describes the contribution on the methodological side, that includes a new approach to kinematic analyses and the development of a child-friendly paradigm to study the development of action control. Thirdly, it develops the thesis' contributions to the field the theoretical side, highlighting the role of executive function development in driving changes in action control throughout the school years. It also discusses the newly developed computational model together with current theories of the development of action control, to propose a framework for understanding how two systems work together to control action across development. Finally, it lays out future research directions.

6.1 Overview of empirical and computational findings

6.1.1 Action-based inhibitory selection mechanisms improve throughout childhood

In Chapter 2, we investigated the interference from distractors when reaching for a target, by re-analysing duration and kinematics data. The findings helped us understand the resolution of competition from potentially conflicting action representations, in children and adults. It was previously established that the environment can have a direct influence on action (e.g., affordances of Gibson, 1977), and that top-down processes related to the intended goals can mediate action selection. Furthermore, it was known that, around preschool age, children's actions tended to follow perceptual inputs more than internal goals or strategy and that this trend decreased across development (e.g., Mitsopoulos et al., 2015). However, the changes in action selection processes across development were unclear. By analysing the precise course of reaching for a target among distractors in children and adults, Chapter 2 found that children's reaching to a target was slowed and suggestively deviated vertically when distractors were present, compared to when they were not, and this was true regardless of the semantic or perceptual compatibility between distractors and the target. This was the case despite the distractors being irrelevant throughout the task (the target remained the same), and this interference effect was not found in adults.

Given the hypothesised role of action-based inhibitory selection mechanisms (e.g., Tipper, 1985), and given that tasks different from our reaching task, using the same age range, have found that the development of inhibitory control helps conflict resolution (e.g., Flanker task in Erb et al., 2018), the novel findings strongly suggest that: (1) even irrelevant distractors pose a challenge in action selection early in childhood, most likely by the simultaneous activation of reaching actions to all the objects present in one's visual field, and (2) improvements in inhibitory control abilities throughout development drive the improvement in action selection when reaching for a pre-defined target among distractors.

6.1.2 The interaction between model-free and model-based types of control can be specified at the algorithmic level

In Chapter 3, we proposed a new computational model that explains data from a pre-existing task (the standard two-stage task, Daw et al., 2005) tapping into the balance of model-free and model-based systems. The systems are partly analogous to the routine and supervisory systems introduced by Norman (1981). First, we prove that it is possible, with an interactive activation network, to model the standard two-stage task at the algorithmic level of Marr (Marr, 1982). The task had previously been described at the computational level of Marr. The novel model provides a more detailed account of the processes of selection between actions and the integration of the two systems. Furthermore, this account has the potential to implement and test various mechanistic hypotheses. Additionally, it is able to make contact with response time data which the original model did not do.

Second, and thanks to the possibility of capitalising on the response time data, we implemented and tested four mechanistic hypotheses about the different performance of the developmental groups of Decker et al. (2016). A possibility is that the recruitment of the model-free and model-based systems is externally arbitrated and favours model-based more and more with age; the other three assume that the model-based system is recruited at the same level throughout development, but that another mechanism explains the apparent increase in the contribution of model-based with age. According to the second hypothesis, the computations get less noisy with age. According to the third, the model-based system becomes faster in performing computations. According to the last, the learning rate increases with age.

We were able to distinguish two hypotheses (the first and second hypotheses which were discussed in the previous paragraph) which were more compatible with the data than the other two. Crucially, we proved that this type of model is useful in distinguishing among competing hypotheses (even if, at this stage, it does not allow to draw conclusions about a single most plausible hypothesis). Thus, to explain the apparent increase in model-based contribution with development in this task, the most plausible hypotheses tested are that the model-free/model-based balance is externally arbitrated, and that the arbitration increasingly favours model-based with age, or that model-based computations become less subject to noise with age.

Focusing on a widely-used task has important advantages and provided some insights behind the changes in the model-free/model-based balance across development. Nonetheless, the task showed limitations in its suitability to investigate the changes within a child group, and furthermore may not be suited for use with children younger than 8 years old. It also appeared to not tap into the entirety of processes relevant for daily-life routine action control. Consequently, the following chapters consisted of experimental work with tasks that were specifically designed for studying children's routine action control.

6.1.3 Fine differences in the updating processes of an ongoing sequence occurs through development

Chapter 4 investigated the routine-like control by 5- to 9-year-olds of a well-learned short sequence, and the children's adaptability to mid-sequence unpredictable changes of the environment. It used a novel touchscreen task, recording the responses time and kinematics of icons being moved across the screen. Two variants of a sequence carried out two different goals. The transition between the first and second action had a low probability of changing the state of the environment in an unpredictable way. Such a change of state subsequently required updating what would have been the best action in the absence of the change. This manipulation allowed us to probe the recruitment of supervisory type of control during a sequence that may be controlled solely by routine type of control.

The youngest age group accurately performed a sequence adapted to the goal, revealing good routine control at this age. Furthermore, this younger age group was also able to monitor and update an ongoing sequence when facing unpredictable changes of state (where a change of the environment that may have been expected, given the action executed, led to an unexpected change of the environment). This indicated that the supervisory functions required to adapt an ongoing sequence were already in place and could be effectively used by participants from as young as the early school years. Nonetheless, this ability was limited in younger children when facing another type of environmental change, in which the goal changed unexpectedly. This result indicated that before 9 years of age, children's not fully-developed supervisory system does not allow an efficient monitoring and/or accurate action update within an ongoing sequence. The results further suggested that the failure to update an action might result from insufficient inhibitory control to interrupt a pre-planned action.

The findings in this chapter, that supervisory control becomes more prominent around school-years, are broadly consistent with the finding from Chapter 3. Chapter 4 further clarifies which components of supervisory control improves during this period: monitoring, updating functions, and possibly inhibition.

6.1.4 The hierarchical control of action sequences and involvement of two systems at various points vary with age

Chapter 5 probed the dual-modes of control during the execution, by 7- to 12-year-olds, of well-learned action sequences that are characteristic of real-life routines. The chapter included two studies, both consisting of an online mouse-tracking task. The main study had two variants of a 9-steps action sequence that featured overlapping actions between the sequences' variants, as well as diverging subsequences. The start of the diverging subsequences constituted a branch point. In addition, a secondary task was used in some trials in order to further understand the role of supervisory control.

The study replicated in adults and extended to children the key finding that adult action selection is slowed at branch points compared to non branch points (Ruh et al., 2010), and that branch points are particularly susceptible to secondary task interference. The findings support the idea that supervisory control is most needed at branch points when the selection between action is most ambiguous and needs to be connected explicitly to the context and goal representations. Other actions, however, may be carried out principally by routine control, which benefits from having learned regular transitions, at the parts of the sequence that overlaps between the two sequences' variants. Chapter 5 also found interesting developmental trends. Younger children appeared to have comparable ability to that of older children to control actions at non branch points, but younger participants' performance was less good than that of older participants at branch points or under cognitive load. This suggests that supervisory control is not yet fully developed at 7-9 years old and continues to improve later on, resulting in improvements in action control through development.

6.2 Methodological contributions

6.2.1 Novel kinematic analyses methods

Kinematics measurements are rich in information, and it has seen a recent increase in its use to probe the mind (see Freeman, 2018 for a review). Popular analytic approaches often face one of the two following limitations: being too little or too much task-specific. In the first case, analyses are tailored to a specific paradigm and research question, and thus have the advantage of capturing a largest amount of the variance, but do not allow to take advantage of being used across diverse studies. They consequently make cross-task comparisons difficult. In the second case, using pre-defined, literature-driven, kinematic features has the advantage of using converging knowledge from multiple studies around the construct affecting the kinematic features, but it presents the risk of missing other kinematics patterns that may be highly informative in the task at hand. We proposed and applied an alternative method. The proposed method defines a large set of possible features (not hypothesis-driven), and selects the most relevant features (that capture the most variance, and possibly additional constraints) extracted on a dataset independent but similar to the dataset of interest (e.g., pilot data). Our alternative method led to examining features that were not the most intuitive ones to look at (such as vertical deviation, or within-trial variability in speed) and provided rich insights on action selection among distractors or in routine sequence production.

6.2.2 Developing child-friendly hand-tracking games

The thesis presents newly-designed tasks, either inspired by a classical reinforcement learning task, or inspired from routine sequence production tasks in adults. The new tasks were gamified to be intuitive and engaging for children, and allowed us to measure a large number of trials of children executing various action sequences. Such tasks are crucial to study action control in children in a way that is as naturalistic as possible yet sufficiently controlled to directly extend results on comparable tasks in adults, verify the applicability of key principles of action control in adults to children and thus gain knowledge on the development of action control.

A methodological challenge when studying action control is that one must take into consideration multiple processes happening in parallel: visual, motor, attentional, memory, planning, etc. Each process studied separately does not necessarily reflect what happens when they are interacting together to control action. Thus, when designing a task to study action control, one must reproduce sufficiently naturalistic sequences that incorporates key characteristics of daily-life activities. We were specifically interested in *routine* action control, which means the task had to be repeated a sufficient number of times.

An additional challenge was designing tasks that children would understand and learn easily (within a single visit- on the lab or online), that children would engage with and have enough attention span to apply until the end of the study. Another crucial point was to ensure that any observed developmental change would not simply be attributable to differences in strategy or motivation with respect to the game, but instead be attributable to developmental changes in action control systems. Therefore, we controlled as much as possible for the valuation of game's rewards by different ages. For that purpose, we integrated entertaining cover stories and animations that would appeal to children across our age range, and promoted children's motivation to complete the task by providing encouraging signs of progression through the task.

The tasks were inspired or derived from existing tasks. They were substantially re-designed (in terms of context, nature of the sequence and of the actions, number of actions, rewards, etc.) and also programmed from scratch to fit this thesis' research purpose. One advantage that comes with the same researcher developing a new task and programming it, is that it ensures that every aspect of the game, and the way they may influence participants (visual salience of items, accessibility on screen, ease of selection, possible and impossible actions, etc.), are carefully thought through. Like computational modelling, programming the game forces one to theorise in detail all the processes that might be at play when participants do the task (not only the processes relevant to the research question, but also non-relevant processes that needs to be controlled for). Another advantage is that knowing how the data was generated and saved makes the analyses more reliable. The disadvantage of such computerised games is their non-ecological nature, despite computer games being relatively more naturalistic than ’keyboard-response’ tasks. Another disadvantage is that developing such a task requires programming skills and takes time.

The tasks were successful as evidenced by: (1) the large number of children completing the studies and the low drop-out rate, and (2) the theoretical insights gained. The insights include replicating findings on adult action control, thus reinforcing existing results and simultaneously bringing additional credibility to the tasks. They can and will also hopefully be tools for further studies, by adapting the games from this thesis, or drawing inspiration from them.

6.3 Theoretical contributions

6.3.1 Children must adapt to varying influences on action control

All four chapters highlight that action selection is subject to many influences, and that both the spatial and temporal context matters. Chapter 2 revealed the vulnerability of a reaching action to interference from surrounding objects, i.e., from the spatial context. Chapter 3, 4, and 5 studied additional dimensions integrated together, and brought the following evidence. Conflict in action selection can arise on the relative location of an action within a sequence, and the existence of alternative actions that may be valid in a similar context. Conflict can also arise from a pre-planned action that is made inappropriate by unexpected changes in the environment. The choice and speed of choice of an action is also influenced by one’s abilities to recruit supervisory control at crucial points of an action sequence, and secondarily of one’s propensity to be more explorative and exploitative.

In other words, competition for action selection can arise because of the physical presence of other objects (as evidence by Chapter 2), or from representation from other actions that may be compatible in a similar context but may achieve a different goal (as observed in Chapter 4 and 5). Difficulties in resolution of the competition processes can be apparent in errors. Alternatively, in the absence of error, they can be apparent in longer response times or in modified kinematic patterns.

6.3.2 The dual-systems is a suitable framework to account for the development of action control

Prior to this thesis, there was no direct evidence of the applicability of the dual-systems framework of Norman and Shallice (1986) to explain children’s action control. The dual-systems account is supported by empirical data in adults and by a computational model that simulates the key empirical phenomena characteristic of sequential action control (Cooper et al., 2014). The experimental paradigm of Ruh et al. (2010) was of particular interest. By designing a task comparable to the task of Ruh et al. (2010) and suitable for children, the thesis brought direct support for the existence of dual-systems type of control in school-aged children, and further indicated how the systems evolve across development.

Furthermore, all three Chapters 3, 4 and 5 looked at the combination of two modes of action control which, although defined slightly differently (model-free or model-based in the reinforcement learning terms of Daw et al. (2011), or routine/supervisory following Norman and Shallice, 1986), share key commonalities. There are more commonalities than differences, but the nuances will be discussed in Section 6.3.4. The findings from the three chapters converge towards the idea of a routine (model-free) mode of action selection already mature at the lower bound of the school-age range (albeit the findings of Chapter 5 do not rule out that the routine control may still slightly improve with age), and a supervisory (model-based) mode of action control significantly developing throughout the school years.

Importantly, the findings widen the possibility to further our understanding of children's action control by leveraging the dual-systems framework.

6.3.3 Executive functions changes drive the developmental changes in performance

The experimental evidence from this thesis and previous studies converge towards the idea that improvements in executive functions are a central driver for improvements of action control throughout childhood, including during school years.

The role of developing executive functions manifests itself at different levels in each chapter. Specifically, in the reaching for objects study (Chapter 2), the younger children (7 years old) were substantially slowed by the presence of distractor objects, and this interference decreased with age. This was argued to be attributable to changes in inhibitory control with age, which allow more efficient suppression of irrelevant action representations linked to the distractors.

The lab-touchscreen study (Chapter 4) and online-computer study (Chapter 5) had sequences differing in their length and challenged posed (Chapter 4: unexpected change within-sequence happening in some trials and not others; or Chapter 5: predictable sequences but with varying difficulty within the sequence). With these two different designs, we found a consistent pattern in which the younger ages in each study sample (5 to about 7 years old in Chapter 4, 7 to 8 years old in Chapter 5) had a similar time cost in action selection than older age (about 7 to 9 years old in Chapter 4, or 9 to 12 years old in Chapter 5) between a baseline condition and a slightly challenging one (common vs rare state transitions in Chapter 4, or non branch point vs branch points without secondary task in Chapter 5), but when in a more challenging one (rare goal transition Chapter 4, or under secondary task

Chapter 5), the cost was significantly higher in the younger compared to the older children.

This indicates that from early in school years, children have the building blocks to learn, execute and update action sequences appropriately, almost or just as well as children later in school ages, when the conditions are not excessively challenging. However, an increase in cognitive load or difficult points in the sequences reveals age differences. Specifically, younger children's performance is impaired while older children deal better with the challenges. Furthermore, when adults' performance was measured, adults showed further improvement compared to the older children's groups, suggesting that the development of executive functions necessary to improve performance in this task may continue beyond school-age years.

The multiple consequences of changes in executive functions on action control are summarised in Figure 6.1. Specifically, in Chapter 3, we found that the influence of the supervisory system overall increased from childhood to adulthood in repeated action selection. In Chapter 2, we found that gains in inhibitory control with age led to more efficient selection of actions in the face of distractors. In Chapter 4, the improvement with age in correcting an action sequence following an unpredictable within-sequence change was attributed to improvements in two functions: the ability to monitor or detect certain changes, and the ability to inhibit a presumably pre-planned action. Finally, in Chapter 5, we found that older children were more efficient than younger children in inhibiting a locally-competing action (that would be relevant at a same point in a sequence in a different context).

6.3.4 Broader developmental shift in the dual-systems control

As revealed by the computational work of Chapter 3, different mechanisms can lead to the same difference in the ratio of model-free/model-based types of control. Although the chapter's work was not able to select a single explicative hypothesis (furthermore, it may not have explored all plausible candidate hypotheses), an important point to keep in mind is that the same behaviour can be obtained by multiple underlying mechanisms.

If we were to use a reinforcement learning type of model on the studies in Chapter 4 and Chapter 5, and if we measured the 'balance model-free/model-based' for the sequence overall, we would probably find results similar to the Decker et al.'s (2016) study, where the balance goes towards more model-based and less model-free influence with age. However, this would be due to the increasing capacities of supervisory (model-based) control with age used to bias points such as branch points, whereas the influence of the supervisory system on the rest of the points would not change substantially with age.

As outlined in the Introduction Chapter (1), the views of Decker et al. (2016) and Janacsek et al. (2012) involve that a developmental shift occurs in late childhood/early adolescence where the 'default' involvement of a model-free strategy balances towards a model-based strategy, via both a maturation of the model-based

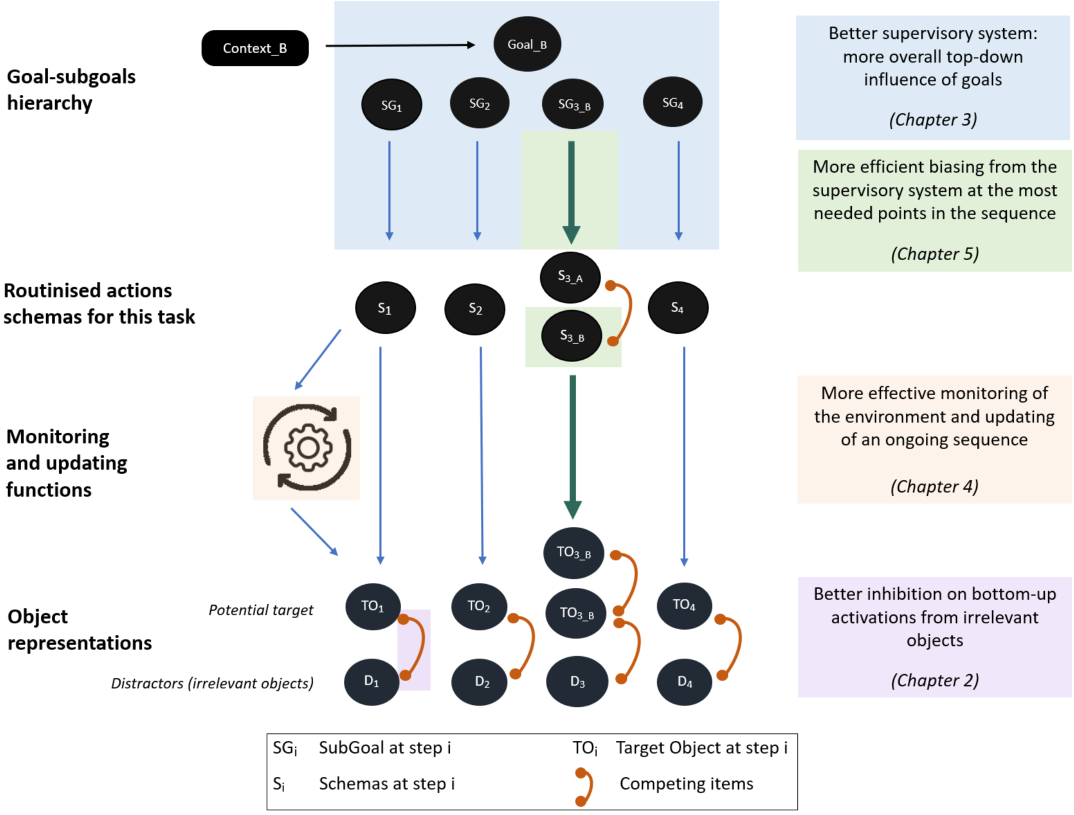


Fig. 6.1: **Summary of the key findings of the thesis: development of supervision functions related to the control of action. Arrows represent the influence of a system or of an item over another; the thickness of the arrows represent the relative degree of influence.**

system and a regression of the model-free system. Instead, we favour the view that, with development, the ability to control action sequences improves with the increasing capacities of the supervisory system, and that the stronger influence of the supervisory system will appear only when required, via a coordinated or conjoint control of routine and supervisory systems. This subtle nuance highlights the fact that many daily action sequences involve the coordinated recruitment of the two systems, each system having advantages at certain parts of the sequence. It is therefore interesting to understand the development of the joint actions of the systems, rather than the prominence of one or the other at different developmental stages.

From this stance, it appears that from early school age, and possibly earlier, children can efficiently use implicit-like or routine type of control when they have acquired experience with an action sequence that bears sufficient regularity. Throughout development, components of supervisory functions gradually improve. As a consequence, action sequences can be controlled more finely, and benefit from additional supervisory influence when it helps performance (e.g., at ambiguous points or in the face of unpredictable environmental changes). This view is summarised in Figure 6.2.

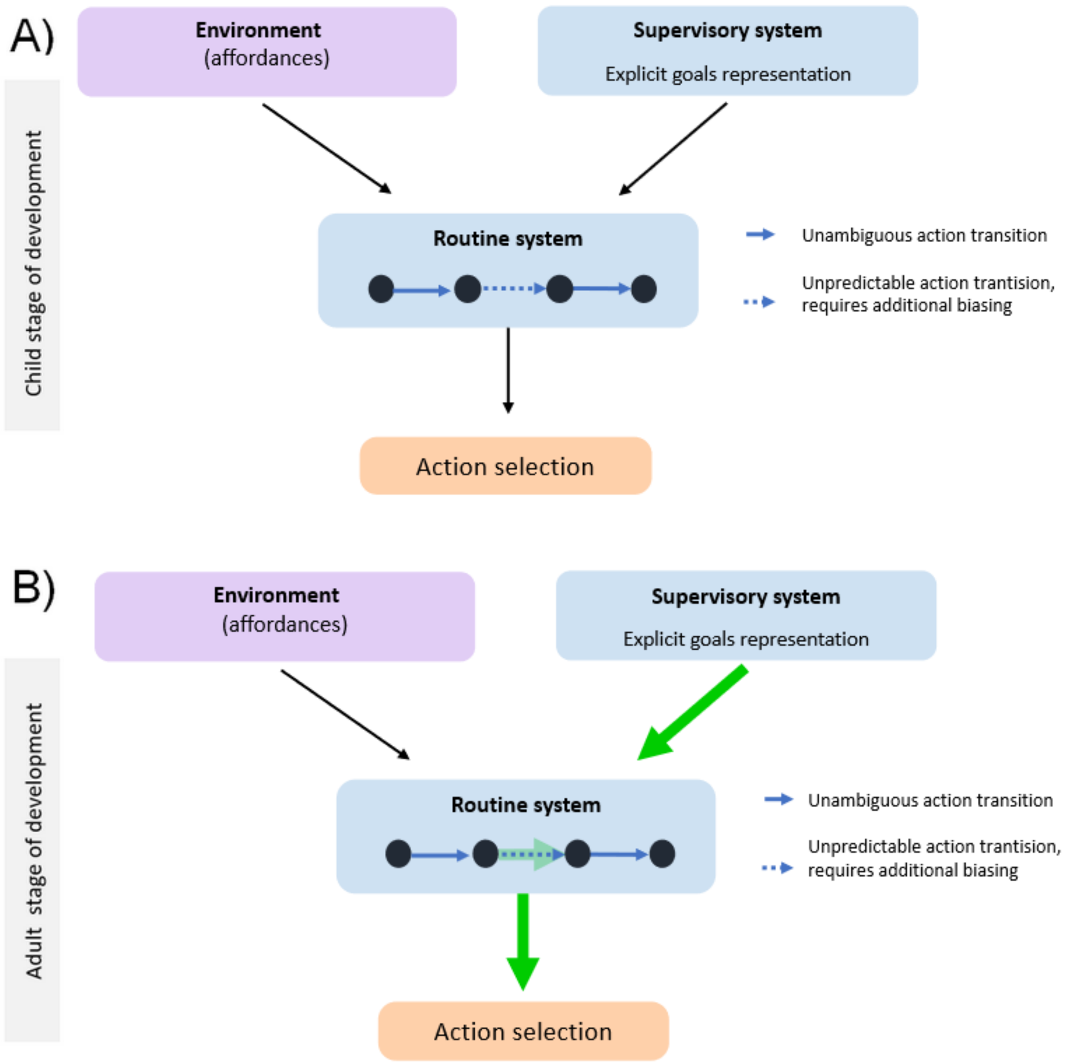


Fig. 6.2: **Summarised view of the differences between action systems within the Dual systems framework (Norman and Shallice, 1986, Cooper et al., 2014) at the early school-years child stage (A), and at the adult stage (B). It is proposed that the influences from the environment and the routine system are comparable in both stages. However, the maturation of the supervisory system brings about the necessary influences at various points, resulting in more efficient action control in adult's compared to children's stage. Arrows thickness represents how strong the influence is. The green arrows represent the developmental changes, including highlighting points at which stronger supervisory influence makes a bigger difference.**

6.4 Future research

6.4.1 Towards more ecological experimental paradigms

All three experimental studies measured kinematics (motion capture, touchscreen or mouse-tracking). The reaching task in Chapter 2 had the advantage of presenting real physical objects in a 3D space, nonetheless the action was the same throughout the task. Chapters 4 and 5 had the advantage of presenting actions embedded in sequence but that suffered from being virtual, and on a 2D surface (touchscreen or mouse-tracking). Chapter 2 revealed that the vertical deviation in reaching was affected by the surroundings (when past studies had only looked at 2 out of 3 dimensions in space), showing that all three dimensions were important measures of reaching. This also means that our 2D kinematic setups potentially miss out on interesting information. Future studies should therefore combine the advantages of the two types of experiments, with action sequences in a 3D realistic environment being routinised, while recording kinematics.

Another artificial aspect of the computer-based games is that every action consisted of going to and selecting icons (via click- or touch-and-drag), which makes the target selection and the act upon that target inseparable (this design choice had the advantage of making the response times directly comparable). In real-life, actions are varied (grasping, turning, pouring, etc.), and one may occasionally perform an inappropriate action to the appropriate object, or the appropriate action to an inappropriate object (an example, from a patient's diary presented in Schwartz (1995), is peeling a can of meatballs instead of peeling an orange). Bringing this dimension would be interesting in naturalistic paradigms. Nonetheless, we do not expect fundamental changes in how school-aged children manipulate objects, given findings that early on children are skilled at using tools and can perform anticipatory and fine adaptation to the type of grasp needed (Adolph and Berger, 2007).

Another point to highlight is that, in all the experiments of this thesis, participants were facing all the task stimuli, that were placed in a relatively limited visual angle (table or screen). However, in real life, action sequences' objects are often dispersed beyond the visual fields (e.g., in a kitchen, items can be on one's sides, behind and/or in cupboards), which adds a substantial object localisation component.

Inevitably, there are advantages and disadvantages to every setup: lab studies, real-life studies (e.g., action slip diaries), and the 'middle-grounds' such as lab studies reproducing real-life settings with less experimental control. Converging evidence from several types of setups is needed to validate a research finding. We have discussed in Chapters 5 how lab-based ecological paradigms bring indirect support for the idea that our findings are valid in real-life settings (and not just on a lab computerised task). Studies such as Yanaoka and Saito (2019)'s are more ecological than ours, but take place only over one session contrary to the four sessions in the task from Chapter 5. It would be interesting to investigate routine actions in more ecological context and with studies that span several days.

Another complementary approach would be to measure children performing actual routines of their daily lives (as opposed to invented sequences such as in Chapters 4 and 5), if possible in a longitudinal design. One could think of choosing a routine that children perform consistently at school, and measuring it across several months in a real school, and/or mimicking such a routine in the lab with standardised material that looks as close to the original as possible. Such studies should quantify performance or errors in action selection, action selection latencies and kinematics to validate our findings on the ability to learn routines and the increasing ability with age to engage supervisory control when appropriate within the routines. Nowadays, it would be technologically possible to place cameras in real life environments and use computer vision algorithms to partially-automatise analyses. The ethical concerns related to such investigation can be successfully addressed, as revealed by studies such as L. B. Smith et al.'s (2015), which used head-mounted cameras and eye-trackers and recorded hours of parents and child interactions in real home environments.

6.4.2 The role of objects' location and visual search in action sequences performance

As mentioned earlier, the object search component is missing from the thesis' studies. Research has found consistent patterns in familiar sequence execution (Land et al., 1999, Pelz and Canoza, 2001): objects are often located ahead in time (and memorised), and subsequently the object to be used is often fixated before the preceding action is terminated. The certainty and rapidity with which the visual system can direct towards the appropriate objects seems to be a crucial component of routine action (Land, 2006). A question to address is whether gazing at the target object of the next action happens earlier for non branch-points than branch points. If this is the case, the late gazing at branch points may contribute to the delay in selecting actions at branch points observed in Chapter 5. An alternative hypothesis would be that gazing at the object happens as early for branch points than non branch points, but that only the action execution takes longer at branch points.

Based on the findings of the thesis, we expect that more difficult action selection at branch points (which require biasing from the supervisory system) would be associated with alternative gazing between candidate options until a target action has been selected. Thus, we hypothesise that anticipatory gazing would happen at non branch points and would be absent at branch points.

Furthermore, in a food-preparation task by Land and Hayhoe (2001), adults exhibited a very low percentage of looking at task-irrelevant objects, supporting the idea that their gaze is driven more by the top-down goal of the task than by the salience of objects in the scene. In Chapter 2, we found distractor objects to affect younger children's reaching, which can be linked to difficulties in maintaining top-down control earlier in development (e.g., Freier et al., 2017). Relatedly, Freier et al. (2015) found that 3- to 5-year-olds incorporated irrelevant items in a sandwich-making task, but no longer when the task was facilitated by spatially arranging the items in the sequence temporal order. Taken together, if we investigated school-aged children naturalistic execution of a well-practised sequence, we can expect that school-aged children would not use, but would gaze at, irrelevant objects (on the contrary to adults, whose fixations have been found to alternate on several target objects but to avoid irrelevant objects). This would indicate an intermediate phase of development, during which top-down control is sufficient to avoid using distractors, but perhaps not sufficient to drive efficient gaze patterns such that of adults.

It is possible then that children would gaze more at irrelevant objects than adults, and it would be interesting to investigate the relationship between developmental changes in viewing patterns (that is, not looking at irrelevant objects, and anticipatory looking at the next target) and actions' accuracy and speed. We hypothesise that the same underlying developmental processes guide viewing patterns and action execution. We thus predict that gazing accuracy (that is, gazing more at relevant than irrelevant objects, and gazing in anticipation at the subsequent target) would improve simultaneously to actions' accuracy and speed.

Finally, using a task identical to that of Chapter 3, Konovalov and Krajbich (2016) found two particular gaze patterns in participants. One pattern was related to model-based choices (e.g., looking at the best option first, and only at that option; as if having made their decision early), and the other pattern was related to model-free valuation (looking at both options multiple times). This is somehow inconsistent with the idea that the more action is under routine (habitual, model-free) control, the more anticipatory is the gaze. This supports the claim that the two-stage task paradigm is not directly translatable to naturalistic action, but further research could clarify the role of anticipatory gaze in sequential action control.

All in all, adding eye-tracking measurements would allow a more comprehensive account of sequence action control that integrates the interaction between anticipatory search for target, object recognition and action.

6.4.3 The speed-accuracy trade-off in the involvement routine and non-routine modes

The role of timing should be further investigated because there is some empirical evidence for a change in the habitual/goal-directed balance related to speed-accuracy trade-offs. Indeed, reinforcement learning tasks that resemble the one in Chapter 3 have found that increasing the speed of response led to increased recruitment of habitual (routine) control, and inversely that slower responses led to increased goal-directed (non-routine) control (Hardwick et al., 2019). Indeed, the habitual (routine) system makes simpler computation and is consequently faster than the goal-directed (non-routine) one. Consequently, timing directly affects one's action control strategy. The study in Chapter 4 could be directly adjusted to influence the speed of responses (e.g., rewarding fast action more by attributing more coins to faster responses, or on the contrary 'blocking' the game for some instants before a step can be carried out) to investigate whether the patterns of errors following common and rare transitions would be altered. We predict that with shorter responding, the 'failure to update' errors would be more frequent (a consequence of the recruitment of the habitual mode of action) while, with longer responding, errors would be less frequent overall. We furthermore predict that the younger the children within the school-age range, the more time pressure would impair their performance. Thus, with the same amount of time pressure, children at the lower bound of school-age years would exhibit more 'failure to update' than children at the upper bound of school-age years.

In addition to affecting the habitual/goal-directed balance, speeding up actions is expected to increase the probability of errors within habitual or routine actions. For example, Schwartz (1995) cites an entry of the Manchester study, in which a participant with the habit of giving a biscuit to her dog when leaving for work, while putting on her earrings, has one day thrown the earrings to the dog and tried to attach a dog biscuit to her ear. This error is an example of the possible errors occurring when speeding up a routine.

6.4.4 Fractionating executive function involvement in sequential performance with different secondary tasks

The study in Chapter 5 uses a secondary task to tax the supervisory resources and consequently dissociate developmental changes in routine and supervisory modes of control. It is based on the assumption that the secondary task increases cognitive load as a whole, as opposed to specific components of supervisory control. However, one could argue that a task requiring to count sounds will tap into different resources from, for example, a task requiring to press a button at each sound occurrence, the latter not relying on memory. Indeed, Cooper et al. (2012) found evidence against the view that taxing resources by a secondary task increases cognitive load indistinctly. Instead, the main tasks in Cooper et al. (2012) were differentially impacted when the secondary task differed in their degree of set-shifting and monitoring components involvement.

It would be interesting to investigate what would be the different impacts on sequence execution of secondary tasks which tap into different cognitive components. We could expect, for example, that a secondary task which has non-negligible memory demands (like counting) would affect more strongly maintaining a higher-level goal in mind (which we know is still a struggle for 4 or 5 years old: Freier et al., 2017, Yanaoka and Saito, 2017; and presumably still later) than a task which has no memory demands. This would result specifically in more errors or slowing down of actions that cannot be predicted by preceding action and need to be related to the higher-level goal, like branch point actions (compared to non branch point actions). On the other hand, a secondary task without memory component but direct responding might slow down all actions equally. Indeed, the latter would require switching between the sequence production task and counting task at every sound, and this would be unaffected by the nature of the action.

6.4.5 Modelling the interaction of routine and non-routine modes of control and the role of developing executive functions

An outstanding line of research is to study whether the newly designed tasks of Chapter 4 and Chapter 5 can be modelled following the Interactive activation model of Chapter 3. It would be interesting to operationalise the role of inhibitory control in a similar fashion across all three tasks, in order to support a unified account of the development of inhibitory control driving changes in action control throughout the school-age years.

Such an account could be implemented with interaction-based processes, using an overall architecture similar to the GCM of Cooper et al. (2014), and that should contain at the minimum the following elements. Units would represent candidate actions at each point of the sequence (denoted *action units*). The actions units would receive direct activation from the objects corresponding to the action (environmental influences) and influences from well-learned representations of the actions of the sequence (either explicit, or implicit like in Cooper et al., 2014). The actions units would also receive influence from goal units that form the supervisory influence. Crucially, the supervisory influence would have a limit, for example, a maximal connection strength. This would implement the fact that, early on in development, the influence of the supervisory system may not be strong enough to overcome the influences from the environment (e.g., distractor objects), nor strong enough to inhibit schemas that are appropriate in another context. The connexion strength would increase across development, modelling the fact that, through development, the supervisory system can more effectively exert the appropriate control over action, in particular at point at which other influences are not sufficient for appropriate action selection.

The presented architecture considers the developmental differences in the efficiency of transferring goal units' input into action control. It assumed that goals units are already defined with respect to a context and goal. This is equal to assuming that the sequence of actions needed to achieve a goal can be learnt and retrieved explicitly, but might not be translated efficiently into action control (e.g., due to environment influences that may be competing with the explicit goals). This account leaves out the process of learning the relevant sequence of actions needed to achieve a goal. In Chapter 4 and Chapter 5, the actions were explicitly instructed to the participant and in real life, a number of sequences are indeed explicitly instructed (e.g., James et al., 1998).

Nonetheless, a number of lower-level actions may be learnt by trial-and-error. In Chapter 3 we saw how reward-based learning could occur and inform model-free and model-based types of action selection. In Chapter 3, learning occurred for each state and action of the sequence separately. To simulate this type of explorative learning in the real world, one must assume an endless number of states and may perform a wide range of possible actions. In that case, learning actions independently will lead to scaling issues; such issues are common on standard reinforcement learning (Botvinick et al., 2009). Hierarchical reinforcement learning may be a solution, and it has been proposed that the brain may implement such form of learning (Botvinick et al., 2009). This is another area for future investigation.

6.5 Conclusion

This thesis proposed new experimental and computational approaches to study routine sequential action control in school-age children and understand its changes across development.

With a variety of experimental contexts (reaching for an object among irrelevant distractors, updating an ongoing action sequence, selecting among multiple compatible actions), findings from the three experimental studies support the argument that in cases when an inappropriate action has to be inhibited, the younger the children are, the more they struggle to do so, resulting in less efficient action selection. Given the important role of executive functions in action selection process (and in particular of inhibitory control), and taken together with the known protracted maturation of the prefrontal cortex, the findings indicate that development in executive functions is a key driver of the developmental changes in sequential action performance.

More broadly, by developing a new computational model of the interaction of model-free and model-based systems as used in the reinforcement learning framework, and comparing existing theoretical accounts of the functioning of two systems for action control, we showed how the dual-systems accounts is suited to explain school-aged children's action control and proposed a developmental trajectory followed by the systems. Early on during school-age years, children appear to have a well-developed routine system that controls well-learned action efficiently at non-challenging points of sequences and under no cognitive load. Throughout development, the supervisory system becomes more efficient and can contribute better to action selection at challenging points, resulting in finer sequential action control.

We hope that the new methods developed (computational modelling, task development and kinematic analyses) as well as the evidence for the applicability of the dual-systems framework to understand children's action control will help future research on children's sequential routine action control, and its change across development.

Appendix A

Appendix

A.1 Appendix for Chapter 4: Action Sequence Control following an Unexpected Change: Two Touchscreen Studies

A.1.1 Script of the instruction videos for the touchscreen game

This is the script read aloud and video recorded to present the instructions of the game of Chapter 4. There was a video presented at the start of the training phase, lasting 1mn44, and one presented at the start of the test phase, lasting 1mn08. The display on the right was directly matching was is being said. We present some screenshots of the video at different times in Figure A.1.

A.1.1.1 Training phase

Note: the same script was existed in two version: the one presented here, and another one where the first cake presented ('white chocolate with grape') is associated to the cat instead of the bunny (and the other one associated to be bunny). The mentions of the cat and bunny are swapped to match the two existing mapping between the characters and their favourite recipes.

"In a moment we are going to play a game together on this touchscreen. We're going to bake cakes for fluffy the bunny, and kitty the cat.

In the kitchen, you will see the following ingredients: white chocolate, dark chocolate, pear, raspberry and grape.

The pets really like cakes and want to eat a lot of them. But they like some more than others.

- The bunny likes the white chocolate with grape; that's its very favourite recipe in the world. Every time you see the bunny, you're gonna make the white chocolate and grape cake.
- The cat likes the dark chocolate with pear; that's its very favourite recipe in the whole wide world. Every time you see the cat, you're gonna make the dark chocolate with pear cake.

To make the cake, we will place the ingredients one by one. Before we decorate the cake with the fruit, we're going to make the chocolate mix. First, we need to pick up the spoon, collect the chocolate, give it to the pet. The pet will place it in the bowl.

Then we can pick up the spoon, get the fruit and place it in the bowl.

Remember they're very hungry, and they want to eat it as soon as possible, so try and make it as quickly as you can.

When the bowl disappears, it means the pet got too hungry and went to the shop because it couldn't wait any longer. But don't worry, they will be back and we can try to make another cake.

Remember, the bunny's favourite cake in the world is the white chocolate and grape cake; and the cat's favourite cake in the world is the dark chocolate and pear cake. Let's start!"

A.1.1.2 Test phase

"Now, we're going to play a similar game, but this time, the pets will attempt to play tricks on you.

They have switched around some of the chocolate wrappers, and when this happens, the chocolate will reveal its true colour in the mixing bowl. This means that, even if you take the white chocolate wrapper, it can be dark chocolate inside. When this happens, you need to make the pet's second favourite cake.

The pets second favourite cake is the one with raspberry. Even if the bunny prefers white chocolate with grape, the bunny will still be happy with dark chocolate and raspberry. The very same goes for the cat.

Sometimes, the bunny and cat will swap places with each other after you've started the cake. They're trying to trick you out to see if you can remember what cakes they like. When this happens, the pets second favourite cake is the one with raspberry.

Now you're ready to play the game! Remember to try and make their very favourite recipe like you did in the first game. But pay close attention to the tricks they are trying to play on you."

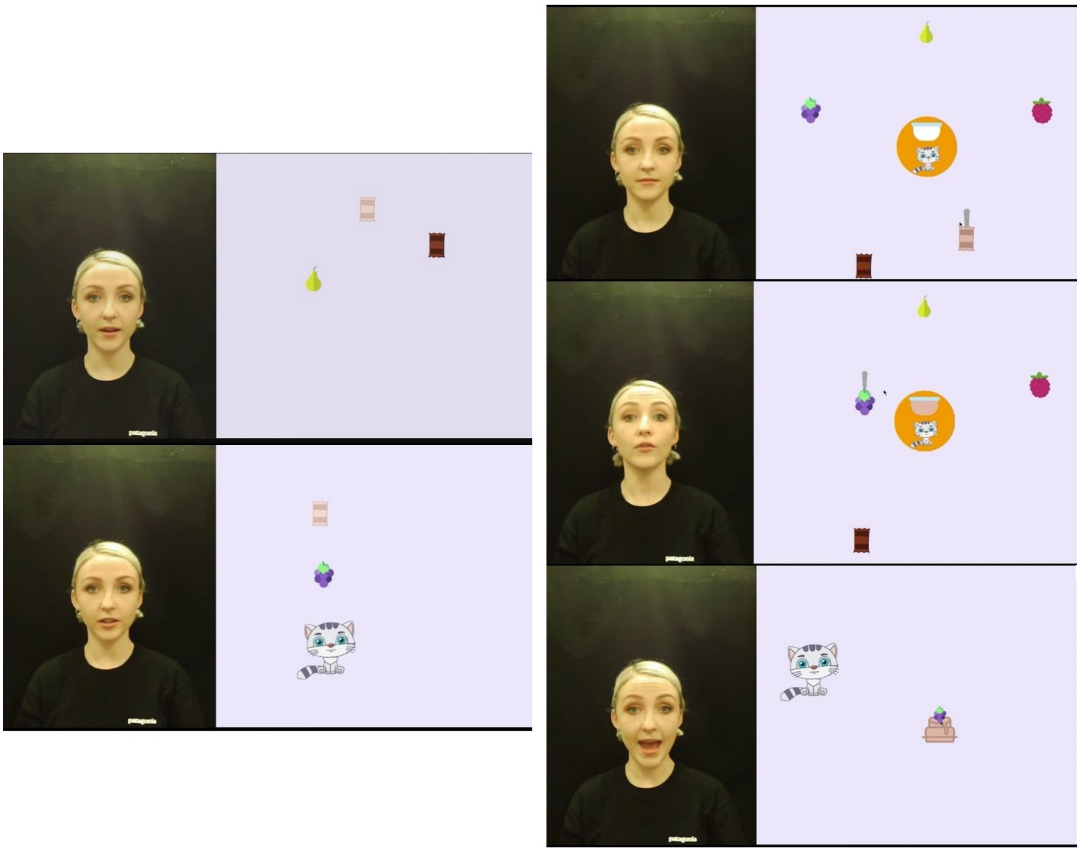


Fig. A.1: **Some snapshots of the instruction videos for the touchscreen game, presenting the items in the game, the goals (left hand side snapshots) and showing the game being played (at step 1, 2 and 3).**

A.2 Appendix for Chapter 5: Hierarchical Control of Action Sequences: Two Online Studies

A.2.1 Experiment 3

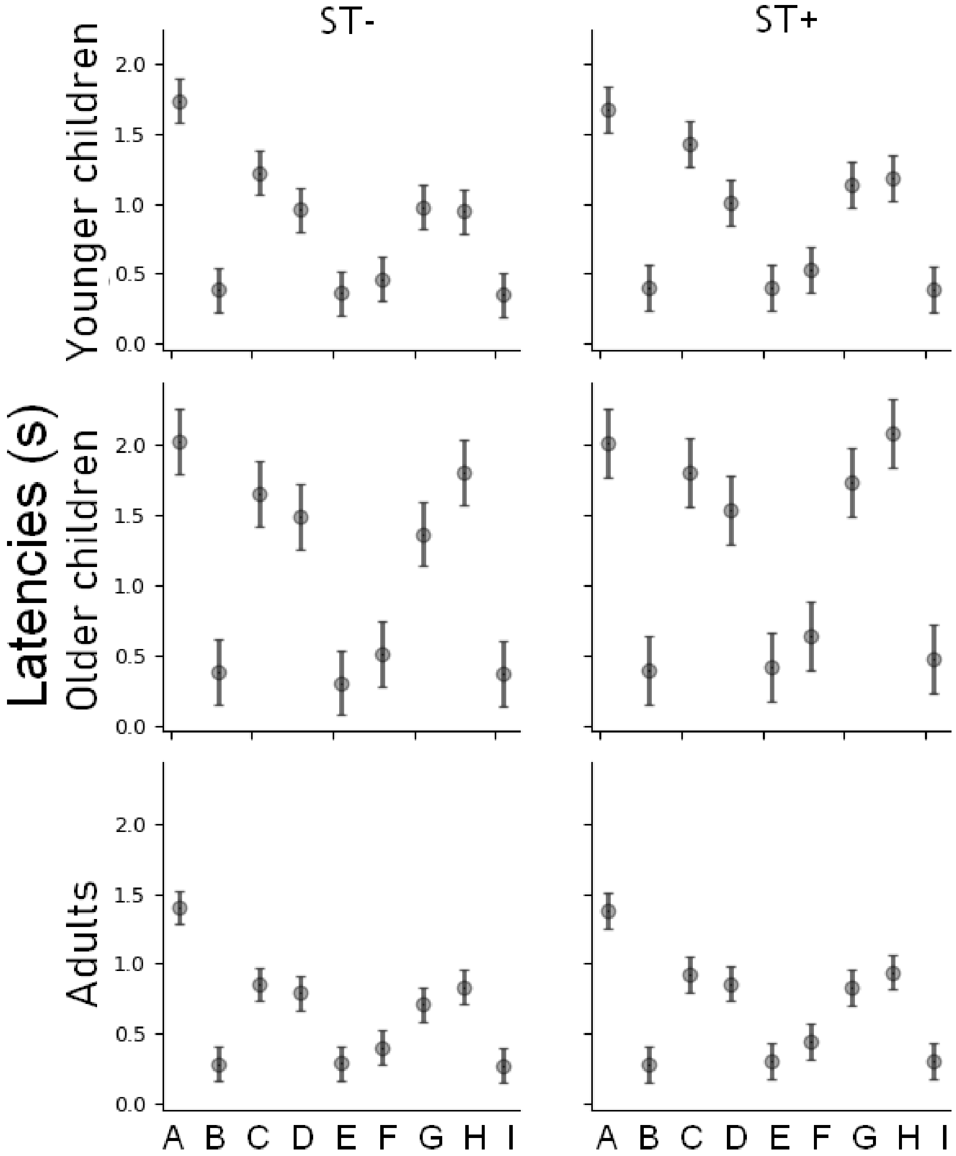


Fig. A.2: **Median step-level action selection latencies in sessions 3 and 4 across age groups, without secondary task (ST_-; left) or with (ST_+; right). Dots indicate the median and error bars indicate s.e.m.**

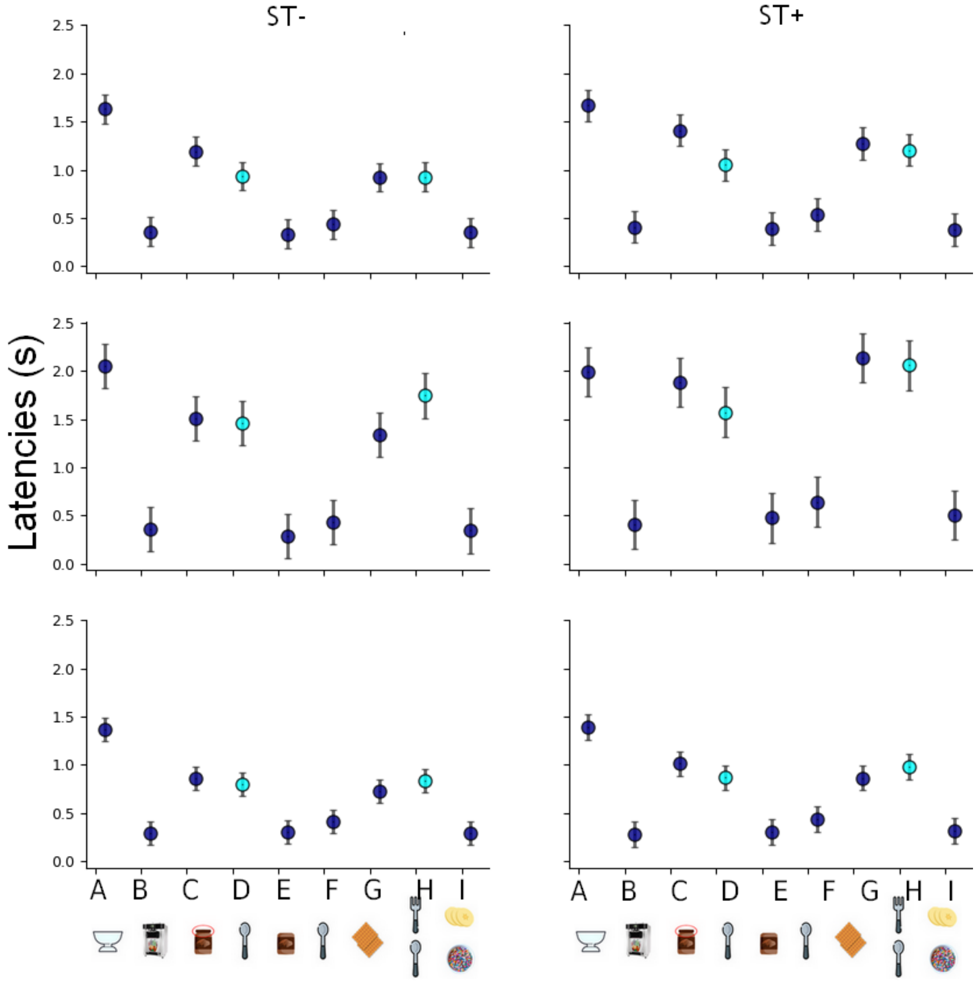


Fig. A.3: **For the sequences with main goal 'spoon-sprinkles' ice-cream, median step-level action selection latencies in sessions 3 and 4 across age groups, without secondary task (ST_-; left) or with (ST_+; right). The cyan-blue dots highlights the steps involving picking-up the spoon. Dots indicate the median and error bars indicate s.e.m.**

Table A.1: **Linear mixed model regression (LMM) coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_+/ST_-) on the action selection latency at Step F.**

	β	CI (l)	CI (u)	Stde	DF	t	p	sig
(Intercept)	-0.81	-0.95	-0.66	0.07	527	-10.89	<0.001	***
ST_+-ST_-	0.09	-0.02	0.20	0.06	527	1.60	0.111	
$ageGp = 9 - 11$	-0.09	-0.37	0.19	0.14	42	-0.67	0.508	
$ageGp = 7 - 8$	0.33	0.01	0.64	0.16	42	2.10	0.042	*
ST_+-ST_-:$ageGp = 9 - 11$	0.43	0.21	0.65	0.11	527	3.79	<0.001	***
ST_+-ST_-:$ageGp = 7 - 8$	-0.01	-0.28	0.27	0.14	527	-0.04	0.97	

β: Estimate (regression coefficient), *CI (l)*: Confidence Interval (lower bound), *CI (u):* Confidence Interval (upper bound), *SE*: standard error, *p*: unadjusted p-value, *p (FDR)*: False Discovery Rate-adjusted p-value, *sig.*: significance levels based on FDR-adjusted p-values:.: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Table A.2: **Linear mixed model regression (LMM) coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_+/ST_-) on the action selection latency at step A.**

	β	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	0.30	0.18	0.41	0.06	527	4.99	<0.001	***
ST_+-ST_-	0.03	-0.03	0.10	0.03	527	0.97	0.331	
$ageGp = 9 - 11$	0.13	-0.09	0.36	0.11	42	1.22	0.229	
$ageGp = 7 - 8$	0.54	0.30	0.77	0.12	42	4.55	<0.001	***
ST_+-ST_-:$ageGp = 9 - 11$	0.27	0.13	0.41	0.07	527	3.90	<0.001	***
ST_+-ST_-:$ageGp = 7 - 8$	-0.01	-0.18	0.16	0.09	527	-0.15	0.878	

Table A.3: **LMM coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_+/ST_-) on the action selection latency at step B.**

	β	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	-1.42	-1.59	-1.26	0.08	527	-16.83	<0.001	***
ST_+-ST_-	0.19	0.10	0.28	0.04	527	4.20	<0.001	***
$ageGp = 9 - 11$	0.31	-0.00	0.62	0.15	42	2.02	0.05	.
$ageGp = 7 - 8$	0.19	-0.14	0.53	0.16	42	1.18	0.245	
ST_+-ST_-:$ageGp = 9 - 11$	0.09	-0.08	0.27	0.09	527	1.04	0.299	
ST_+-ST_-:$ageGp = 7 - 8$	0.31	0.09	0.53	0.11	527	2.77	0.006	**

Table A.4: **LMM coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_+/ST_-) on the action selection latency at step C.**

	β	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	-0.04	-0.19	0.12	0.08	527	-0.47	0.635	
ST_+-ST_-	-0.01	-0.12	0.10	0.06	527	-0.12	0.905	
$ageGp = 9 - 11$	0.20	-0.09	0.50	0.15	42	1.38	0.176	
$ageGp = 7 - 8$	0.61	0.28	0.94	0.16	42	3.76	0.001	***
ST_+-ST_-:$ageGp = 9 - 11$	0.40	0.18	0.63	0.11	527	3.52	<0.001	***
ST_+-ST_-:$ageGp = 7 - 8$	0.16	-0.11	0.44	0.14	527	1.16	0.245	

Table A.5: **LMM coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_+/ST_-) on the action selection latency at step D.**

		Value	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	-0.11	-0.27	0.05	0.08	528	-1.35	0.178		
ST_+-ST_-	-0.15	-0.24	-0.05	0.05	528	-2.96	0.003	**	
$ageGp = 9 - 11$	0.04	-0.26	0.34	0.15	42	0.26	0.798		
$ageGp = 7 - 8$	0.58	0.25	0.90	0.16	42	3.57	0.001	***	
ST_+-ST_-:$ageGp = 9 - 11$	0.42	0.22	0.62	0.10	528	4.11	<0.001	***	
ST_+-ST_-:$ageGp = 7 - 8$	0.08	-0.17	0.32	0.12	528	0.62	0.536		

Table A.6: **LMM coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_+/ST_-) on the action selection latency at step E.**

	β	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	-1.33	-1.53	-1.14	0.10	528	-13.71	<0.001	***
ST_+-ST_-	0.13	0.04	0.21	0.04	528	2.96	0.003	**
$ageGp = 9 - 11$	0.23	-0.13	0.59	0.18	42	1.30	0.199	
$ageGp = 7 - 8$	0.27	-0.10	0.64	0.18	42	1.47	0.15	
ST_+-ST_-:$ageGp = 9 - 11$	0.09	-0.09	0.26	0.09	528	0.99	0.325	
ST_+-ST_-:$ageGp = 7 - 8$	0.03	-0.18	0.25	0.11	528	0.30	0.763	

Table A.7: **LMM coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_{+}/ST_{-}) on the action selection latency at step G.**

	β	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	0.14	-0.03	0.31	0.08	527	1.65	0.1	
ST_{+}-ST_{-}	-0.38	-0.49	-0.26	0.06	527	-6.24	<0.001	***
$ageGp = 9 - 11$	0.19	-0.13	0.51	0.16	42	1.21	0.234	
$ageGp = 7 - 8$	0.42	0.07	0.78	0.18	42	2.40	0.021	*
ST_{+}-ST_{-}:$ageGp = 9 - 11$	0.34	0.10	0.58	0.12	527	2.76	0.006	**
ST_{+}-ST_{-}:$ageGp = 7 - 8$	0.19	-0.10	0.49	0.15	527	1.29	0.198	

Table A.8: **LMM coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_{+}/ST_{-}) on the action selection latency at step H.**

	β	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	0.06	-0.09	0.21	0.08	527	0.74	0.46	
ST_{+}-ST_{-}	-0.16	-0.27	-0.06	0.05	527	-3.21	0.001	**
$ageGp = 9 - 11$	0.09	-0.20	0.37	0.14	42	0.60	0.552	
$ageGp = 7 - 8$	0.69	0.37	1.00	0.16	42	4.38	<0.001	***
ST_{+}-ST_{-}:$ageGp = 9 - 11$	0.39	0.19	0.60	0.10	527	3.75	<0.001	***
ST_{+}-ST_{-}:$ageGp = 7 - 8$	-0.04	-0.29	0.21	0.13	527	-0.32	0.753	

Table A.9: **LMM coefficients indicating the effect of age group (ageGp) and presence/absence of secondary task (ST_{+}/ST_{-}) on the action selection latency at step I.**

	β	CI (l)	CI (u)	SE	DF	t	p	sig
(Intercept)	-1.26	-1.45	-1.07	0.10	527	-12.96	<0.001	***
ST_{+}-ST_{-}	-0.04	-0.13	0.05	0.05	527	-0.88	0.377	
$ageGp = 9 - 11$	0.20	-0.15	0.56	0.18	42	1.16	0.253	
$ageGp = 7 - 8$	0.45	0.08	0.82	0.19	42	2.44	0.019	*
ST_{+}-ST_{-}:$ageGp = 9 - 11$	0.17	-0.01	0.35	0.09	527	1.82	0.07	.
ST_{+}-ST_{-}:$ageGp = 7 - 8$	-0.18	-0.40	0.05	0.11	527	-1.54	0.123	

A.2.2 Experiment 4

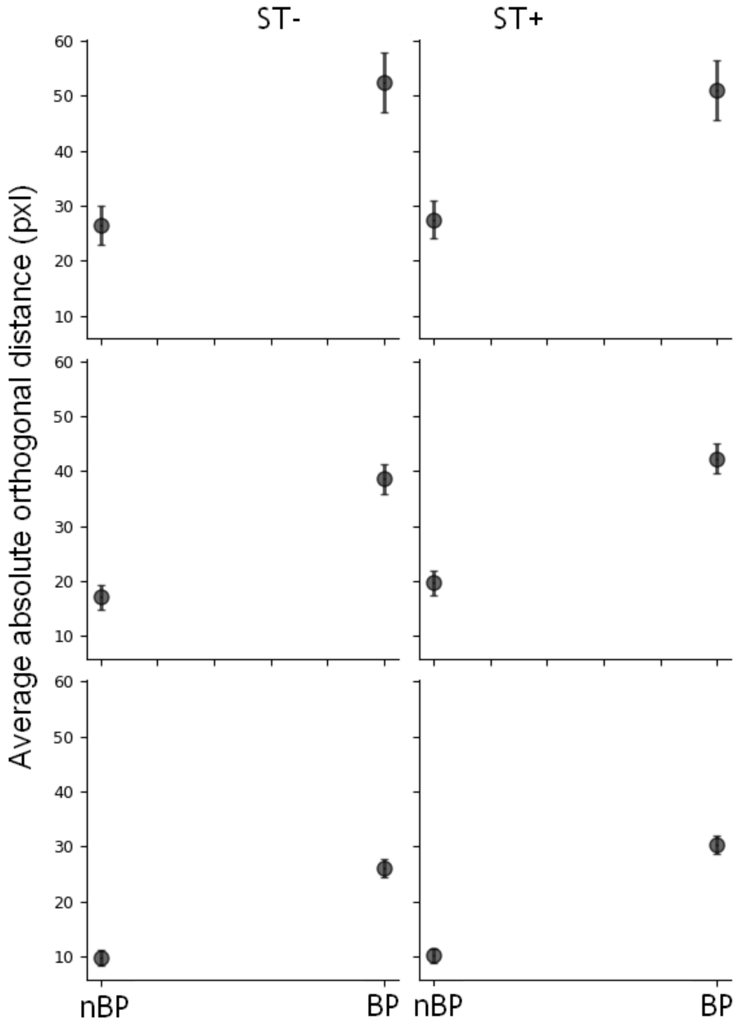


Fig. A.4: **Deviation: Average absolute orthogonal distance to straight path). Median values. Error bars represent s.e.m.**

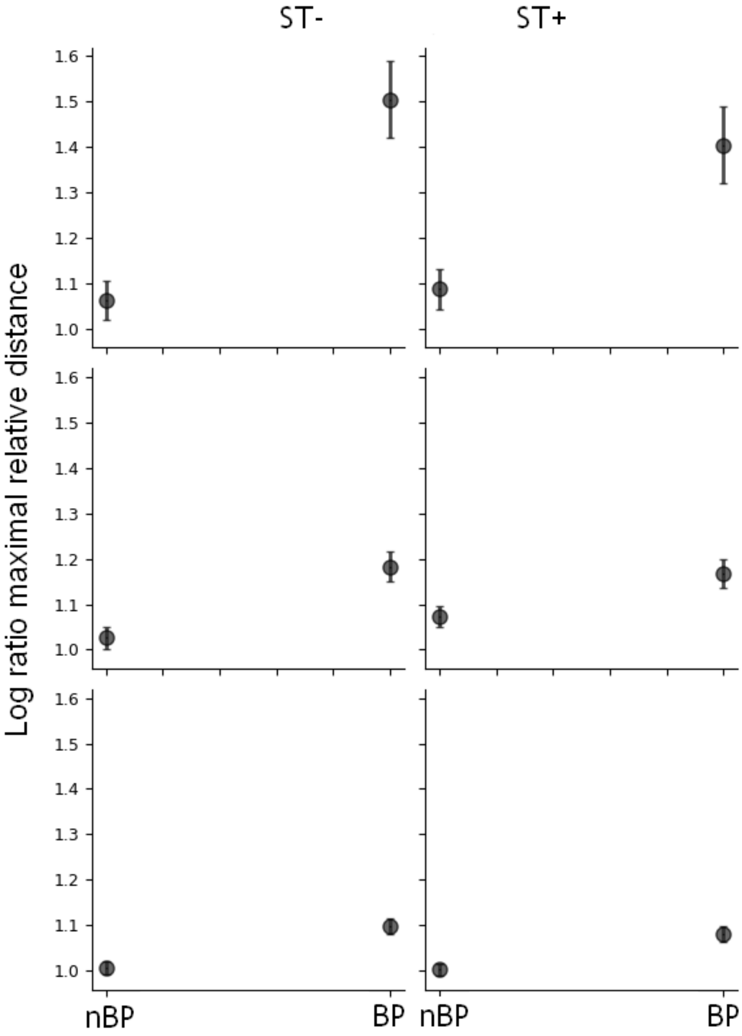


Fig. A.5: **Deviation: Log ratio maximal relative distance (maxDist/dist start end). Median values. Error bars represent s.e.m.**

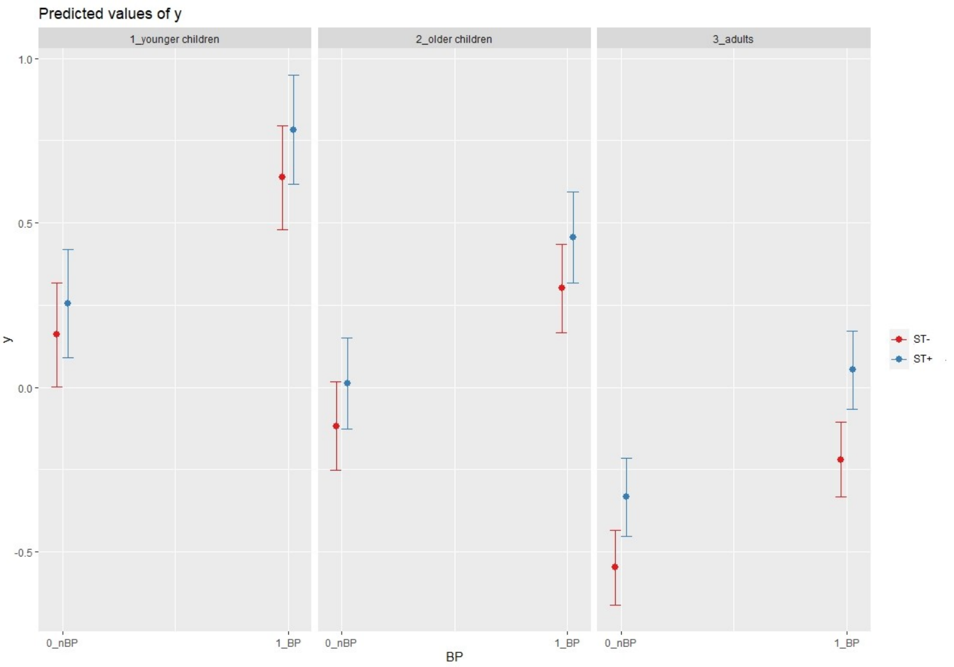


Fig. A.6: **Linear Mixed Model predicted values of the Step latencies. ST_{-}: without secondary task, ST_{+}: with secondary task. Error bars represent 95% confidence intervals.**

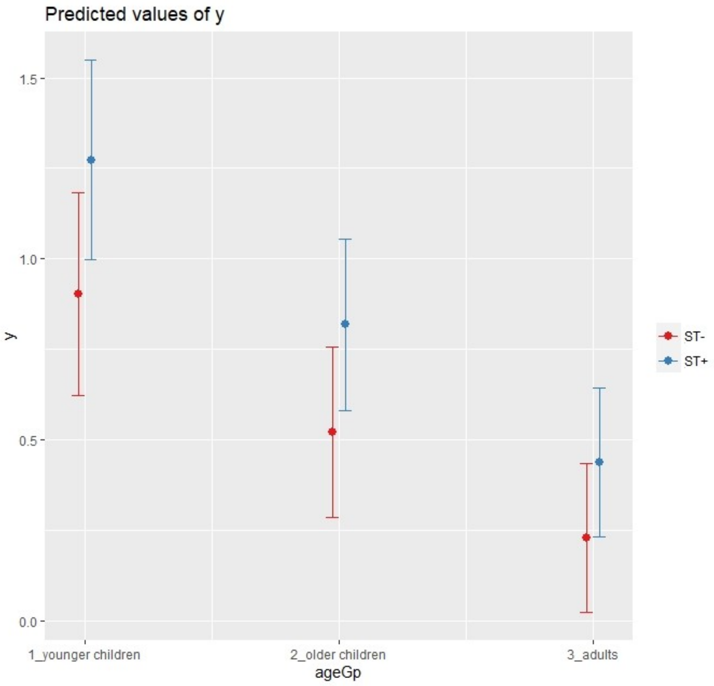


Fig. A.7: **Linear Mixed Model predicted values of the $Delta\ BP - nBP$. ST_{-}: without secondary task, ST_{+}: with secondary task. Error bars represent 95% confidence intervals.**

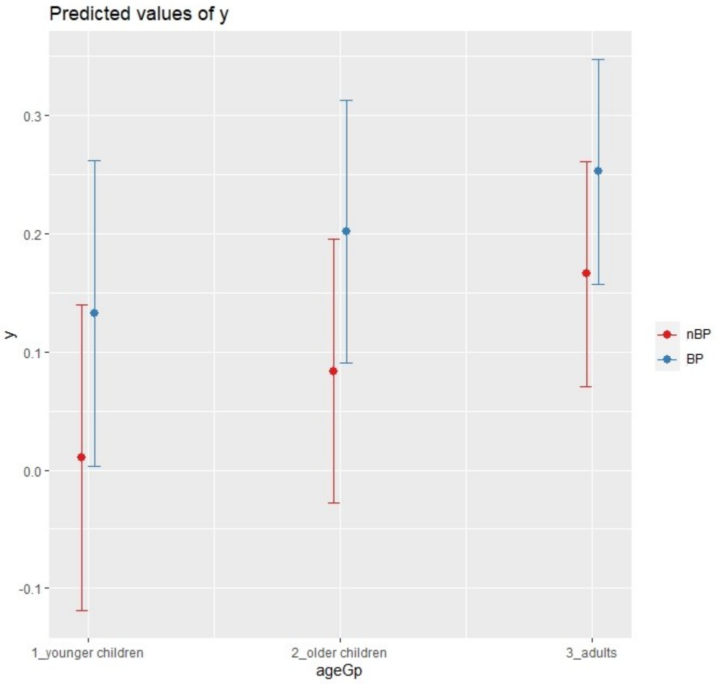


Fig. A.8: **Linear MixedModel predicted values of the $DeltaST_{+} - ST_{-}$. ST_{-}: without secondary task, ST_{+}: with secondary task. BP: branch points, nBP: non branch points. Error bars represent 95% confidence intervals.**

Bibliography

Abdul, A., Carroll, D. J., Andrews, K., & Wiebe, S. A. (2017). Neural Correlates of Response Inhibition in Early Childhood : Evidence From a Go / No-Go Task. *Developmental Neuropsychology*, *42*(5), 336–350. https://doi.org/10.1080/87565641.2017.1355917

Adolph, K. E. (2008). Learning to move. *Current Directions in Psychological Science*, *17*, 213 218.

Adolph, K. E., & Berger, S. E. (2007). Motor Development. In *Handbook of Child Psychology*. American Cancer Society. https://doi.org/10.1002/9780470147658.chpsy0204

Adolph, K. E., & Berger, S. E. (2011). Physical and motor development. In M. Bornstein & M. Lamb (Eds.), *Developmental science: An advanced textbook* (pp. 241–302). Psychology Press.

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J., & Brown, J. (2011). A model-based fmri analysis with hierarchical bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 95–110. https://doi.org/10.1037/a0020684

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the Automated Working Memory Assessment. *Educational Psychology*, *28*(7), 725–734. https://doi.org/10.1080/01443410802243828

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action.in h. In Heuer & A. Sanders (Eds.). Hillsdale, NJ, Erlbaum.

Anderson, J. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, *94*, 192–210.

Anwyl-Irvine, A. L., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2020). Online participants in the wild: Realistic precision & accuracy of platforms, web-browsers, and devices. *in press, PsyArxiv Preprints.*

Arnold, A., Wing, A. M., & Rotshtein, P. (2017). Building a lego wall: Sequential action selection. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(5), 847–852. https://doi.org/10.1037/xhp0000382

Baldauf, D., Wolf, M., & Deubel, H. (2006). Deployment of visual attention before sequences of goal-directed hand movements. *Vision Research*, *46*(26), 4355–4374. https://doi.org/10.1016/j.visres.2006.08.021.

Baldauf, D., & Deubel, H. (2009). Attentional Selection of Multiple Goal Positions Before Rapid Hand Movement Sequences: An Event-related Potential Study.

Journal of Cognitive Neuroscience, *21*(1), 18–29. https://doi.org/10.1162/jocn.2008.21021

Ballard, I. C., & McClure, S. M. (2019). Joint modeling of reaction times and choice improves parameter identifiability in reinforcement learning models. *Journal of Neuroscience Methods*, *317*, 37–44. https://doi.org/10.1016/j.jneumeth.2019.01.006

Barrett, T., Traupman, E., & Needham, A. (2008). Infants' visual anticipation of object structure in grasp planning. *Infant Behavior and Development*, *31*, 1–9. https://doi.org/10.1016/j.infbeh.2007.05.004

Basso, M., & Wurtz, R. (1997). Modulation of neuronal activity by target uncertainty. *Nature*, *389*(6646), 66–69. https://doi.org/10.1038/37975.

Beck, M., Lohrenz, M., & Trafton, J. (2010). Measuring search efficiency in complex visual search tasks: Global and local clutter. *Journal of Experimental Psychology: Applied*, *16*(3), 238–250.

Benedetti, V., Gavazzi, G., Giovannelli, F., Bravi, R., Giganti, F., Minciacchi, D., Mascalchi, M., Cincotta, M., & Viggiano, M. P. (2020). Mouse tracking to explore motor inhibition processes in go/no-go and stop signal tasks. *Brain Sciences*, *10*(7), 1–15. https://doi.org/10.3390/brainsci10070464

Berns, G., McClure, S., Pagnoni, G., & Montague, P. (2001). Predictability modulates human brain response to reward. *J. Neurosci*, *21*, 2793–2798.

Bettencourt, K., & Somers, D. (2009). Effects of target enhancement and distractor suppression on multiple object tracking capacity. *Journal of Vision*, *9*(7). https://doi.org/10.1167/9.7.9.

Blakemore, S., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(3-4), 296–312.

Blaye, A., & Chevalier, N. (2011). The role of goal representation in preschoolers flexibility and inhibition. *Journal of Experimental Child Psychology*, *108*, 469–483.

Bock, O., & Eversheim, U. (2000). The mechanisms of movement preparation: A precuing study. *Behavioural Brain Research*, *108*(1), 85–90. https://doi.org/10.1016/S01664328(99)00134-5.

Bolenz, F., Reiter, A. M. F., & Eppinger, B. (2017). Developmental Changes in Learning: Computational Mechanisms and Social Influences. *Frontiers in Psychology*, *0*. https://doi.org/10.3389/fpsyg.2017.02048

Borst, J. P., Nijboer, M., Taatgen, N. A., Van Rijn, H., & Anderson, J. R. (2015). Using data-driven model-brain mappings to constrain formal models of cognition. *PLoS ONE*, *10*(3), 5–7. https://doi.org/10.1371/journal.pone.0119673

Bos, W., Cohen, M., Kahnt, T., & Crone, E. (2012). Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cereb. Cortex*, *22*, 1247–1255. https://doi.org/10.1093/cercor/bhr198

Botvinick, M. M. (2007). Multilevel structure in behaviour and in the brain: A model of Fuster's hierarchy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1485), 1615–1626. https://doi.org/10.1098/rstb.2007.2056

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, *12*(5), 201–208. https://doi.org/10.1016/j.tics.2008.02.009

Botvinick, M. M. (2012). Hierarchical reinforcement learning and decision making. *Current Opinion in Neurobiology*, *22*(6), 956–962. https://doi.org/10.1016/j.conb.2012.05.008

Botvinick, M. M., & Bylsma, L. (2005). Distraction and action slips in an everyday task: Evidence for a dynamic representation of task context. *Psychonomic Bulletin & Review*, *12*(6), 1011–1017. https://doi.org/10.3758/BF03206436

Botvinick, M. M., Carter, C. S., Braver, T. S., Barch, D. M., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, *113*(3), 262–280. https://doi.org/10.1016/j.cognition.2008.08.011

Botvinick, M. M., & Plaut, D. C. (2004). Doing Without Schema Hierarchies: A Recurrent Connectionist Approach to Normal and Impaired Routine Sequential Action. *Psychological Review*, *111*(2), 395–429. https://doi.org/10.1037/0033-295X.111.2.395

Bradley, J. V. (1969). Optimum Knob Crowding. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *11*(3), 227–237. https://doi.org/10.1177/001872086901100304

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, 1–29. https://doi.org/10.7717/peerj.9414

Broadbent, H. J., Osborne, T., Rea, M., Peng, A., Mareschal, D., & Kirkham, N. Z. (2018). Incidental category learning and cognitive load in a multisensory environment across childhood. *Developmental Psychology*, *54*(6), 1020–1028. https://doi.org/10.1037/dev0000472

Bruner, J. S. (1973). Organization of Early Skilled Action. *Child Development*, *44*(1), 1–11. http://www.jstor.org/stable/1127671

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence*, *47*, 34–43. https://doi.org/10.1016/j.intell.2014.08.010

Bullock, M., & Lutkenhaus, P. (1988). The development of volitional behavior in the toddler years. *Child Development*, *59*, 664–674.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). Hove, UK, Psychology Press.

Burgess, P. W., & Stuss, D. T. (2017). Fifty Years of Prefrontal Cortex Research: Impact on Assessment. *Journal of the International Neuropsychological Society*, *23*(9-10), 755–767. https://doi.org/10.1017/S1355617717000704

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (2006). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing*. https://epubs.siam.org/doi/abs/10.1137/0916069

Carlson, S., Moses, L., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? contributions of inhibitory control and working memory. *Infant and Child Development*, *11*, 73–92.

Carlson, S., & Zelazo, P. (2014). *Minnesota executive function scale test manual*. Saint Paul, MN, Reflection Sciences.

Carter, c. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior Cingulate Cortex , Error Detection , and the Online Monitoring of Performance, *280*(May), 747–749.

Carver, A. C., Livesey, D. J., & Charles, M. (2001). Age Related Changes in Inhibitory Control as Measured by Stop Signal Task Performance. *International Journal of Neuroscience*, *107*(1-2), 43–61. https://doi.org/10.3109/00207450109149756

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005a). Imaging the developing brain: What have we learned about cognitive development? *Trends in Cognitive*, *9*(3), 104–110.

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005b). Imaging the developing brain: What have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3 SPEC. ISS.), 104–110. https://doi.org/10.1016/j.tics.2005.01.011

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., Castellanos, F. X., Haxby, J. V., Noll, D. C., Cohen, J. D., Forman, S. D., Dahl, R. E., & Rapoport, J. L. (1988). A Developmental Functional MRI Study of Prefrontal Activation during Performance of a Go-No-Go Task.

Castiello, U. (1996). Grasping a fruit: Selection for action. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(3), 582–603. https://doi.org/10.1037/0096-1523.22.3.582

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

Chang, S. W., & Abrams, R. (2004). Hand movements deviate toward distracters in the absence of response competition. *The Journal of general psychology*, *131*(4), 328–44.

Chapman, C., & Goodale, M. (2008). Missing in action: The effect of obstacle position and size on avoidance while reaching. *Experimental Brain Research*, *191*(1), 83–97. https://doi.org/10.1007/s00221-008-1499-1.

Chen, Y.-p., Keen, R., Rosander, K., & Von Hofsten, C. (2010). Movement Planning Reflects Skill Level and Age Changes in Toddlers. *Child Development*, *81*(6), 1846–1858. https://doi.org/10.1111/j.1467-8624.2010.01514.x

Cheng, P. C.-H., & van Genuchten, E. (2018). Combinations of Simple Mechanisms Explain Diverse Strategies in the Freehand Writing of Memorized Sentences. *Cognitive Science*, *42*(4), 1070–1109. https://doi.org/10.1111/cogs.12606

Chevalier, N., & Blaye, A. (2009). Setting Goals to Switch Between Tasks: Effect of Cue Transparency on Children's Cognitive Flexibility. *Developmental Psychology*, *45*(3), 782–797. https://doi.org/10.1037/a0015409

Chu, V. W., Park, S.-W., Sanger, T. D., & Sternad, D. (2016). Children With Dystonia Can Learn a Novel Motor Skill: Strategies That are Tolerant to High Variability. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *24*(8), 847–858. https://doi.org/10.1109/TNSRE.2016.2521404

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: A computational model. *Journal of Neuroscience*, *26*(38), 9761–9770. https://doi.org/10.1523/JNEUROSCI.5605-05.2006.

Cisek, P., & Kalaska, J. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron*, *45*(5), 801–814. https://doi.org/10.1016/j.neuron.2005.01.027.

Claxton, L. J., Keen, R., & McCarty, M. E. (2003). Evidence of motor planning in infant reaching behavior. *Psychological Science*, *14*(4), 354–356. https://doi.org/10.1111/1467-9280.24421

Clifton, R., Rochat, P., Litovsky, R., & Perris, E. (1991). Object representation guides infants' reaching in the dark. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 323–329. https://doi.org/10.1037/0096-1523.17.2.323

Clifton, R., Rochat, P., Robin, D., & Berthier, N. (1994). Multimodal perception in the control of infant reaching.journal of experimental psychology. *Human Perception and Performance*, *20*, 876–886.

Cohen, J., Barch, D., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology*, *108*, 120–133.

Collins, A., & Frank, M. (2012). How much of reinforcement learning is working memory, not reinforcement learning? a behavioral, computational, and neurogenetic analysis. *Eur. J. Neurosci*, *35*, 1024–1035.

Coluccini, M., Maini, E. S., Martelloni, C., Sgandurra, G., & Cioni, G. (2007). Kinematic characterization of functional reach to grasp in normal and in motor disabled children. *Gait & Posture*, *25*(4), 17270446, 493–501. https://doi.org/10.1016/j.gaitpost.2006.12.015

Cooper, R. P., Fox, J., Farringdon, J., & Shallice, T. (1996). A systematic methodology for cognitive modelling. *Artificial Intelligence*, *85*(1), 3–44. https://doi.org/10.1016/0004-3702(95)00112-3

Cooper, R. P., & Guest, O. (2014). Implementations are not specifications: Specification, replication and experimentation in computational cognitive modeling. *Cognitive Systems Research*, *27*, 42–49. https://doi.org/10.1016/j.cogsys.2013.05.001

Cooper, R. P., Ruh, N., & Mareschal, D. (2014). The goal circuit model: A hierarchical multi-route model of the acquisition and control of routine sequential action in humans. *Cognitive Science*, *38*(2), 244–274. https://doi.org/10.1111/cogs.12067

Cooper, R. P., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, *17*(4), 297–338. https://doi.org/10.1080/026432900380427

Cooper, R. P., & Shallice, T. (2006). Hierarchical Schemas and Goals in the Control of Sequential Behavior, *113*(4), 887–916. https://doi.org/10.1037/0033-295X.113.4.887

Cooper, R. P., Wutke, K., & Davelaar, E. J. (2012). Differential contributions of set-shifting and monitoring to dual-task interference. *Quarterly Journal of Experimental Psychology*, *65*(3), 587–612. https://doi.org/10.1080/17470218.2011.629053

Courtney, S. (2004). Attention and cognitive control as emergent properties of information representation in working memory. *Cogn. Affect. Behav. Neurosci*, *4*, 501–516.

Cox, R. F., & Smitsman, A. W. (2006). Action planning in young children's tool use. *Developmental Science*, *9*(6), 628–641. https://doi.org/10.1111/j.1467-7687.2006.00541.x

Cragg, L., & Nation, K. (2009). Shifting development in mid-childhood: The influence of between-task interference. *Developmental Psychology*, *45*(5), 1465–1479. https://doi.org/10.1037/a0015360

Craighero, L., Fadiga, L., Umiltà, C., & Rizzolatti, G. (1996). Evidence for visuomotor priming effect. *NeuroReport*, *8*, 347–349.

Crone, E. A., Bunge, S. A., Van Der Molen, M. W., & Richard Ridderinkhof, K. (2006). Switching between tasks and responses: A developmental study. *Developmental Science*, *9*(3), 278–287. https://doi.org/10.1111/j.1467-7687.2006.00490.x

Da Silva, C. F., & Hare, T. A. (2018). A note on the analysis of two-stage task results: How changes in task structure affect what model-free and model-based strategies predict about the effects of reward and transition on the stay probability. *PLoS ONE*, *13*(4), 1–13. https://doi.org/10.1371/journal.pone.0195328

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory and Cognition*, *35*(1), 15–28. https://doi.org/10.3758/BF03195938

Davidson, M., Amso, D., Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037–2078.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *In Affect, Learning and Decision Making, Attention and Performance XXIII, E.A. Phelps, T.W. Robbins and M. Delgado, eds. Oxford University Press.*

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*, 199–204.

Daw, N. D., Gershman, J., Seymour, B., Dayan, P., & Dolan, R. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control - Nature Neuroscience. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Curr Opin Neurobiol*, *18*(2), 185–196.

Decker, J., Otto, A., Daw, N., & Hartley, C. (2016). From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning. *Psychological Science*, *27*(6), 848–858. https://doi.org/10.1177/0956797616639301

Decker, J., Lourenco, F., Doll, B., & Hartley, C. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cogn. Affect. Behav. Neurosci*, *15*, 2, 310–320.

Dehaene, S., & Changeux, J.-P. (1997). A hierarchical neuronal network for planning behavior, In *Proceedings of the national academy of sciences of the united states of america*.

Dekker, T., Mareschal, D., Sereno, M., & Johnson, M. (2010). Dorsal and ventral stream activation and object recognition performance in school-age children. *NeuroImage*, *57*, 659–670. https://doi.org/10.1016/j.neuroimage.2010.11.005

Della Sala, S., Marchetti, C., & Spinnler, H. (1991). Right-sided anarchic (alien) hand: A longitudinal study. *Neuropsycholoquia*, *29*(11), 1113–1127.

Desmurget, M., Jordan, M., Prablanc, C., & Jeannerod, M. (1997). Constrained and Unconstrained Movements Involve Different Control Strategies. *Journal of Neurophysiology*. https://doi.org/10.1152/jn.1997.77.3.1644

Desmurget, M., & Turner, R. S. (2010). Motor Sequences and the Basal Ganglia: Kinematics, Not Habits. *Journal of Neuroscience*, *30*(22), 7685–7690. https://doi.org/10.1523/JNEUROSCI.0163-10.2010

Dezfouli, A., & Balleine, B. W. (2013). Actions, Action Sequences and Habits: Evidence That Goal-Directed and Habitual Action Control Are Hierarchically Organized. *PLoS Computational Biology*, *9*(12). https://doi.org/10.1371/journal.pcbi.1003364

Dezfouli, A., Lingawi, N. W., & Balleine, B. W. (2014). Habits as action sequences: hierarchical action control and changes in outcome value. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2013.0482

Diamond, A. (1990a). Developmental time course in human infants and infant monkeys, and the neural bases of inhibitory control in reaching.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. Stuss & R. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). New York, Oxford University Press.

Diamond, A., Carlson, S., & Beck, D. (2005). Preschool childrens performance in task switching on the dimensional change card sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology*, *28*, 689–729.

Diamond, A. (1990b). Developmental Time Course in Human Infants and Infant Monkeys, and the Neural Bases of, Inhibitory Control in Reachinga. *Annals of the New York Academy of Sciences*, *608*(1), 637–676. https://doi.org/10.1111/j.1749-6632.1990.tb48913.x

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diedrichsen, J., & Kornysheva, K. (2015). Motor skill learning between selection and execution. *Trends in Cognitive Sciences*, *19*(4), 227–233. https://doi.org/10.1016/j.tics.2015.02.003

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Dotan, D., Meyniel, F., & Dehaene, S. (2018). On-line con fi dence monitoring during decision making. *Cognition*, *171*, 112–121. https://doi.org/10.1016/j.cognition.2017.11.001

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw*, *12*, 961–974.

Dunbar, K., & Sussman, D. (1995). Toward a Cognitive Account of Frontal Lobe Function: Simulating Frontal Lobe Deficits in Normal Subjects. *Annals of the New York Academy of Sciences*, *769*(1), 289–304. https://doi.org/10.1111/j.1749-6632.1995.tb38146.x

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., & Cox, S. (2008). Goal neglect and spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology: General*, *137*(1), 131–148.

Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, *3*(3), 271–290.

Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). *Manual for kit of factorreferenced cognitive tests*. Princeton, NJ, Educational Testing Service.

Ellis, R., Tucker, M., Symes, E., & Vainio, L. (2007). Does selecting one visual object from several require inhibition of the actions associated with non-selected objects? *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 670–691. https://doi.org/10.1037/0096-1523.33.3.670.

Elman, J. L. (1990). Sequential machine. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, *14*, 531–546.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, *7*, 195–225.

Eppinger, B., Heekeren, H., & Li, S.-C. (2012). When two birds in the bush are better than one in the hand: Age-related impairments in learning to predict future rewards, In *Annual meeting of the society for neuroscience*, New Orleans, LA). Orleans, LA.

Eppinger, B., Walter, M., Heekeren, H. R., & Li, S.-C. (2013). Of goals and habits: age-related and individual differences in goal-directed decision-making. *Frontiers in Neuroscience*, *0*. https://doi.org/10.3389/fnins.2013.00253

Erb, C., Moher, J., Sobel, D., & Song, J. (2016). Reach tracking reveals dissociable processes underlying cognitive control. *Cognition*, *152*, 114–126.

Erb, C., Moher, J., Song, J., & Sobel, D. (2018). Reach tracking reveals dissociable processes underlying inhibitory control in 5- to 10-year-olds and adults. *Developmental Science*, *21*(2), 1–14. https://doi.org/10.1111/desc.12523.

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, *16*, 143–149.

Fan, J., McCandliss, B., Sommer, T., Raz, M., & Posner, M. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*, 340–347.

Fein, G. (1981). Pretend play in childhood: An integrative review. *Child Development*, *52*, 1095.

Fenson, L., & Ramsay, D. (1981). Effects of modeling action sequences on the play of twelve-, fifteen-, and nineteen-month old children. *Child Development*, *52*, 1028–1036.

Fischer, K. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, *87*(6), 477–531. https://doi.org/10.1037/0033-295X.87.6.477

Fontanesi, L., Palminteri, S., & Lebreton, M. (2019). Decomposing the effects of context valence and feedback information on speed and accuracy during reinforcement learning: a meta-analytical approach using diffusion decision modeling. *Cognitive, Affective and Behavioral Neuroscience*, *19*(3), 490–502. https://doi.org/10.3758/s13415-019-00723-1

Frank, M., & Kong, L. (2008). Learning to avoid in older age. *Psychol. Aging*, *23*, 392–398. https://doi.org/10.1037/0882-7974.23.2.392

Freeman, J. B. (2018). Doing Psychological Science by Hand. https://doi.org/10.1177/0963721417746793

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241. https://doi.org/10.3758/BRM.42.1.226

Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, *2*(APR), 1–6. https://doi.org/10.3389/fpsyg.2011.00059

Freier, L. (2016). *The development of bottom-up and top-down interaction in the processing of goal-directed action* (Doctoral dissertation). https://eprints.bbk.ac.uk/id/eprint/40176

Freier, L., Cooper, R. P., & Mareschal, D. (2015). The planning and execution of natural sequential actions in the preschool years. *Cognition*, *144*, 58–66. https://doi.org/10.1016/j.cognition.2015.07.005

Freier, L., Cooper, R. P., & Mareschal, D. (2017). Preschool children's control of action outcomes. *Developmental Science*, *20*(2), 1–13. https://doi.org/10.1111/desc.12354

Friedman, N., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.

Fuster, J. (1990). Prefrontal cortex and the bridging of temporal gaps in the perceptionaction cycle [doi:10.1111/j.1749-6632.1990. tb48901.x]. *Ann. NY Acad. Sci*, *608*, 318–329.

Gallivan, J. P., Cavina-Pratesi, C., & Culham, J. C. (2009). Is That within Reach? fMRI Reveals That the Human Superior Parieto-Occipital Cortex Encodes Objects Reachable by the Hand. *Journal of Neuroscience*, *29*(14), 4381–4391. https://doi.org/10.1523/JNEUROSCI.0377-09.2009

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, *3*(11), 410–419. https://doi.org/10.1016/S1364-6613(99)01388-1

Gazzaniga, M., Bogen, J., & Sperry, R. (1962). Some functional effects of sectioning the cerebral commisures in man., In *Proceedings of the national academy of sciences, 48*.

Gershman, S. J. (2018). Deconstructing the human algorithms for exploration. *Cognition*, *173*, 34–42. https://doi.org/10.1016/j.cognition.2017.12.014

Gerstadt, C., Hong, Y., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3.57 years old on a stroop-like daynight test. *Cognition*, *53*, 129–153.

Gibson, J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67–82). Hillsdale, NJ, Lawrence Erlbaum.

Gibson, J. (1979). *The ecological approach to visual perception*. MIT.

Giedd, J., Blumenthal, J., Jeffries, N., Castellanos, F., Liu, H., Zijdenbos, A., Paus, T., Evans, A., & Rapoport, J. (1999). Brain development during childhood and adolescence: A longitudinal mri study. *Nature Neuroscience*, *2*(10), 861–863.

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology*, *44*(3), 297–337. https://doi.org/10.1006/cogp.2001.0770

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *ELife*, *5*26928075, e11305. https://doi.org/10.7554/eLife.11305

Giovannetti, T., Schwartz, M. F., & Buxbaum, L. J. (2007). The Coffee Challenge: A new method for the study of everyday action errors. *Journal of Clinical and Experimental Neuropsychology*, *29*(7), 690–705. https://doi.org/10.1080/13803390600932286

Gnanaseelan, R., Gonzalez, D. A., & Niechwiej-Szwedo, E. (2014). Binocular advantage for prehension movements performed in visually enriched environments requiring visual search. *Frontiers in Human Neuroscience*, *8*(NOV). https://doi.org/10.3389/fnhum.2014.00959

Goldberg, G., Mayer, N., & Toglia, J. (1981). Medial frontal cortex infarction and the alien hand sign. *Archives of Neurology*, *38*, 683–686.

Goldman-Rakic, P. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology: The nervous system* (pp. 373–417). Bethesda, MD, American Physiological Society.

Gottwald, J. M., De Bortoli Vizioli, A., Lindskog, M., Nyström, P., Ekberg, T. L., von Hofsten, C., & Gredebäck, G. (2017). Infants prospectively control reaching based on the difficulty of future actions: To what extent can infants' multiple-step actions be explained by Fitts' law? *Developmental Psychology*, *53*(1), 4–12. https://doi.org/10.1037/dev0000212

Graybiel, A. M. (2008). Habits, Rituals, and the Evaluative Brain. *Annual Review of Neuroscience*, *31*(1), 359–387. https://doi.org/10.1146/annurev.neuro.29.051605.112851

Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation - Nature Human Behaviour. *Nature Human Behaviour*, *3*(12), 1252–1262. https://doi.org/10.1038/s41562-019-0725-0

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., Imamizu, H., & Kawato, M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: A functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci*, *24*, 1660–1665.

Heathcote, A., Popiel, S., & Mewhort, D. (1991). Analysis of response time distributions: An example using the stroop task. *Psychological Bulletin*, *109*(2), 340–347. https://doi.org/10.1037/0033-2909.109.2.340

Heckhausen, H., & Beckmann, J. (1990). Intentional action and action slips. *Psychological Review*, *97*(1), 36–48. https://doi.org/10.1037/0033-295X.97.1.36

Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes and Intergroup Relations*, *18*(3), 384–401. https://doi.org/10.1177/1368430214538325

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinion in Neurobiology*, *12*(2), 217–222.

Hofer, T., Hauf, P., & Aschersleben, G. (2005). Infants perception of goal-directed actions performed by a mechanical device. *Infant Behavior and Development*, *28*(4), 466–480. https://doi.org/10.1016/j.infbeh.2005.04

Hofer, T., Hauf, P., & Aschersleben, G. (2007). Infants' perception of goal-directed actions on video. *British Journal of Developmental Psychology*, *25*(3), 485–498. https://doi.org/10.1348/026151006X170308

Hofsten, C. (1993). Prospective control: A basic aspect of action development. *Human*, *Development*,*36*, 253–270.

Hofsten, C., & Ronnqvist, L. (1988). Preparation for grasping an object: A developmental study. *Journal of Experimental Psychology: Human Perception and*, *Performance*,*14*, 610–621.

Hommel, B., & Elsner, B. (2009). Acquisition, representation, and control of action. In E. Morsella, J. Bargh, & P. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 371–398). New York, Oxford University Press.

Houghton, G. (1990). The problem of serial order: Opponent mechanisms in sequencing and selective attention. *Current research in natural language generation*, 287–319.

Howard, L. A., & Tipper, S. P. (1997). Hand deviations away from visual cues: Indirect evidence for inhibition experimental brain research.

Hudson, J., & Fivush, R. (1991). Planning in the preschool years: The emergence of plans from general event knowledge. *Cognitive Development*, *6*, 393–415. https://doi.org/10.1016/0885-2014(91)90046-G

Hudson, J., Shapiro, L., & Sosa, B. (1995). Planning in the real world: Preschool childrens scripts and plans for familiar events. *Child Development*, *66*, 984–998. https://doi.org/10.2307/1131793

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009). Tracking Executive Function Across the Transition to School: A Latent Variable Approach. *Developmental Neuropsychology*, *35*(1), 20–36. https://doi.org/10.1080/87565640903325691

Huizinga, M., & van der Molen, M. W. (2011). Task switching and shifting between stopping and going: Developmental change in between-trial control adjustments. *Journal of Experimental Child Psychology*, *108*(3), 484–503. https://doi.org/10.1016/j.jecp.2010.10.004

Humes, G., Welsh, M., Retzlaff, P., & Cookson, N. (1997). Towers of hanoi and london: Reliability and validity of two executive function tests. *Assessment*, *4*, 249–257.

Humphreys, G. W., & Forde, E. M. (1998). Disordered action schemas and action disorganisation syndrome. *Cognitive Neuropsychology*, *15*, 771–811.

Humphreys, G. W., Forde, E. M., & Francis, D. (2000). The organization of sequential actions. *Attention and Performance*, *18*(1995), 424–442.

Ikeda, Y., Okuzumi, H., & Kokubun, M. (2014). Age-related trends of inhibitory control in Stroop-like big small task in 3 to 12-year-old children and young adults, *5*, 1–6. https://doi.org/10.3389/fpsyg.2014.00227

Jackson, S. R., Jackson, G. M., & Rosicky, J. (1995). Are non-relevant objects represented in working memory? The effect of non-target objects on reach and grasp kinematics. *Experimental Brain Research*, *102*(3), 519–530. https://doi.org/10.1007/BF00230656

James, A., Jenks, C., & Prout, A. (1998). *Theorizing childhood*. New York, NY, Teachers College Press.

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, *15*(4), 496–505. https://doi.org/10.1111/j.1467-7687.2012.01150.x

Jaspers, E., Feys, H., Bruyninckx, H., Harlarr, J., Molenaers, G., & Desloovere, K. (2011). Upper limb kinematics: Development and reliability of a clinical protocol for children. *Gait & Posture*, *33*, 279–285.

Javadi, A., Schmidt, D., & Smolka, M. (2014). Adolescents adapt more slowly than adults to varying reward contingencies. *J. Cogn. Neurosci*, *26*(12), 2670–2681.

Jenkins, I., Brooks, D., Nixon, P., Frackowiak, R., & Passingham, R. (1994). Motor sequence learning: A study with positron emission tomography. *Journal of Neuroscience*, *14*, 3775–3790.

Jennings, K. D. (2004). Development of goal-directed behaviour and related self-processes in toddlers. *International Journal of Behavioral Development*, *28*(4), 319–327. https://doi.org/10.1080/01650250444000036

Jensen, J. L. (2005). The puzzles of motor development: How the study of developmental biomechanics contributes to the puzzle solutions, *14*(5), 501–511. https://doi.org/10.1002/icd.425

Jersild, A. (1927). Mental set and shift. *Archives of Psychology, Whole*, (89).

Jervis, C., Bennett, K., Thomas, J., Lim, S., & Castiello, U. (1999). Semantic category interference effects upon the reach to grasp movement. *Neuropsychologia*, *37*, 857 868.

Jones, L., Rothbart, M., & Posner, M. (2003). Development of executive attention in preschool children. *Developmental Science*, *6*(5), 498–504.

Jordan, M. (1986). Attractor dynamics and parallelism in a connectionist.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S., & Passingham, R. E. (1997). Anatomy of motor learning. i. frontal cortex and attention to action. *Journal of Neurophysiology*, *77*(3), 1313–1324. https://doi.org/10.1152/jn.1997.77.3.1313

Kagan, J. (1981). *The second year: The emergence of self-awareness*. Cambridge,MA, Harvard University Press.

Kahrs, B., Jung, W. P., & Lockman, J. J. (2013). Motor Origins of Tool Use. *Child development*, *84*(3). https://doi.org/10.1111/cdev.12000

Kahrs, B., Jung, W., & Lockman, J. (2012). What is the role of infant banging in the development of tool use? *Experimental Brain Research*, *218*, 315–320.

Kahrs, B., Jung, W., & Lockman, J. (2014). When does tool use become distinctively human? hammering in young children. *Child Development*, *85*, 1050–1061.

Kaller, C. P., Rahm, B., Spreer, J., Mader, I., & Unterrainer, J. M. (2008). Thinking around the corner: The development of planning abilities. *Brain and Cognition*, *67*, 360–370.

Kaller, C. P., Rahm, B., Spreer, J., & Mader, I. (2008). Thinking around the corner : The development of planning abilities, *67*, 360–370. https://doi.org/10.1016/j.bandc.2008.02.003

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, *23*(2), 95–147. https://doi.org/10.1016/0010-0277(86)90040-5

Karni, A. (1995). Functional mri evidence for adult motor cortex plasticity during motor skill learning. *Nature*, *377*, 155–158.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-barrera, M. A. (2018). The Unity and Diversity of Executive Functions : A Systematic Review and Re-Analysis of Latent Variable Studies, *144*(11), 1147–1185.

Keele, S., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, *110*(2), 316–339.

Kelly, A., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex*, *15*, 1089–1102.

Kenward, B., Folke, S., Holmberg, J., Johansson, A., & Gredebäck, G. (2009). Goal Directedness and Decision Making in Infants. *Developmental Psychology*, *45*(3), 809–819. https://doi.org/10.1037/a0014076

Keramati, M., Dezfouli, A., & Piray, P. (2011a). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology*, *7*(5), e1002055. https://doi.org/10.1371/journal.pcbi.1002055

Keramati, M., Dezfouli, A., & Piray, P. (2011b). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS computational biology*, *7*(5), e1002055. https://doi.org/10.1371/journal.pcbi.1002055

Keulen, R. F., Adam, J. J., Fischer, M. H., Kuipers, H., & Jolles, J. (2004). Selective Reaching: Distracter Effects on Movement Kinematics as a Function of TargetDistracter Separation. *Journal of General Psychology*, *131*(4), 345–364. https://doi.org/10.3200/GENP.131.4.345-364

Kieslich, P. J., & Henninger, F. (2017). Mousetrap: An integrated, open-source mouse-tracking package. *Behavior Research Methods*, *49*(5), 1652–1667. https://doi.org/10.3758/s13428-017-0900-z

Kieslich, P. J., Schoemann, M., Grage, T., Hepp, J., & Scherbaum, S. (2020). Design factors in mouse-tracking: What makes a difference? *Behavior Research Methods*, *52*(1), 317–341. https://doi.org/10.3758/s13428-019-01228-y

Kimberg, D., DEsposito, M., & Farah, M. (1997). Effects of bromocriptine on human subjects depend on working memory capacity. *NeuroReport*, *6*, 3581–3585.

Klauer, K. C., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2010). Understanding the role of executive control in the Implicit Association Test: Why flexible people have small IAT effects. *Quarterly Journal of Experimental Psychology*, *63*(3), 595–619. https://doi.org/10.1080/17470210903076826

Klenberg, L., Korkman, M., & Lahti-Nuuttila. (2010). Differential development of attention and executive functions in 3- to 12-year-old finish children. *Developmental Neuropsychology*, *20*, 407–428.

Klossek, U. M. H., & Dickinson, A. (2012). Journal of Experimental Child Rational action selection in 1¡- to 3-year-olds following an extended training experience. *Journal of Experimental Child Psychology*, *111*(2), 197–211. https://doi.org/10.1016/j.jecp.2011.08.008

Kochari, A. R. (2019). Conducting Web-Based Experiments for Numerical Cognition Research. *Journal of Cognition*, *2*(1), 1–21. https://doi.org/10.5334/joc.85

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture ofcognitive control in the human prefrontal cortex. *Science*, *302*, 1181–1185.

Konczak, J., & Dichgans, J. (1997). The development toward stereotypic arm kinematics during reaching in the first 3 years of life. *Experimental Brain Research*, *117*(2), 346–354. https://doi.org/10.1007/s002210050228

Konovalov, A., & Krajbich, I. (2016). Gaze data reveal distinct choice processes underlying model-based and model-free reinforcement learning. *Nature Communications*, *7*, 1–11. https://doi.org/10.1038/ncomms12438

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When Does Model-Based Control Pay Off? *PLoS Computational Biology*, *12*(8), 1–34. https://doi.org/10.1371/journal.pcbi.1005090

Kool, W., Gershman, S. J., & Cushman, F. A. (2017a). Cost-Benefit Arbitration Between Multiple Reinforcement-Learning Systems. *Psychological Science*, *28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Kool, W., Gershman, S. J., & Cushman, F. A. (2017b). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, *28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Kopp, C. (1991). Young childrens progression to self-regulation (MBullock, Ed.). *Contributions to human development*, *22*, 38–54.

Kuhtz Buschbeck, J., Stolze, H., Johnk, K., Boczek, A., & Illert, M. (1998). Development of prehension movements in children: A kinematic study. *Experimental Brain Research*, *122*, 424 432.

Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, *40*(6), 998–1009. https://doi.org/10.1002/ejsp.674

Laming, D. (1979). Choice reaction performance following an error. *Acta Psychologica*, *43*(3), 199–224. https://doi.org/10.1016/0001-6918(79)90026-X

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, *25*(3), 296–324. https://doi.org/10.1016/j.preteyeres.2006.01.002

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25-26), 11718795, 3559–3565. https://doi.org/10.1016/s0042-6989(01)00102-x

Land, M. F., Mennie, N., & Rusted, J. (1999). The Roles of Vision and Eye Movements in the Control of Activities of Daily Living. *Perception*, *28*(11), 1311–1328. https://doi.org/10.1068/p2935

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

Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models. *Journal of Mathematical Psychology*, *55*(1), 1–7. https://doi.org/10.1016/j.jmp.2010.08.013

Lee, S., Shimojo, S., & ODoherty, J. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron*, *81*, 687–699.

Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. https://doi.org/10.1348/026151003321164627

Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents : Insights from anatomical magnetic resonance imaging, *30*, 718–729. https://doi.org/10.1016/j.neubiorev.2006.06.001

Leon-Carrion, J., Garcia-Orza, J., & Perez-Santamaria, F. (2004). Development of the inhibitory component of the executive functions in children and adolescents. *International Journal of Neuroscience*, *114*(10), 1291–1311. https://doi.org/10.1080/00207450490476066

Lewis, F. C., Reeve, R. A., Kelly, S. P., & Johnson, K. A. (2017). Journal of Experimental Child Evidence of substantial development of inhibitory control and sustained attention between 6 and 8 years of age on an unpredictable Go / No-Go task. *Journal of Experimental Child Psychology*, *157*, 66–80. https://doi.org/10.1016/j.jecp.2016.12.008

Lhermitte, F. (1983). Utilisation behaviour and its relation to lesions of the frontal lobes. *Brain*, *106*, 237–255.

Lockman, J., Ashmead, D., & Bushnell, E. (1984). The development of anticipatory hand orientation during infancy. *Journal of Experimental Child Psychology*, 37, 176–186.

Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA, Academic Press.

Logan, G. D., & Crump, M. J. C. (2011). Hierarchical Control of Cognitive Processes: The Case for Skilled Typewriting. In *Psychology of Learning and Motivation* (pp. 1–27). Cambridge, MA, USA, Academic Press. https://doi.org/10.1016/B978-0-12-385527-5.00001-2

Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E. J. (2014). On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review*, *121*(1), 66–95. https://doi.org/10.1037/a0035230

Luna, B. (2009). Developmental changes in cognitive control through adolescence. *Advances in Child Development and Behavior*, *37*, 233–278.

Luria, A. (1966). *Higher cortical functions in man*. New York, Basic Books.

Lutkenhaus, P., Bullock, M., & Geppert, U. (1987). Toddlers' Actions: Knowledge, Control, and the Self. In *Motivation, Intention, and Volition* (pp. 145–161). Berlin, Germany, Springer. https://doi.org/10.1007/978-3-642-70967-8_11

Macdonald, J. A., Beauchamp, M. H., Crigan, J. A., & Anderson, P. J. (2014). Age-related differences in inhibitory control in the early school years. *Child Neuropsychology*, *20*(5), 509–526. https://doi.org/10.1080/09297049.2013.822060

Magimairaj, B. M. (2017). Psycholinguistic mechanisms supporting school-age children's comprehension of spoken directions: A preliminary study. *Speech, Language and Hearing*, *20*(1), 38–48. https://doi.org/10.1080/2050571X.2016.1206167

Maldonado, M., Dunbar, E., & Chemla, E. (2019). Mouse tracking as a window into decision making, (2011).

Maldonado, T., Orr, J. M., Goen, J. R. M., & Bernard, J. A. (2020). Age Differences in the Subcomponents of Executive Functioning, *75*(6). https://doi.org/10.1093/geronb/gbaa005

Marcovitch, S., Boseovski, J., & Knapp, R. (2007). Use it or lose it: Examining preschoolers difficulty in maintaining and executing goals. *Developmental Science*, *10*, 559–564.

Marcovitch, S., Boseovski, J. J., & Knapp, R. J. (2007). Use it or lose it : examining preschoolers ' difficulty in maintaining and executing a goal, *5*, 559–564. https://doi.org/10.1111/j.1467-7687.2007.00611.x

Mareschal, D., & Johnson, M. H. (2003). The what and where of object representations in infancy. *Cognition*, *88*(3), 259–276. https://doi.org/10.1016/S0010-0277(03)00039-8

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Henry Holt; Co.

Master, S. L., Eckstein, M. K., Gotlieb, N., Dahl, R., Wilbrecht, L., Anne, G., & Collins, E. (2020). Developmental Cognitive Neuroscience Distentangling the systems contributing to changes in learning during adolescence. *Developmental Cognitive Neuroscience*, *41*, 100732. https://doi.org/10.1016/j.dcn.2019.100732

Master, S. L., Eckstein, M. K., Gotlieb, N., Dahl, R., Wilbrecht, L., & Collins, A. G. E. (2020). Disentangling the systems contributing to changes in learning during adolescence. *Developmental Cognitive Neuroscience*, *41*, 100732. https://doi.org/10.1016/j.dcn.2019.100732

Mayr, U., & Keele, S. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General*, *129*, 4–26.

McCarty, M., Clifton, R., Ashmead, D., Lee, P., & Goubet, N. (2001). How infants use vision for grasping objects. *Child Development*, *72*, 973–987.

McClelland, J. (1992). Toward a theory of information processing in graded, random, interactive networks. In D. Meyer & S. Kornblum (Eds.), *Attention and performance xiv: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 655–688). Cambridge, MA, MIT Press.

McClelland, J., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. an account of basic findings. *Psychological Review*, *88*(5), 375–407.

Meegan, D. V., & Tipper, S. P. (1999a). Visual search and target-directed action. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1347–1362. https://doi.org/10.1037/0096-1523.25.5.1347

Meegan, D. V., & Tipper, S. P. (1999b). Visual search and target-directed action. *Journal of Experimental Psychology: Human Perception & Performance*, *25*(5), 1347–1362.

Mehnert, J., Akhrif, A., Telkemeyer, S., Rossi, S., Schmitz, C., Steinbrink, J., & Neufang, S. (2013). Developmental changes in brain activation and functional connectivity during response inhibition in the early childhood brain.

Mehnert, J., Akhrif, A., Telkemeyer, S., Rossi, S., Schmitz, C. H., Steinbrink, J., Wartenburger, I., Obrig, H., & Neufang, S. (2013). Developmental changes in brain activation and functional connectivity during response inhibition in the early childhood brain. *Brain and Development*, *35*(10), 894–904. https://doi.org/10.1016/j.braindev.2012.11.006

Michalczyk, K., Malstadt, N., Worgt, M., Konen, T., & Hasselhorn, M. (2013). Age Differences and Measurement Invariance of Working Memory in 5- to 12-Year-Old Children. *European Journal of Psychological Assessment*.

Miles, J., & Proctor, R. (2011). Colour correspondence effects between controlled objects and targets. *The Quarterly Journal of Experimental Psychology*, *64*, 2044–2064. https://doi.org/10.1080/17470218.2011.582130

Mikowski, M., Hensel, W. M., & Hohol, M. (2018). Replicability or reproducibility? On the replication crisis in computational neuroscience and sharing only relevant detail. *Journal of Computational Neuroscience*, *45*(3), 163–172. https://doi.org/10.1007/s10827-018-0702-z

Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review*, *126*(2), 292–311. https://doi.org/10.1037/rev0000120

Mitsopoulos, C., Mareschal, D., & Cooper, R. P. (2015). Model-based analysis of the tower of london task. *In proceedings of :Reinforcement Learning and Decision Making, University of Alberta*, 198–202.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000a). The Unity and Diversity of Executive Functions and Their Contributions to Complex Frontal Lobe " Tasks : A Latent Variable Analysis and, *100*, 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000b). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Mohler, J., Anderson, B., & Song, J. (2015). Dissociable effects of salience on attention and goal-directed action. *Current Biology*, *25*, 2040–2046.

Montague, P., Dolan, R., Friston, K., & Dayan, P. (2012). Computational psychiatry. *Trends Cogn Sci*, *16*, 72–80. https://doi.org/10.1016/j.tics.2011.11.018

Morey, R., & Rouder, N. (2018). Bayesfactor: Computation of bayes factors for common designs [Retrieved from]. https://CRAN.R-project.org/package=BayesFactor.

Morris, A., & Cushman, F. (2019). Model-Free RL or Action Sequences? *Frontiers in Psychology*, *0*. https://doi.org/10.3389/fpsyg.2019.02892

Morris, N., & Jones, D. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*, 111–121.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.

Nevo, E., & Breznitz, Z. (2013). The development of working memory from kindergarten to first grade in children with different decoding skills. *Journal of Experimental Child Psychology*, *114*(2), 217–228. https://doi.org/10.1016/j.jecp.2012.09.004

Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, *88*(1), 1–15. https://doi.org/10.1037/0033-295X.88.1.1

Norman, D. A., & Shallice, T. (1986). Attention to Action. In *Consciousness and Self-Regulation: Advances in Research and Theory Volume 4* (pp. 1–18). Boston, MA, USA, Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-0629-1_1

Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research? *Developmental*

Cognitive Neuroscience, *40*, 100733. https://doi.org/10.1016/j.dcn.2019.100733

Ottes, F., Van Gisbergen, J., & Eggermont, J. (1987). Collicular involvement in a saccadic colour discrimination task. *Experimental Brain Research*, *66*, 465–478.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, *110*(52), arXiv NIHMS150003, 20941–20946. https://doi.org/10.1073/pnas.1312011110

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2014). Cognitive Control Predicts Use of Model-based Reinforcement Learning. *Journal of Cognitive Neuroscience*, *27*(2), 319–333. https://doi.org/10.1162/jocn_a_00709

Owen, A. (1997). Cognitive planning in humans: Neuropsychological, neuroanatomical, and neuropharmacological perspectives. *Prog Neurobiol*, *53*, 431–450.

Palminteri, S., Lefebvre, G., Kilford, E. J., & Blakemore, S. (2017). Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. *PLOS Computational Biology*, *13*(8), e1005684. https://doi.org/10.1371/journal.pcbi.1005684

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The Importance of Falsification in Computational Cognitive Modeling. *Trends in Cognitive Sciences*, *21*(6), 425–433. https://doi.org/10.1016/j.tics.2017.03.011

Pascual-Leone, A., Wassermann, E. M., Grafman, J., & Hallett, M. (1996). The role of the dorsolateral prefrontal cortex in implicit procedural learning. *Experimental Brain Research*, *107*(3), 479–485. https://doi.org/10.1007/BF00230427

Paulus, M., & Sodian, B. (2015). Which way to take? Infants select an efficient path to their goal. *Journal of Experimental Child Psychology*, *137*, 111–124. https://doi.org/10.1016/j.jecp.2015.03.014

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin and Review*, *24*(4), 1234–1251. https://doi.org/10.3758/s13423-016-1199-y

Pelz, J., & Canoza, R. (2001). Oculomotor behavior and perceptual categories in complex tasks. *Vision Res*, *41*, 3587–3596.

Perez, O. D., Dickinson, A., Perez, O. D., & Dickinson, A. (2020). Psychological Review A Theory of Actions and Habits : The Interaction of Rate Correlation and Contiguity Systems in Free-Operant A Theory of Actions and Habits : The Interaction of Rate Correlation and Contiguity Systems in Free-Operant Behavior.

Perone, S., Anderson, A. J., & Youatt, E. A. (2020). Don't forget your lunch: Age and individual differences in how children perform everyday tasks. *Cognitive Development*, *54*. https://doi.org/10.1016/j.cogdev.2020.100879

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behavior. *Neuropsychologia*, *12*, 323–330.

Petitclerc, A., Briggs-Gowan, M. J., Estabrook, R., Burns, J. L., Anderson, E. L., McCarthy, K. J., & Wakschlag, L. S. (2015). Contextual variation in young children's observed disruptive behavior on the DB-DOS: implications for early identification. *Journal of Child Psychology and Psychiatry*, *56*(9), 1008–1016. https://doi.org/doi:10.1111/jcpp.12430

Phillips, L. (1997). Do 'frontal tests' measure executive function? issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 191–213). Hove, UK, Psychology Press.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2020). nlme: Linear and Nonlinear Mixed Effects Models. *R package version 3.1-148.*, *Retrieved from https://CRAN.R-project.org/package=nlme.*

Plumert, J., Kearney, J., & Cremer, J. (2007). Childrens road crossing: A window into perceptual-motor development. *Current Directions in Psychological Science,16*, *255e258*.

Potter, T. C., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Developmental Cognitive Neuroscience*, *25*, 272–280. https://doi.org/10.1016/j.dcn.2016.10.005

Pouget, P., Murthy, A., Stuphorn, V., Curie-paris, M., & Pouget, P. (2017). Cortical control and performance monitoring of interrupting and redirecting movements.

Pratt, J., & Abrams, R. (1994). Action-centred inhibition: Effects of distractors on movement planning and execution. *Human Movement Science*, *12*(2), 245–254.

Pratt, J., & Abrams, R. A. (1994). Action-centered inhibition: Effects of distractors on movement planning and execution. *Human Movement Science*, *13*(2), 245–254. https://doi.org/10.1016/0167-9457(94)90039-6

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Reason, J. (1979). Actions not as planned: The price of automatization. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness (ch. 4* (pp. 67–89). London, Academic Press.

Reason, J. (1984). Lapses of attention in everyday life. In W. Parasuraman & R. Davies (Eds.), *Varieties of attention (ch. 14* (pp. 515–549). Orlando, FL, Academic Press.

Reason, J. (1990). *Human error*. Cambridge, Cambridge University Press.

Reason, J., & Mycielska, K. (1982). *Absent minded? the psychology of mental lapses and everyday errors*. Englewood Cliffs, NJ, Prentice Hall.

Reck, S. G., & Hund, A. M. (2011). Journal of Experimental Child Sustained attention and age predict inhibitory control during early childhood. *Journal of Experimental Child Psychology*, *108*(3), 504–512. https://doi.org/10.1016/j.jecp.2010.07.010

Rhodes, B. J., Bullock, D., Verwey, W. B., Averbeck, B. B., & Page, M. P. (2004). Learning and production of movement sequences: Behavioral, neurophysiological, and modeling perspectives. *Human Movement Science*, *23*(5), 699–746. https://doi.org/10.1016/j.humov.2004.10.008

Rizzolatti, G., Riggio, L., Dascola, I., & Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1, Part 1), 31–40. https://doi.org/10.1016/0028-3932(87)90041-8

Roberts, R., Hager, L., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*, 374–393.

Rogers, R. D., & Monsell, S. (1995a). Costs of a Predictable Switch Between Simple Cognitive Tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Rogers, R. D., & Monsell, S. (1995b). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*, 207–231.

Roy, A., Kefi, M.-Z., Bellaj, T., Fournet, N., Le Gall, D., & Roulin, J. (2018). The Stroop test: A developmental study in a French children sample aged 7 to 12 years. *Psychologie Francaise*, *63*(2), 129–143. https://doi.org/10.1016/j.psfr.2016.08.001

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2003.12.012

Ruh, N., Cooper, R. P., & Mareschal, D. (2008). The Hierarchies and Systems that Underlie Routine Behavior : Evidence from an Experiment in Virtual Gardening. *Proceedings of the Annual Meeting of the Cognitive Science*.

Ruh, N., Cooper, R. P., & Mareschal, D. (2010). Action selection in complex routinized sequential behaviors. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 955–975. https://doi.org/10.1037/a0017608

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of cognitive-motor performance. *Cognitive Science*, *6*, 1–36.

Schacter, D., Addis, D., Hassabis, D., Martin, V., Spreng, R., & Szpunar, K. (2012). The future of memory: Remembering, imagining, and the brain. *Neuron*, *76*, 677–694.

Schank, R. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge, Cambridge University Press.

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hove, UK, Lawrence Erlbaum Associates Ltd.

Schiller, P. (1952). Innate constituents of complex responses in primates. *Psychological Review*, *59*(3), 177–191. https://doi.org/10.1037/h0062854

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*(4). https://doi.org/10.1037/h0076770

Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. *Human Factors*, *30*, 539–566.

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition*, *2*(2-3), 331–376. https://doi.org/10.1080/13506289508401737

Schoemann, M., Lüken, M., Grage, T., Kieslich, P. J., & Scherbaum, S. (2019). Validating mouse-tracking : How design factors influence action dynamics in intertemporal decision making.

Schröer, L., Cooper, R. P., & Mareschal, D. (2021). Science with Duplo: Multilevel goal management in preschoolers' toy house constructions. *Journal of Experimental Child Psychology*, *206*. https://doi.org/10.1016/j.jecp.2020.105067

Schultz, W., Dayan, P., & Montague, P. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for Rewards Like a Child Means Less Generalization and More Directed Exploration. *Psychological Science*, *30*(11), 1561–1572. https://doi.org/10.1177/0956797619863663

Schwartz, M. F. (1995). Re-examining the Role of Executive Functions in Routine Action Production". *Annals New York Academy of sciencess*.

Schwartz, M. F., Buxbaum, L. J., Montgomery, M. W., Fitzpatrick-DeSalme, E., Hart, T., Ferraro, M., Lee, S. L., & Coslett, H. B. (1999). Naturalistic action production following right hemisphere stroke. *Neuropsychologia*, *37*, 51–66.

Schwartz, M. F., Buxbaum, L. J., Montgomery, M. W., Fitzpatrick-Desalme, E., Hart, T., Ferraro, M., Lee, S. S., & Coslett, H. B. (1998). Naturalistic action production following right hemisphere stroke. *Neuropsychologia*, *37*(1), 51–66. https://doi.org/10.1016/S0028-3932(98)00066-9

Schwartz, M. F., Montgomery, M. W., Buxbaum, L. J., Lee, S. S., Carew, T. G., Coslett, H., Ferraro, M., Fitzpatrick-DeSalme, E., Hart, T., & Mayer, N. (1998). Naturalistic action impairment in closed head injury. *Neuropsychology*, *12*(1), 13.

Schwartz, M. F., Reed, E. S., Montgomery, M., Palmer, C., & Mayer, N. H. (1991). *The quantitative description of action disorganisation after brain damage: A case study* (Vol. 8). https://doi.org/10.1080/02643299108253379

Seabold, S., & Perktold, J. (2010). Statsmodels : Econometric and Statistical Modeling with Python, (Scipy), 57–61.

Sebold, M., Schad, D. J., Nebe, S., Garbusow, M., Junger, E., Kroemer, N. B., Kathmann, N., Zimmermann, U. S., Smolka, M. N., Rapp, M. A., Heinz, A., & Huys, Q. J. M. (2016). Don't Think, Just Feel the Music: Individuals with Strong Pavlovian-to-Instrumental Transfer Effects Rely Less on Model-based Reinforcement Learning. *Journal of Cognitive Neuroscience*, *28*(7), 985–995. https://doi.org/10.1162/jocn_a_00945

Shahar, N., Hauser, T. U., Moutoussis, M., Moran, R., Keramati, M., Consortium, N. S., & Dolan, R. J. (2019). Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Computational Biology*, *15*(2), 1–25. https://doi.org/10.1371/journal.pcbi.1006803

Shallice, T. (1972). Dual functions of consciousness. *Psychological Review*, *79*(5), 383–393.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *298*(1089), 199–209. https://doi.org/10.1098/rstb.1982.0082

Shallice, T. (2006). Contrasting domains in the control of action: the routine and the non-routine. In *In Y. Munakata & M. H. Johnson (Eds.), Processes of change in brain and cognitive development. Attention and performance XXI* (pp. 3–29). Oxford: University Press.

Shallice, T., Burgess, P. W., & Robertson, I. (1996). The domain of supervisory processes and temporal organization of behaviour. *Philosophical Transactions: Biological Sciences*, *351*(1346), 1405–1412.

Shallice, T., Burgess, P. W., Schon, F., & Baxter, D. M. (1989). The origins of utilisation behaviour. *Brain*, *112*, 1587–1598.

Shallice, T., Stuss, D., Picton, T., Alexander, M., & Gillingham, S. (2008). Mapping task switching in frontal cortex through neuropsychological group studies. *Frontiers in Neuroscience*, *2*, 79–85.

Shapiro, H. M., Wong, L. M., & Simon, T. J. (2013). A cross-sectional analysis of the development of response inhibition in children with chromosome 22q11 . 2 deletion syndrome, *4*, 1–10. https://doi.org/10.3389/fpsyt.2013.00081

Sharp, M. E., Foerde, K., Daw, N. D., & Shohamy, D. (2015). Dopamine selectively remediates model-based reward learning: A computational approach [PMID: 26685155]. *Brain*, *139*, 355–64. https://doi.org/10.1093/brain/awv347

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: Ii. perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S.-C., & Davidson, M. C. (2010). Memory Maintenance and Inhibitory Control Differentiate from Early Childhood to Adolescence. *Developmental Neuropsychology*, *35*(6), 679–697. https://doi.org/10.1080/87565641.2010.508546

Simone, & Baylis. (1997). - selective attention in a reaching task: Effect of normal aging and alzheimer's disease.

Smigasiewicz, K., Ambrosi, S., Blaye, A., & Burle, B. (2020). Inhibiting errors while they are produced: Direct evidence for error monitoring and inhibitory control in children. *Developmental Cognitive Neuroscience*, *41*, 100742. https://doi.org/10.1016/j.dcn.2019.100742

Smith, L. B., Yu, C., Yoshida, H., & Fausey, C. M. (2015). Contributions of Head-Mounted Cameras to Studying the Visual Environments of Infants and Young Children. *Journal of Cognition and Development*, *16*(3), 407–419. https://doi.org/10.1080/15248372.2014.933430

Smith, P., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience*, *27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006.

Solum, M., Lorås, H., & Pedersen, A. V. (2020). A Golden Age for Motor Skill Learning? Learning of an Unfamiliar Motor Task in 10-Year-Olds, Young Adults, and Adults, When Starting From Similar Baselines. *Frontiers in Psychology*, *11*, 1–8. https://doi.org/10.3389/fpsyg.2020.00538

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition*, *95*(1), arXiv arXiv:1011.1669v3, 1–30. https://doi.org/10.1016/j.cognition.2003.12.004

Song, J. H., & Nakayama, K. (2006). Role of focal attention on latencies and trajectories of visually guided manual pointing. *Journal of Vision*, *6*(9), 982–995. https://doi.org/10.1167/6.9.11

Song, J. H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in Cognitive Sciences*, *13*(8), arXiv 0004057, 360–366. https://doi.org/10.1016/j.tics.2009.04.009

Speekenbrink, M. (2015). Uncertainty and Exploration in a Restless Bandit Problem, *7*, 351–367. https://doi.org/10.1111/tops.12145

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(29), 10393–8. https://doi.org/10.1073/pnas.0503903102

Stedron, J. M., Sahni, S. D., & Munakata, Y. (2005). Common Mechanisms for Working Memory and Attention: The Case of Perseveration with Visible Solutions. *Journal of Cognitive Neuroscience*, *17*(4), 623–631. https://doi.org/10.1162/0898929053467622

Steele, F. (2014). Multilevel modelling of repeated measures data. *LEMMA VLE Module*, *15*, 1–64. http://www.bristol.ac.uk/cmm/learning/course.html

Sternberg, S., Knoll, R., Monsell, S., & Wright, C. (1988). Motor Programs and Hierarchical Organization in the Control of Rapid Speech. *Phonetica*, *45*(2-4), 175–197. https://doi.org/10.1159/000261825

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Sumner, E., Li, A. X., Perfors, A., Hayes, B., Navarro, D., & Sarnecka, B. W. (2019). The exploration advantage: Childrens instinct to explore allows them to find information that adults miss. *PsyArXiv*.

Sutton, R. S., & Barto, A. (1998). *Reinforcement learning: An introduction*. Massachusetts, MIT Press,Cambridge.

Sutton, R. S., Precup, D., & Singh, S. (1999). Between mdps and semi- mdps: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, *112*, 181–211.

Thomas, M. S., & Johnson, M. H. (2008). New advances in understanding sensitive periods in brain development. *Current Directions in Psychological Science*, *17*(1), 1–5. https://doi.org/10.1111/j.1467-8721.2008.00537.x

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives - Nature. *Nature*, *407*(6805), 742–747. https://doi.org/10.1038/35037588

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology Section A*, *37*(4), 571–590. https://doi.org/10.1080/14640748508400920

Tipper, S. P., Howard, L. A., & Houghton, G. (1998). Action-based mechanisms of attention. *Philosophical Transactions of the Royal Society B: Biology*, *353*, 1385–1393.

Tipper, S. P., Howard, L. A., & Houghton, G. (2000). Behavioural consequences of selection form neural population codes. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance xviii* (pp. 222–245). The MIT Press.

Tipper, S. P., Howard, L. A., & Jackson, S. R. (1997). Selective reaching to grasp: Evidence for distractor interference effects. *Visual Cognition*, *4*, 1–38.

Tipper, S. P., Lortie, C., & Baylis, G. C. (1992). Selective reaching: Evidence for action-centred attention. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 891–905.

Tonidandel, S., & LeBreton, J. M. (2011). Relative Importance Analysis: A Useful Supplement to Regression Analysis. *Journal of Business and Psychology*, *26*(1), 1–9. https://doi.org/10.1007/s10869-010-9204-3

Towse, J., Lewis, C., & Knowles, M. (2007). When knowledge is not enough: The phenomenon of goal neglect in preschool children. *Journal of Experimental Child Psychology*, *96*, 320–332.

Toyama, A., Katahira, K., & Ohira, H. (2019). Biases in estimating the balance between model-free and model-based learning systems due to model misspecification. *Journal of Mathematical Psychology*, *91*, 88–102. https://doi.org/10.1016/j.jmp.2019.03.007

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5.

Tsay, J. S., Lee, A. S., Ivry, R. B., & Avraham, G. (2021). Moving outside the lab: The viability of conducting sensorimotor learning studies online. *bioRxiv*. https://doi.org/10.1101/2021.01.30.181370

Turner, M., & Engle, R. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharsk, S., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2020). The jasp guidelines for conducting and reporting a bayesian analysis. *Psychonomic Bulletin and Review*, 1–14. https://doi.org/10.3758/s13423-020-01798-5.

Vaughn, B., Kopp, C., & Krakow, J. (1984). The emergence and consolidation of self-control from eighteen to thirty months of age: Normative trends and individual differences. *Child Development*, *55*, 990–1004.

Verbruggen, F., Aron, A. R., Band, G. P., Beste, C., Bissett, P. G., Brockett, A. T., Brown, J. W., Chamberlain, S. R., Chambers, C. D., Colonius, H., Colzato, L. S., Corneil, B. D., Coxon, J. P., Dupuis, A., Eagle, D. M., Garavan, H., Greenhouse, I., Heathcote, A., Huster, R. J., … Boehler, C. N. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife*, *8*, 1–26. https://doi.org/10.7554/eLife.46323

Verschoor, S. A., Paulus, M., Spapé, M., Biro, S., & Hommel, B. (2015). The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models. *Cognition*, *138*, 64–78. https://doi.org/10.1016/j.cognition.2015.01.005

Von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Sciences*, *8*(6), 266–272. https://doi.org/10.1016/j.tics.2004.04.002

Voon, V., Baek, K., Enander, J., Worbe, Y., Morris, L. S., & Harrison, N. A. (2015). Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder [PMID: 26529423]. *Translational Psychiatry*, *5*, 670–8. https://doi.org/10.1038/tp.2015.165

Voon, V., Derbyshire, K., CRu, Irvine, M., Worbe, Y., & Enander, J. (2014). Disorders of compulsivity: A common bias towards learning habits [PMID: 24840709]. *Molecular Psychiatry*, *20*, 345–52. https://doi.org/10.1038/mp.2014.44

Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*(3), 279–292. https://doi.org/10.1007/BF00992698

Weisner, T. (1996). The 5 to 7 transition as an ecocultural project. In A. Sameroff & M. T. J. D. C. T (Eds.), *Macarthur foundation series on mental health and development. the five to seven year shift: The age of reason and responsibility* (pp. 295–326). University of Chicago Press.

Welsh, T. N., Elliott, D., & Weeks, D. J. (1999). Hand deviations toward distractors: Evidence for response competition. *Experimental Brain Research*, *127*(2), 207–212. https://doi.org/10.1007/s002210050790

Welsh, T. N., & Elliott, D. (2004). Movement Trajectories in the Presence of a Distracting Stimulus: Evidence for a Response Activation Model of Selective Reaching. *The Quarterly Journal of Experimental Psychology Section A*, *57*(6), 1031–1057. https://doi.org/10.1080/02724980343000666

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and Error: Common Neuronal Architecture for the Processing of Errors and Novelty. *Journal of Neuroscience*, *32*(22), 7528–7537. https://doi.org/10.1523/JNEUROSCI.6352-11.2012

White, E. J., Hutka, S. A., Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Frontiers in Systems Neuroscience*, *7*, 1–18. https://doi.org/10.3389/fnsys.2013.00090

Wifall, T., Buss, A. T., Farmer, T. A., Spencer, J. P., & Hazeltine, E. (2017). Reaching Into Response Selection : Stimulus and Response Similarity Influence Central Operations, *43*(3), 555–568.

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental psychology*, *35*(1), 205–213. https://doi.org/10.1037/0012-1649.35.1.205

Willingham, D. (1998). A neuropsychological theory of motor skill learning. *Psychol. Rev*, *105*, 558–584.

Willoughby, M., Wirth, R., Blair, C., & Investigators, F. L. P. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, *24*(2), 418–431. https://doi.org/10.1037/a0025779

Wirth, R., Foerster, A., Kunde, W., & Pfister, R. (2020). Design choices: Empirical recommendations for designing two-dimensional finger-tracking experiments. *Behavior Research Methods*, *52*(6), 2394–2416. https://doi.org/10.3758/s13428-020-01409-0

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-Month-Old Infants Interpret Action in Context. *Psychological Science*, *11*(1), 73–77. https://doi.org/10.1111/1467-9280.00218

Wunderlich, K., Smittenaar, P., & Dolan, R. (2012). Dopamine enhances model-based over model-free choice behavior [PMID: 22884326]. *Neuron*, *75*, 418–424. https://doi.org/10.1016/j.neuron.2012.03.042

Yanaoka, K., & Saito, S. (2017). Developing control over the execution of scripts: The role of maintained hierarchical goal representations. *Journal of Experimental Child Psychology*, *163*, 87–106. https://doi.org/10.1016/j.jecp.2017.06.008

Yanaoka, K., & Saito, S. (2019). Repeated sequential action by young children: Developmental changes in representational flexibility of task context. *Developmental Psychology*, *55*(4), 780–792. https://doi.org/10.1037/dev0000678

Yanaoka, K., & Saito, S. (2020). Contribution of Executive Functions to Learning Sequential Actions in Young Children. *Child Development*, *00*(0), 1–18. https://doi.org/10.1111/cdev.13489

Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T., & Wang, X. J. (2019). Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience*, *22*(2), 297–306. https://doi.org/10.1038/s41593-018-0310-2

Yin, H., Knowlton, B., & Balleine, B. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci*, *19*, 181–189.

Yntema, D. (1963). Keeping track of several things at once. *Human Factors*, *5*, 7–17.

Young, M. E., Sutherland, S. C., & Mccoy, A. W. (2018). Optimal go / no-go ratios to maximize false alarms, 1020–1029. https://doi.org/10.3758/s13428-017-0923-5