How Does Sense of Agency Develop Across Adolescence?

Hannah Slack, BSc, MSc, AFHEA

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

April 2023

Abstract

A sense of agency (SoA) refers to an individual's awareness of their control over their voluntary actions and the sensory consequences of those actions. Experiencing a veridical SoA is imperative to basic functioning as it facilitates effective goal-directed action. Despite this, a consensus on the trajectory at which the capacity to experience a SoA develops from childhood to adulthood has remained absent from past literature. To resolve this issue, SoA development was investigated by evaluating the influence of age on the functional efficiency of the forward model; the cognitive framework believed to generate a SoA. More specifically, the current research examined the extent to which children, adolescents and adults could, i) accurately predict the outcome of their action, and ii) update their action-outcome knowledge following post-action feedback; two skills indicative of a precise forward model.

A synchronisation-continuation task (chapter 3) was used to assess the impact of age on both the capacity to form veridical action-outcome predictions and update action-outcome knowledge in children, adolescents and adults. To isolate the effect of age on action-outcome prediction, a cued reaction time task (chapter 4) and a goal-switching task (chapter 5) were also administered to children, adolescents, and adults. Likewise, an outcome learning task (chapter 6) was used to assess how post-action learning changes from adolescence to adulthood. It was revealed that the frequency at which individuals engage in action-outcome prediction (chapter 4) and the quality of those predictions (chapters 3 and 5) improves with age. Similarly, the accuracy (chapter 3) and magnitude (chapter 6) to which individuals can update action-outcome knowledge in response to feedback was also found to refine with age. Moreover, the results of this thesis extend prior knowledge by suggesting that forward model precision, and thereby, the capacity to experience a SoA, develops with age across childhood, adolescence, and young adulthood.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Professor Stephen Jackson, Professor Georgina Jackson, and Dr Danielle Ropar, for their advice, patience, and support throughout this project. I have learned a great deal from this experience and I am indebted to you for giving me this opportunity. I am also very grateful to my viva examiners, Professor Clare Press and Professor Ellen Townsend, for generously providing their knowledge and expertise. Additionally, I would also like to thank the Economic and Social Research Council for funding my research; I am extremely grateful to have been given this life-changing opportunity.

I would also like to give my sincerest thanks to all the PsychoPy staff and contributors to the https://discourse.psychopy.org forums, in particular Dr Jon Peirce, Wakefield Morys-Carter, and Dr Rebecca Hirst, for their support in programming online behavioural tasks and code debugging. Switching from offline to online experiments half-way through the project was an anxiety-provoking experience; I would not have been able to progress without your endless assistance. Furthermore, I would also like to thank the staff at Summer Scientist Month, Belper School, Tuxford Academy, and Tourettes Action for all their help with recruitment, and everyone who participated in my research, without whom this project would not have been possible. I also wish to express my gratitude to Professor Michael Browning for his advice on the design of the outcome learning task, Dr Beverly Brown for her help in understanding the two-horse linear rise-to-threshold model, Dr Thomas Pronk for his guidance on using web-based eye-tracking software, and Dr Neil Roach and Lucy McKeown for their support in implementing Bayesian learning models.

I would also like to thank my friends in the department, in particular, Lucy McKeown, Karl Miller, and Felix Lewandowski, for their continued support and for our many interesting conversations. I am also grateful to Dr Emma Whitt for her generous encouragement and support throughout my studies.

I would like to say a huge thank you to my incredible parents for their understanding and support throughout my academic studies. Finally, a special thank you to Jay, who has been there for me throughout my PhD with an abundance of encouragement, compassion, and laughter.

Covid-19 Impact Statement

In March 2020, the UK government announced restrictions on all non-essential contact in response to the Covid-19 pandemic. This included the immediate suspension of all face-to-face data collection and teaching activities (Brown & Kirk-Wade, 2021). The following statement will outline the implications of these restrictions for the current research.

As will be discussed within the empirical chapters, the current research primarily consisted of online behavioural experiments. However, the presented studies are not reflective of the original research plan for this project. In 2019, preparations had been made to run in-person behavioural experiments exclusively. More specifically, offline versions of the synchronisation-continuation task (see chapter 3) and cued RT task (see chapter 4) had been designed in MATLAB using Psychtoolbox (Kleiner et al., 2007). An electroencephalogram (EEG) was also intended to be used to record participants' contingent negative variation (CNV) during the cued RT task, as implemented in prior developmental research (e.g., Bender et al., 2005; Thillay et al. 2015). The CNV is believed to be indicative of stimulus expectancy and motor preparation (Kononowicz & Penney, 2016), and hence, would have provided an additional measure of participants' capacity to form outcome predictions. Furthermore, there was a plan to run an experiment to explore participants' sensory adaptation, or in other words, their ability to adjust the direction of their reaching movements in response to sensory feedback. This would have involved using a hand-held robot manipulandum device to exert artificial perturbations on participants' expected action-outcome contingencies, similar to the device used in research by Kim et al. (2019b).

In response to the announced restrictions on in-person contact, the current research underwent a complete shift in focus from offline behavioural experiments to online behavioural studies, i.e., experiments that could be completed remotely without face-to-face contact with a researcher. As a result, the synchronisation-continuation task and cued RT task were entirely reprogrammed within PsychoPy to facilitate the online data collection. To achieve this, a considerable amount of time was spent completing self-paced online courses in Python; the programming language needed to recreate the planned tasks within PsychoPy. In addition, given that participants were unlikely to have access to an EEG device or a hand-held robot manipulandum device within their home, plans to utilise a CNV measure in the cued RT task and conduct a sensory adaptation study were abandoned.

In addition to the suspension of all face-to-face research activities, all UK primary and secondary schools were closed intermittently throughout 2020, and some of 2021, as part of government-imposed restrictions on all in-person teaching. This meant that school educators and pupils had to navigate online teaching and learning for the first time, leading to increases in perceived workload, stress and burnout for both groups (Beames et al., 2021; Commodari & La Rosa, 2021; Kim et al., 2022). This caused a considerable delay to the recruitment of individuals aged 13-17 for all four of the current studies. Attempts were made to contact several high schools regarding the recruitment of adolescents for the current project during this time. However, quite understandably, many educators commented that they did not have the capacity or resources available to support any research studies. Even when schools did reopen, the majority of educators contacted noted that their workload still left little space to support any external projects, as they adjusted back to in-person teaching. To speak candidly, this created a great deal of anxiety for me. As the end of 2021 arrived, it was unclear whether any adolescent participants would be recruited at all.

Fortunately, at the beginning of 2022, educators at two different secondary schools confirmed that they were able to facilitate participant recruitment for the current project. However, it should be noted that the difficulty experienced in gaining access to adolescent participants restricted the quantity of adolescent participants that were recruited. Likewise, given that all primary schools were also closed during the pandemic, it was only possible to recruit child participants via Summer Scientist Month, an event which ran online in 2020 and 2021. Again, this limited the number of child participants recruited, as well as the diversity of ages that I was able to collect data for.

Finally, the current research intended to recruit adolescents with Tourette Syndrome for all four of the present studies. Tourette Syndrome is a developmental disorder associated with involuntary movements and vocalisations (Cavanna et al., 2009). Previous research has shown that individuals with Tourette Syndrome exhibit both a deficit in their ability to update their forward model (Kim et al., 2019b) and an impaired sense of agency (Zapparoli et al., 2020). Therefore, by comparing the extent to which those with and without Tourette Syndrome could predict the most probable consequences of their actions and update their action-outcome knowledge, the current research intended to gain a more comprehensive understanding of how the neurotypical adolescent forward model system operates. Recruitment of individuals with Tourette Syndrome occurred via social media and through the charity, Tourettes Action. Unfortunately, an insufficient number of participants with Tourette Syndrome were recruited. Notably, government-imposed restrictions on face-to-face contact limited the avenues through which young people with Tourette Syndrome could be accessed. For instance, many support groups for individuals with Tourette Syndrome were not operating according to their normal schedule. Moreover, four planned chapters of this thesis were discarded.

In summary, the Covid-19 pandemic and associated government-imposed restrictions warranted alterations to the planned research methods and introduced delays and constraints to participant recruitment. Thank you for taking this consideration when reading this thesis.

Table of Contents

List of Figures

List of Tables

Chapter 1: General Introduction

Chapter Summary

Chapter 1 takes the form of a narrative review, beginning with an overview of how sense of agency has been conceptualised in past literature. Current knowledge regarding how a sense of agency might develop will then be discussed. Following this, the concept of a forward model will be introduced and the relevance of forward models to the topic of agency will be outlined. Finally, past research on how the quality of an individual's forward model might improve with age will be discussed, which will then lead to the aims of this thesis.

What is a Sense of Agency?

A sense of agency (SoA) refers to an individual's awareness of their control over their voluntary actions and the sensory consequences of those actions (Haggard & Chambon, 2012). Given that an individual will author numerous voluntary actions over the course of a single day, experiencing a SoA is imperative to basic functioning (Haggard, 2017). In particular, it enables the individual to identify and continuously monitor the impact of their own actions upon their environment (Liljeholm, 2021). Through this knowledge, they can then maintain an up-to-date understanding of how they can interact with the physical world in order to achieve their goals (Moore & Obhi, 2012). In addition to shaping our knowledge of the external world, possessing a SoA is also paramount to maintaining an accurate sense of self as it provides a means through which to differentiate our own actions from those of other agents (David et al., 2011).

When attempting to specify the precise characteristics of an agency experience, it is important to first acknowledge the idea that SoA can be subdivided into two separate constructs: a feeling of agency (FoA) and a judgement of agency (JoA; Synofzik et al., 2008). A FoA is defined as a low-level, internal 'buzz' sensation which informs the individual of their control over their voluntary action and its effect. It has been suggested that a FoA is formed through an internal computation where the expected consequence of a voluntary action is compared with what actually occurred after the action was performed. When a match is detected, a FoA is experienced and the event is classified as self-authored (Synofzik et al., 2013). The FoA has been argued to rely predominantly on the cerebellum and sensorimotor areas, such as the pre-supplementary motor area (pre-SMA; Seghezzi et al., 2019) due to their role in predicting the sensory consequences of intended actions (Seghezzi & Zapparoli, 2020; Welniarz et al., 2021) and detecting disparity between predicted and observed sensory outcomes (Gabitov et al., 2020; van Kemenade et al., 2019).

In contrast to the FoA, the JoA is characterised as a higher-order, explicit belief regarding the most likely cause of an observed outcome. JoAs are believed to be produced through a process of conscious reasoning where the individual's personal beliefs regarding their ability to execute the action, the perceived alignment between their intention and the observed outcome, and their perception of the contextual information available post-action execution, such as the spatial and temporal proximity of other potential sources of the sensory event, are taken into account (Desantis et al., 2011; Weiss et al., 2014). The JoA has been associated with activation in prefrontal areas, such as the dorsal medial prefrontal cortex (Sperduti et al., 2011), an area linked to the processing of salient context cues (Moorman & Aston-Jones, 2015) and hypothetical reasoning about others' motives (Molenberghs et al., 2016), and the ventromedial prefrontal cortex (Subramaniam, 2021), which has been argued to contribute to self-referential processing and the maintenance of self- vs other-related cues in working memory (Yin et al., 2021).

Evidently, where a FoA is merely concerned with establishing whether or not the individual themselves was the author of an observed event, a JoA moves a step beyond this and seeks to establish, if not the agent themselves, then who or what caused the event to happen (Haggard, 2017). Unsurprisingly, researchers grounded in action literature tend to focus exclusively on examining the FoA, centring their investigations on understanding the cognitive and neural architecture underlying this internal agency sensation (e.g., Haggard & Clark, 2003). Whereas, phenomenological and social psychologists tend to be more concerned with understanding the JoA than the FoA, often framing their findings in the context of moral and legal responsibility (Lacey, 2016; Pulkkinen & Aaltonen, 2003; Sidarus et al., 2020), as well as perceived free-will and wellbeing (Bandura, 1989; Saarikallio et al., 2020).

Research directly focused on determining the maturation of either facet of agency is notably lacking from past literature (Choudhury et al., 2007). Furthermore, of the research that has explicitly investigated SoA development, it may be argued that the primary focus has been on understanding children's and adolescents' JoA (e.g., Nobusako et al., 2020, 2022; van Elk et al., 2015; Weijs et al., 2021) with relatively few studies concerned directly with revealing the nature of their FoA (e.g., Cavazzana et al., 2014, 2017). Arguably, this is likely the result of the relative ease with which a JoA can be measured in comparison to a FoA (Synofzik et al., 2013). To examine the accuracy of an individual's JoA, a researcher will often manipulate the visual feedback available after a participant's action, either temporally (Nobusako et al., 2020) or spatially (Metcalfe et al., 2010). They then need only ask the participant the extent to which they attribute the observed event to their own action or to that of another agent (Saito et al., 2015). Whereas, a FoA is not directly observable as it occurs internally and is thus harder for researchers to quantify (Haggard, 2017).

The lack of direct research into the developmental trajectory of the FoA is a prominent issue given that a FoA is arguably more commonly experienced than a JoA and is therefore of greater use in everyday life (Haggard, 2017). Individuals are rarely required to make explicit evaluations regarding their sense of control over their actions in day-to-day life. Whereas, an internal experience of the fluidity between a planned action and its external effect is frequently experienced (Kühn et al., 2013). In addition, given that the FoA has been suggested to rely on different brain regions than the JoA, it follows that these two constructs will likely develop at differing rates (Saito et al., 2015). Indeed, research examining brain development from childhood to adulthood has consistently shown that cortical grey matter develops in a region-specific manner with motor and sensory systems maturing earliest and higher-order areas demonstrating a protracted development period (Gogtay et al., 2004; Lenroot & Giedd, 2006), thus suggesting that the FoA may develop at a faster rate than the JoA. Therefore, the current research aimed to determine the developmental trajectory of the FoA, as its presence, or lack thereof, is essential to individuals' everyday interactions with their environment. For simplicity and to remain consistent with previous action literature (e.g., Haggard et al., 2002), the term SoA will be used to refer to the FoA exclusively throughout the remainder of this thesis.

The Development of Agency

We will now turn to current knowledge regarding how a SoA might develop. Despite being essential for daily life, few studies have directly investigated how SoA reaches an adult-like level of precision (Choudhury et al., 2007). Furthermore, of the research that has explicitly investigated SoA development, the primary focus has centred on comparing the experiences of children with those of adults (e.g., Cavazzana et al., 2014, 2017), thereby omitting adolescence from the discussion. Notably, of the two studies which have examined how an experience of agency differs in childhood relative to adulthood, both concluded that children have a reduced SoA compared to adults (Cavazzana et al., 2014, 2017). This implies that SoA develops linearly with age, meaning that adolescents will likely outperform children but perform comparatively worse than adults. Indeed, this idea is consistent with past neuroimaging research, which has highlighted adolescence as a critical development period featuring significant maturational changes in regions of the brain previously linked to SoA, such as the pre-SMA, the parietal cortex and the cerebellum (Blakemore et al., 2012; Fuhrmann et al., 2015; Shaw et al., 2008; Sowell et al., 1999; Wierenga et al., 2014; Zito et al., 2020).

Thus far, this assertion of an age-related improvement in SoA across adolescence has not been supported. To date, only two studies have directly examined SoA in adolescence (Aytemür et al., 2021; Aytemür & Levita, 2021). In research by Aytemür and Levita (2021), children (9-10), mid-adolescents (13-14), late-adolescents (18-20) and adults (25-28) were instructed to make a keypress at any time of their choice whilst viewing an analogue clock. After a delay, a tone was delivered and participants were asked to report the time on the clock at the moment which the tone was heard. The extent to which participants exhibited the outcome binding effect was then recorded. The outcome binding effect refers to a phenomenon in which an individual judges a sensory event to have shifted temporally towards their action. This effect is believed to only occur when the individual believes that their action caused the event to transpire (Render & Jansen, 2021). Hence, the outcome binding effect has been suggested to measure the extent to which an individual experiences agency over a given event (Borhani et al., 2017). Aytemür and Levita (2021) found that the outcome binding effect declined from childhood to late-adolescence before returning to a level similar to that observed in children at adulthood. Therefore, contrary to the proposed notion of a linear improvement in SoA with age across ontogeny, this finding suggests that SoA development follows a U-shaped trajectory, with a marked decrement in adolescence.

Intriguingly, when the same researchers measured the outcome binding effect again in mid-adolescents (13-14), late-adolescents (18-20) and adults (25-28) using a near identical task, they failed to replicate their previous findings (Aytemür et al., 2021). It was found that the outcome binding effect was greater in mid-adolescents compared to adults, whilst outcome binding in late-adolescents did not differ from the levels of binding observed in mid-adolescents or adults. Initially, this appears to directly contradict their previous conclusion of a poorer SoA amongst adolescents relative to adults. However, Aytemür et al. (2021) argued that the reason for the disparity between the findings of their two studies was that they had implemented a longer delay of 450ms between participants' actions and the tone in the current task compared to the 250ms delay used in Aytemür and Levita (2021).

In support of the idea proposed by Aytemür et al. (2021) to explain their results, previous research has argued that there is a cognitive mechanism, known as a temporal binding window, which determines the maximum delay that can exist between an action and an effect for the two to still be perceived as causally related (Jaime et al., 2014). It has been shown that the temporal binding window narrows with age through childhood, adolescence and young adulthood (Hillock-Dunn & Wallace, 2012; Nobusako et al., 2018). Therefore, the results obtained by Aytemür et al. (2021) demonstrate that mid-adolescents are more willing to accept a longer delay between an action and an effect than adults. Arguably, this tendency suggests that adolescents are more likely to incorrectly bind unrelated actions and effects together, and thereby falsely perceive an event as self-produced compared to adults. This interpretation then implies that the results of Aytemür et al. (2021) are indeed consistent with the conclusion drawn by Aytemür and Levita (2021), as it suggests that adolescents have a less precise and more error-prone SoA compared to adults.

Alternatively, if it were the case that adolescents' expanded temporal binding window meant that they were more likely to perceive a tone that occurred 450ms after their action as self-caused than adults in research by Aytemür et al. (2021), then it is unclear why adolescents in research by Aytemür and Levita (2021) were less able to perceive a tone that onset 250ms after their action as self-produced compared to adults. Intuitively, it follows that, if adolescents' temporal binding window can accommodate a 450ms delay between an action and its effect, then a shorter 250ms delay should fit adequately within that same window. For this reason, it can be argued that the explanation that was proposed by Aytemür et al. (2021) to explain their findings was specious, as it is incompatible with the results of Aytemür and Levita (2021). Subsequently, given that adolescents' experience of agency has only been investigated in two prior studies with arguably contradictory results, the precise manner through which SoA matures across this development period remains ambiguous. Hence, the ultimate goal of the current research was to determine whether SoA matures at a linear rate across ontogeny, as suggested by past child studies (e.g., Cavazzana et al., 2014; 2017) and neuroimaging research (e.g., Blakemore et al., 2012).

Notably, all four of the noted studies that have sought to determine the developmental trajectory of agency have employed a version of the intentional binding effect to measure SoA (Cavazzana et al., 2014, 2017; Aytemür et al., 2021; Aytemür & Levita, 2021). The intentional binding effect refers to a temporal compression which occurs between the perceived timing of an action and its sensory effect when the effect is thought to be self-authored (Haggard, 2017). To clarify, the outcome binding effect recorded by Aytemür and Levita (2021) and Aytemür et al. (2021) is often conceptualised as a subcomponent of the intentional binding effect that focuses exclusively on the perceived shift in the timing of the sensory effect relative to the voluntary action (Render & Jansen, 2021). It has been argued that the intentional binding effect is reliant on the same cognitive system as a SoA, as creating a mismatch between the expected and actual timing of the action consequence via the addition of a temporal delay has previously been found to diminish the intentional binding effect (Wen, 2019). For this reason, the intentional binding effect has commonly been used as a measure of agency within past research (Haggard, 2017).

In criticism of the noted reliance on the intentional binding effect in past studies, it has been argued that this effect may not measure SoA exclusively (Suzuki et al., 2019). This can be said as it has also been associated with causal inference and multisensory integration in absence of any volitional action (Kirsch et al., 2019; Lorimer et al., 2020; Suzuki et al., 2019). Therefore, by studying how the functionality of the cognitive framework underlying SoA changes with age across childhood, adolescence and young adulthood, it will be possible to gain a more direct understanding of SoA development than has been achieved in prior intentional binding studies.

Forward Models and Their Relevance to Agency

The internal computation through which predicted and observed action outcomes are compared in order to produce a SoA is believed to occur within a cognitive framework, known as a forward model (Haggard & Chambon, 2012). Forward models were originally conceptualised in the context of optimal motor control (e.g., Wolpert et al., 1995). Our environment is an inherently dynamic space. Hence, the brain often receives sensory input from multiple different sources simultaneously (Wolpert & Flanagan, 2001). Therefore, in order to regulate the impact of our motor actions on the environment, it has been argued that the brain maintains two classes of internal, computational model: forward sensory models and forward dynamic models (Wolpert et al., 1995). When an individual intends to perform an action, the brain first generates a motor command. This motor command contains instructions for the muscles on how to manipulate the current body state in order to execute the desired action. A forward sensory model uses an efference copy of the motor command to form a prediction regarding the most probable sensory consequences of the intended action. Whereas, a forward dynamic model uses the motor command to predict the most likely end body state that will be reached post-action execution (Wolpert & Ghahramani, 2000).

When these internal models were first conceptualised, the precise mechanism through which these predictions are created was not explicitly specified (e.g., Wolpert et al., 1995). Given that our sensory receptors are unable to perceive sensory events with perfect acuity, it has been suggested that each signal arriving into the brain will be corrupted by a degree of perpetual and neural noise (van Beers et al., 2002; Neri, 2010; Wallace & Stevenson, 2014). Outside of the action context, perception researchers have argued that the brain is able resolve the uncertainty introduced by this noise and thereby, predict probable future sensory events, by weighting relevant knowledge gained from past experience against current contextual information using Bayes' theorem (Vilares & Körding, 2011; see equation 1.1, as outlined by Shi et al., 2013). For example, when faced with an unidentifiable figure in a dimly lit room, pre-acquired knowledge, such as the identity of all the other occupants in the house, is combined with contextual information, such as the physical stature of the figure, in order to presume the identity of the mystery individual.

To briefly explain the notation shown in equation 1.1, E_{pr} refers to an average estimate of the sensory evidence accumulated from past experience, otherwise known as the prior. Whereas, E_{li} refers to the likelihood, or in other words, the individual's perceptual estimate of the available contextual cues. W_{pr} and W_{li} represent the weight awarded to the prior and the likelihood, respectively. Finally, $Posterior$ refers to the prediction produced as a result of the weighted combination of the prior and the likelihood.

$$Posterior = (W_{pr} * E_{pr}) + (W_{li} * E_{li}) \tag{1.1}$$

Crucially, in order to minimise the level of noise present in the final prediction, it has been argued that the likelihood and the prior are not necessarily integrated with an equal weighting to produce the posterior (Moore & Fletcher, 2012). For instance, if the individual's sensory system is unable to make a reliable estimate of one or more sensory cues available in the environment, then less weight will be awarded to the likelihood information relative to the prior evidence (Ernst & Banks, 2002; Yon & Frith, 2021). On the other hand, if the individual's prior knowledge is unreliable due to a lack of relevant experience, then the computed posterior will be based more on the likelihood than on the prior (Knill & Pouget, 2004). Therefore, it is believed that by attuning the weight assigned to the prior and likelihood, the veridicality of the posterior prediction is maximised.

Subsequent literature has also sought to apply Bayes' theorem to action research, thus suggesting that the predictions of forward sensory models and forward dynamic models are formed through a weighted combination of a prior and a likelihood (e.g., Faisal et al., 2008; Franklin & Wolpert, 2011; Legaspi & Toyoizumi, 2019). In the context of the forward sensory model, a prior refers to a belief regarding the typical result of the intended action, based on an average estimate of all previous iterations of the planned action. Whereas, a likelihood refers to an estimate of the sensory information available in the environment before action execution, such as the perceived properties of an encountered object (Berniker & Körding, 2011; Di Luca & Rhodes, 2016). Therefore, from a Bayesian perspective, the predictions of a forward sensory model are formed through the combination of both internal prior knowledge cues and external likelihood cues (Moore & Fletcher, 2012).

After the motor action has been completed, it is believed that a comparator mechanism within the forward sensory model then compares the predicted outcome with the actual sensory information observed (Carruthers, 2012). If a disparity is detected between the expected and observed outcomes, then one of two consequences occurs: either the forward sensory model is updated or the model remains unchanged and instead, a different action is chosen for any subsequent movements (Chambon et al., 2014). The magnitude to which the forward sensory model is modified varies according to the outcome of the comparator mechanism. The larger the discrepancy detected by the comparator, the greater the extent to which the probabilistic associations between the action and its potential effects are altered. The updated model is then used to generate new outcome predictions when the individual performs the target action again in the future (Franklin & Wolpert, 2011).

A SoA is believed to arise as a by-product of the internal computation that occurs at the comparator mechanism; the greater the alignment between expectations and observations, the greater the probability that a SoA will be experienced (Sato & Yasuda, 2005). Subsequently, the accuracy with which an individual can experience a SoA over an observed sensory event is dependent on the precision of their forward model prediction. In turn, the veridicality of those predictions is thought to be reliant on the individual's ability to utilise the information gained from past action to maintain an up-to-date conceptualisation of the action-outcome contingencies relevant to the current context (Asai, 2017).

In support of the assertion that both the optimisation of motor control and SoA rely on the precision of the forward sensory model, past research has shown that both processes rely on the same areas of the brain, such as the cerebellum, the pre-supplementary motor area, and the angular gyrus (Tanaka et al., 2020; Welniarz et al., 2021). Similarly, past research has demonstrated that introducing a spatial or temporal discrepancy between a voluntary action and its outcome creates an incongruency between the anticipated and actual consequence of the action, and thereby diminishes the individual's sense of control over the observed event (Haggard et al., 2002; Kirsch et al., 2016; Ruess et al., 2018). Taken together, these findings suggest that through studying the functionality of the forward sensory model, it will be possible to obtain a proxy measure of a participant's SoA. For simplicity, the forward sensory model will be referred to as the forward model for the remainder of this thesis.

Incidentally, it should briefly be acknowledged that some past studies have adopted a Bayesian account to describe how outcome feedback can alter a learned action-outcome association (e.g., Berniker & Körding, 2011; Hohwy, 2017; Körding & Wolpert, 2006). Yon et al. (2020) argued that both the predicted and actual sensory information are entered into a weighted comparison. When the individual's perceptual estimate of the observed outcome is more precise than their expectation, greater weight is assigned to this new evidence over their prediction. Their prior action-outcome knowledge will then remain unchanged and a new action will be selected for any subsequent actions. Whereas, when a prediction is less reliable than the sensory feedback observed, less weight is attributed to the prediction relative to the observed feedback. As a result, the individual will then incorporate the incoming sensory feedback into their prior estimate. Given that the prior is believed to be an average estimate of a range of past observations, it has been argued that, the larger the dissimilarity between the newly observed outcome and this distribution of past observations, the greater the extent to which the prior estimate is altered by this new outcome information (Berniker & Körding, 2011; Hohwy, 2017). Evidently, this Bayesian account of motor control appears to be compatible with the more traditional comparator mechanism account; larger discrepancies between predicted and observed outcomes result in more substantial changes to the conceptualised action-outcome relationship.

On a similar note, some studies have also taken a Bayesian approach to explain the manner in which a SoA can emerge (e.g., Legaspi & Toyoizumi, 2019). It has been argued that, the higher the probability that the expected and perceived outcomes were produced from the same prior estimate, the greater the individual's SoA over the observed event (Legaspi & Toyoizumi, 2019). Again, it appears that this Bayesian perspective of SoA production is in agreement with the traditional comparator mechanism account; the higher the correspondence between an anticipated and an observed outcome, the greater the probability that a SoA will be experienced. Ergo, this reinforces the decision to measure the functionality of the forward model as an index of one's capacity to experience agency within this thesis.

The Development of the Forward Model

Attention will now turn to current knowledge regarding the developmental trajectory of the forward model, given its suggested role in constructing a SoA. In particular, this subsection will discuss how the ability to i) appropriately update action-outcome knowledge in response to incoming action feedback and ii) form accurate action-outcome predictions matures with age, as both abilities demonstrate the precision of the forward model (Desmurget & Grafton, 2000).

In order to predict the consequence of a planned action, it has been suggested that an individual must first be able to conceptualise the causal associations between their own actions and observed sensory events (Assaiante, 2012). This action-outcome knowledge is not innate; infants will often flail their limbs in the absence of a distinct goal (Adolph & Franchak, 2017). However, these initial explorative movements provide infants with the opportunity to build associations between specific actions and their effects (Paulus et al., 2012). Indeed, it has been demonstrated that the capacity to learn about action-effect contingencies after repeated exposure, and use this knowledge to guide further action, is present from early infancy (Gredebäck et al., 2018; Kretch & Adolph, 2012). For example, Watanabe and Taga (2006) demonstrated that, when the previously observed result of their leg kick suddenly became absent, 2-month-old infants increased the frequency of their subsequent kicking movements. This suggests that, from 2-months-old, infants can acquire an understanding of the relationship between their kick and its effect from past action feedback. Not only this, but this also illustrates that infants can use this knowledge to predict the outcome of their action, as evidenced by their increased leg-kicking frequency when their expectation was violated. Therefore, this suggests that individuals can construct and operationalise a forward model from early infancy.

Given evidence to suggest that the capacity to produce and utilise a forward model to support goal-directed action is already present from as early as 2-months-old (Watanbe & Taga, 2006), one could conclude that the objective of this thesis is redundant. However, the mere presence of a forward model does not guarantee that the model will operate with adult-like precision. Indeed, although it has been demonstrated that children can combine learned knowledge and new sensory evidence to generate causal predictions (Bejjanki et al., 2020; Griffiths et al., 2011; Sobel & Munro, 2006), it has been argued that the accuracy of those predictions are poorer compared to those of adults (Klevberg & Anderson, 2002; Plumert, 1995). For example, when younger children (aged 4-7 years), older children (8-11 years), and adults were asked to walk through doorways of varying widths, Franchak (2019) found that participants' accuracy in predicting whether or not they could fit through the available doorway space improved with age. This supports the idea that children possess a less sophisticated forward model compared to adults, as this finding suggests that children are less able to form accurate predictions regarding the most likely outcome of their action.

One potential reason for the noted lack of precision in children's outcome predictions is that the quantity of information that they are able to incorporate into their prior estimate has been argued to be poorer compared to that of adults (Chambers et al., 2018; Gopnik & Bonawitz, 2015). For example, in research by Barash et al. (2019), children aged 5-8 and adults drew yellow and green balls from an urn with replacement. Before each ball was drawn, participants predicted the ball's colour. It was found that, compared to adults, children tended to base their predictions on the outcome of only the most recent trials, as opposed to an average estimate of all the balls seen so far. As a result, it was reported that children demonstrate poorer choice accuracy than adults. This suggests that, relative to adults, children are less able to appropriately update their prior action-outcome knowledge to incorporate new outcome evidence, thus resulting in erroneous subsequent predictions. Consequently, this implies that children's inferior prediction accuracy results from a poorer ability to learn from post-action feedback when compared to adults. Therefore, this finding provides further support for the idea that children possess a less developed forward model system than adults.

As the quantity of information that children are able to incorporate into their prior is believed to be relatively limited compared to that of adults, it has been suggested that their predictions will often be biased towards the likelihood (Chambers et al., 2018). As previously noted, the sensory system is not an infallible machine able to perceive sensory events with impeccable acuity (Wolpert & Flanagan, 2001). This is particularly evident in children; adults have been shown to outperform children on tasks where they are asked to discriminate between different auditory and visual stimuli based on perceptual features, such as intensity, frequency, and duration (Bishop et al., 2011; Droit-Volet et al., 2007; Jensen & Neff, 1993; Zélanti & Droit-Volet, 2012). Likewise, it has been reported that children below the age of 10 are unable to integrate multiple sensory cues in an optimal manner (Ernst & Banks, 2002). Taken together, this suggests that children are less able to access an accurate perceptual estimate of the sensory information present within their environment compared to adults. Therefore, in addition to an unreliable prior, children's poor prediction accuracy can also be attributed to a tendency to over-rely on imprecise likelihood information. Additionally, this lack of perceptual acuity can also be suggested to undermine their capacity to evaluate the congruency between predicted and actual feedback and thereby, update action-outcome knowledge. Ergo, this provides further evidence in support of the idea that children have a less precise forward model system than adults.

Gaps in the Current Literature

In parallel to previous SoA literature (e.g., Cavazzana et al., 2014, 2017), the majority of past studies which have contributed to our understanding of forward model development have tended to compare how children and adults differ in their ability to predict action consequences (e.g., Franchak, 2019; Perchet & Garcia-Larrea, 2005) and adapt behaviour in response to sensory feedback (e.g., Wilson & Hyde, 2013; Tahej et al., 2012; Scheerer et al., 2016). In contrast, relatively few studies have explicitly examined the manner in which the forward model system develops in adolescence (Quatman-Yates et al., 2012; Barlaam et al. 2012; Dahl et al. 2018). Hence, the objective of the current thesis was to determine the full trajectory at which the forward model develops from childhood to adulthood, including across adolescence.

Adolescence is a developmental period which spans from the onset of puberty to the start of adulthood, typically occurring between the ages of 13-17-years-old (Jaworska & MacQueen, 2015). This period is believed to be marked by substantial structural and functional changes within the brain (Sisk & Gee, 2022; Smith et al., 2011) as a result of increased synaptic pruning and myelination (Whitford et al., 2007). These maturational changes have been argued to cause a progressive shift throughout adolescence from a bottom-up, stimulus-driven motor control strategy to a more top-down, proactive method of action control (Braver, 2012; Decker et al., 2016). In line with this idea, it has previously been argued that adolescents are better able to prepare appropriate motor responses in anticipation of target stimuli compared to children (Padilla et al., 2014; Van Gerven et al., 2016); a skill which necessitates the estimation of probable action-consequences (Wolpert & Flanagan, 2001). Furthermore, variation in adolescents' anticipatory motor control has previously been linked to age-related changes in the structure of their prefrontal cortex (Vijayakumar et al., 2014), a region of the brain implicated in both the proactive orientation of attentional resources towards expected stimuli (Bechara et al., 1996; Snyder et al., 2021) and the anticipatory representation of goal-related information (Chatham et al., 2009). Taken together, this suggests that the ability to predict the most probable outcome of a planned action should improve across adolescence, in line with developmental changes occurring within the brain.

In further support of the idea that adolescents' ability to anticipate action-consequences improves with age, it may be suggested that adolescents are able to construct more reliable prior and likelihood estimates compared to children. For instance, past research has shown that adolescents demonstrate improved sensory perception abilities compared to children, as evidenced by their enhanced performance on sensory discrimination tasks (Herman et al., 1996; Ladouceur et al., 2007). This implies that adolescents are better able to form an accurate perceptual estimate of the sensory information present within their environment than children. Hence, this suggests that the quality of the likelihood estimate that individuals are able to use when generating their forward model predictions improves with age from childhood to adolescence. Similarly, it has been reported that adolescents tended to use a greater volume of past trial outcomes to inform their subsequent choices on learning tasks compared to children (e.g., Barash et al., 2019; Master et al., 2020). This suggests that the reliability of individuals' prior estimate also undergoes age-related improvements from childhood to adolescence. Therefore, it may be argued that both the quality of forward model predictions and the ability to appropriately update learned action-outcome mappings should refine with age in adolescence, before reaching full maturity in adulthood.

Summary and Research Aims

To summarise, the purpose of this thesis was to rectify the absence of adolescents from prior SoA development literature, and thereby, establish the full developmental trajectory of a SoA from childhood to young adulthood. Given that a SoA is believed to be produced via a forward model (Haggard & Chambon, 2012), and a SoA itself cannot be directly observed (Haggard, 2017), the proficiency of individuals' forward model system was chosen as a proxy measure of agency. More specifically, the functionality of an individual's forward model system was examined via their ability to i) accurately predict the outcome of their action, and ii) update learned action-outcome knowledge in response to post-action feedback; two skills indicative of a precise forward model.

The primary goals of this thesis were twofold:

1. Based on past literature (e.g., Van Gerven et al., 2016), the first goal of this thesis was to test the idea that the ability to form accurate action-outcome predictions improves with age from childhood to adulthood.

2. In addition, the second goal of this thesis was to evaluate the suggestion that the ability to appropriately update learned action-outcome associations in light of post-action feedback improves with age from childhood to adulthood, as suggested by past research (e.g., Master et al., 2020).

Evidently, achieving these two initial goals will allow us to attain our ultimate thesis aim:

3. To assess the assertion that SoA matures at a linear rate from childhood to adulthood, as suggested by past child studies (e.g., Cavazzana et al., 2014; 2017) and neuroimaging research (e.g., Blakemore et al., 2012), in light of the contradictory evidence presented by Aytemür and Levita (2021).

The empirical studies within this thesis will address the first and second thesis goals. The implications of the findings drawn from the reported studies for the third thesis objective will then be discussed in chapter 7. Furthermore, the next chapter will outline how the first and second goals were investigated within this thesis, alongside the relative merits and limitations of the methods employed to accomplish this endeavour.

Chapter 2: A Comprehensive Discussion of the Methodology Used in the Current Research

Chapter Summary

Chapter 2 outlines how prediction accuracy and post-action outcome learning were assessed empirically within the current research, in addition to the relative merits and limitations of the methods used to achieve these goals. More precisely, this chapter begins by briefly discussing the specific online behavioural tasks that were employed across this thesis and how they facilitated the quantification of participants' prediction accuracy and post-action outcome learning. Next, the disadvantages of conducting behavioural experiments online are discussed, alongside their implications for the robustness of the current research and how the identified issues were mitigated. Following this, the parent- and self-report scales that were used to assess participants' level of trait impulsivity are described and their use in the current research supported with appropriate prior evidence. The benefits and challenges that were encountered when recruiting participants for the current online studies are then discussed. Finally, the approach to data analysis taken throughout this thesis is outlined and justified.

The Specific Behavioural Tasks Implemented Within the Current Research

Across chapters 3 - 6 of this thesis, four online behavioural tasks are presented, each of which were used to assess how the forward model reaches an adult-like level of precision. A relatively brief overview of each task, including their relevance to the forward model and why they were chosen for the current research, will now be presented. A more detailed discussion of how prediction accuracy and/or post-action outcome learning were quantified in each task can be found within each individual empirical chapter.

Chapter 3 - The Synchronisation-Continuation Task

The study described in chapter 3 used a synchronisation-continuation task to measure the accuracy and consistency of sensorimotor continuation in participants aged 4-25. Sensorimotor continuation refers to an individual's ability to maintain a specified temporal interval between each of their motor responses (McPherson et al., 2018). In keeping with past synchronisation-continuation studies (Repp & Su, 2013), participants were first instructed to press the spacebar in synchrony with a series of isochronous tones that were delivered at either a high, medium or low frequency. They were then required to continue pressing the spacebar at the same pace after the tones had been removed. The accuracy and consistency with which participants could maintain the same temporal interval between each of their keypresses as was present between the tones was then recorded as an indicator of their sensorimotor continuation skill.

Previous research has argued that effective sensorimotor continuation is dependent on the precision of the forward model (Maes, 2016). In order to accurately and consistently reproduce the target response pace, participants must use their prior knowledge of the target inter-response-interval to predict the optimal time at which to make their next response. They must then monitor for, and correct, any disparity between their produced inter-response-interval and the target inter-response-interval (Maes, 2016). Evidently, the greater the accuracy and consistency with which participants could maintain the target response pace, the better both their prediction accuracy and their post-action outcome learning. Therefore, the results of chapter 3 are relevant to both the first and second objectives of this thesis.

Aside from providing an effective method of indexing forward model precision, the synchronisation-continuation task was also selected for the current research because it is believed to be sufficiently simplistic for young children to understand, as demonstrated by its implementation in previous developmental studies (e.g., Monier & Droit-Volet, 2019). In addition, despite the fact that synchronisation-continuation tasks have only been run in offline contexts within prior research (e.g., Drewing et al., 2006), the task was also chosen because it was reasonably effortless to implement online. This can be said as no additional hardware was needed for participants to complete the study; only access to a keyboard and a stable Wi-Fi connection were required. Therefore, the synchronisation-continuation task was an ideal paradigm to utilise in the current research.

Chapter 4 - The Cued Reaction Time Task

Chapter 4 details an experiment in which a cued reaction time (RT) task was used to assess predictive motor timing in participants aged 4-25. Predictive motor timing refers to the ability to manipulate the timing of an intended action such that its occurrence aligns with the predicted onset of an imminent stimulus (Debrabant et al., 2012; Tanaka et al., 2021). Inspired by the tasks implemented in research by Brown (2019) and Burnett Heyes et al. (2012), participants were first presented with a red stimulus for a fixed interval during the cued RT task. After this, an amber cue stimulus was presented for a variable interval. Finally, the amber cue stimulus was replaced by a green target stimulus. Participants' objective was to make a mouse click response as soon as the target stimulus became visible.

Crucially, participants could achieve the task objective by either making an anticipatory response or a reactive response (Burnett Heyes et al., 2012). Anticipatory responses required participants to predict the most likely onset time of the target stimulus and thus, when best to respond such that their keypress temporally aligned with the target stimulus' onset. Whereas, reactive responses were triggered by the onset of the target stimulus, and thus, required no internal action preparation via the forward model in advance of the target stimulus' arrival (Braver, 2012). The ratio of anticipatory to reactive responses produced by each participant was recorded. Intuitively, anticipatory responses were more advantageous than reactive responses, as an anticipatory response would achieve faster reaction time relative to a reactive response. Consequently, the higher the anticipatory to reactive response ratio, the greater the participant's predictive motor timing, and thereby, the better their ability to form forward model predictions. Hence, unlike the synchronisation-continuation task described in chapter 3, the cued RT task was used to measure participants' prediction accuracy exclusively. Therefore, the results of chapter 4 address the first goal of this thesis.

Evidently, it can be argued that the objective of the cued RT task was fairly straightforward for both younger and older participants to understand, as evidenced by prior developmental research (e.g., Brown, 2019). This simplicity minimises the potentially confounding impact of age-related differences in task comprehension on the results, thus making it an ideal paradigm for the current research. Furthermore, past research has demonstrated that the cued RT task can easily be framed as a car race, given that the roles of the cue and target stimuli mirror those of real-world amber and green traffic lights (Burnett Heyes et al., 2012). Therefore, the cued RT was believed to be effective in maintaining the attention of younger participants, thus further justifying its suitability for the current research.

Chapter 5 - The Goal-Switching Task

The study outlined in chapter 5 used a goal-switching task to measure task-switching in participants aged 5-21. Task-switching refers to an individual's ability to flexibly shift between two or more different task objectives (Barcelo et al., 2006). Inspired by the task administered by Jung et al. (2015), on each trial of the goal-switching task, participants were presented with either a red or green stimulus positioned on the left or right side of the screen. When the stimulus was green, participants had to perform a pro-saccade, which involved moving their gaze towards the side of the screen that contained the stimulus. Whereas, when the stimulus was red, an anti-saccade was required; participants needed to move their gaze to the opposite side of the screen, away from the presented stimulus. Whether or not participants shifted their gaze in the correct direction was recorded for each trial. Hence, the goal-switching task required participants to shift between performing pro-saccades and anti-saccades across the trials.

To make a correct response on the goal-switching task, the participant needed to combine their prior knowledge of the causal action-outcome associations with the colour of the presented stimulus in order to determine how best to respond. Therefore, similar to chapter 4, the results of chapter 5 also address the first goal of this thesis. Switch costs and mixing costs were calculated for both pro-saccade trials and anti-saccade trials based on participants' response accuracy. Switch costs revealed the cost to accuracy of having to switch between action-outcome pairings when assessing how best to respond. Whereas, mixing costs demonstrated the cost to accuracy of having to maintain, and select between, different action-outcome associations (Manzi et al., 2011). Moreover, both switch costs and mixing costs provide an effective means through which to quantify an individual's ability to use appropriate prior knowledge to guide their current actions and suppress incorrect, automatic responses. Similar offline tasks have previously been used to assess task-switching in developmental studies (e.g., Reimers and Maylor, 2005). Moreover, this reinforced the decision to utilise the goal-switching task within the current research.

Chapter 6 - The Outcome Learning Task

Chapter 6 describes an experiment in which an outcome-learning task was administered to participants aged 14-24. Outcome learning refers to an individual's ability to alter their knowledge of the action-outcome contingencies present within their environment in response to post-action feedback (Kawato & Wolpert, 2007). In keeping with the task used by Browning et al. (2015), participants selected between two boxes during the outcome-learning task, one of which contained a reward. The task contained two conditions. In the stable condition, the relative probability that each box would lead to a reward outcome was fixed. Whereas, in the volatile condition, these probabilities shifted between the two boxes over time. Each participant's choices were then recorded. From this, one learning rate was calculated per condition for each participant. The learning rate revealed the extent to which the participant's action-outcome knowledge, and subsequent choice behaviour, was modified in response to the most recently observed trial outcomes. The higher the learning rate, the greater the influence of recent outcomes on learned action-outcome knowledge, relative to the wider history of observed feedback.

To maintain an up-to-date understanding of the probabilistic associations between actions and their effects, an individual must incorporate observed action feedback into their prior estimate (Berniker & Körding, 2011). Crucially, the rate at which these modifications to prior knowledge are made must be modulated according to the volatility of the current context (Behrens et al., 2007). In a relatively stable context, where probabilistic action-outcome relationships remain fixed over time, it is optimal to possess a low learning rate. As a result, each action outcome will only trigger a minor update to the individual's action-outcome knowledge (Behrens et al., 2008). Whereas, in a more volatile context, where action-outcome associations are subject to frequent change, a high learning rate is favourable. Consequently, a recent action outcome will trigger a substantial update to one's action-outcome knowledge (Browning et al., 2015). Therefore, this suggests that, by examining an individual's ability to optimally modify their learning rate according to the relative stability of the current context, it is possible to measure their capacity to make appropriate updates to their forward model. Hence, the results of chapter 6 address the second objective of this thesis. Notably, similar learning tasks have been used to assess individuals' capacity to update action-outcome knowledge in prior developmental research (e.g., Eckstein et al., 2020), making this task ideal for the current research.

Brief Summary

To summarise, participants' performance on a synchronisation-continuation task, a cued RT task and a goal-switching task were used to quantify their ability to predict the consequences of their own actions. Additionally, the synchronisation-continuation task and an outcome learning task revealed participants' capacity to update their action-outcome knowledge in response to past action feedback. All four tasks were selected for the current research based on their relative simplicity, thus meaning that they were easy for young participants to understand and straightforward to implement online.

Online Behavioural Experiments

Attention will now turn to establishing, more generally, the implications of conducting behavioural experiments in an online space and the specific impact that this had on the current research. An online behavioural experiment typically refers to a behavioural task, which has been programmed to run within a web browser. Online tasks are typically stored on a server, from which participants can access the task remotely via their own computer. Participants' responses are then uploaded to the same server and made available to the researcher (Grootswagers, 2020). For context, each of the four tasks outlined within this thesis were programmed using PsychoPy software and hosted online via Pavlovia. PsychoPy is a free, open-source application, which allows researchers to design behavioural experiments for both offline and online use (Peirce et al., 2019). Whereas, Pavlovia is a web-based platform through which researchers can upload behavioural experiments created using PsychoPy to a secure server and collect data from participants remotely (Peirce et al., 2022).

It has been argued that the robustness of the data collected via online behavioural tasks is determined by the functioning of numerous interconnected technological systems. These can include: the server, which hosts the task; the internet service provider, which delivers the required task files from the server to the participant's computer; and the browser, which presents the stimuli to the participant and records their responses (Anwyl-Irvine et al., 2020). Subsequently, it has been suggested that any delay in the operating of these technologies can undermine the accuracy of both stimulus presentation times and recorded reaction times (Jia et al., 2018). For instance, when a task requires that new stimuli files be continuously downloaded from a server prior to their presentation, any lag in a participant's internet connection speed can cause alterations in the intended stimulus display times (Anwyl-Irvine et al., 2020). This was a particularly pertinent issue for the current research given that both the synchronisation-continuation task and cued RT task were dependent on precise stimulus onset timings and response time recordings. Therefore, this implies that the reliability of the current results was likely confounded by fluctuations in the efficiency of the computer, internet service provider, and browser used by each participant.

Contrary to the idea that results obtained through browser-based tasks are inherently unreliable due to high levels of variation in temporal precision, recent studies have reported precise stimuli and response timings across various software packages available for conducting online behavioural experiments (Anwyl-Irvine et al., 2021). For instance, Bridges et al. (2020) directly compared the precision at which five different software packages commonly used for online study implementation were able to present visual stimuli durations online. It was found that the majority of the tested packages had an inter-trial variability of less than 5ms, regardless of the browser or operating system that was used to run the task. Similarly, across all of the tested browser and operating system combinations, all five software packages demonstrated an inter-trial variability of under 10ms for recorded response times. Taken together, this suggests that online behavioural experiments demonstrate minimal temporal delays in their stimuli and response timings. In support of this idea, several past studies have reported comparable reaction times between lab-based and browser-based studies (Armitage & Eerola, 2020; Barnhoorn et al., 2015; Crump et al., 2013; Gould et al., 2015; Hilbig, 2016; Kim et al., 2019a; De Leeuw & Motz, 2015), even when online studies are completed within domestic settings (Miller et al., 2018). Ergo, this suggests that the robustness of the data obtained from the current online studies was unlikely to be compromised by any unintended temporal delays.

Aside from deviations in presentation and response timings, it has also been argued that the quality of the data collected via online behavioural experiments can be jeopardised by a lack of control over potential distractors (Sauter et al., 2020). Traditional lab-based experiments tend to be conducted within a standardised environment where potential distractors are minimised, thus facilitating participants' ability to attend to the task. Whereas, online tasks are usually completed within participants' own home or school where it is often not possible for the researcher to control, nor even observe, the influence of distractors on participants' attention (Kochari, 2019). Indeed, when examining data from 16 studies, Drody et al. (2023) found that participants frequently engaged in media-based multitasking whilst completing an online study within their home environment, such as watching a television or listening to music, with an average prevalence rate of 38%. Multi-tasking in this manner has previously been suggested to limit individuals' attention, and thereby, diminish their task performance (Aagaard, 2019). Taken together, this raises concern over the reliability of the current findings, as they may be more reflective of variation in the distractive nature of participants' surroundings, as opposed to the precision of their forward model.

In order to mitigate the suggested confounding effect of distractors on the quality of the collected data, several methods of promoting participants' attention were implemented in the current research. For example, in accordance with the recommendations outlined by Rhodes et al. (2020), each of the current tasks was framed as a game with age-appropriate animations and images used to maintain participants' focus on the task objective. In addition, text-based prompts were used to stimulate participants' attention following trials where no response was given, as suggested by Crump et al. (2013). Alternative options were also considered, such as requiring the researcher to remain present via a video call whilst participants completed the task in order to give encouragement and monitor their attention level (Forsberg et al., 2021). However, it was decided that the researcher's presence would have introduced additional technical and ethical challenges (Howlett, 2022), which ultimately rendered this approach unsuitable for the current research. Admittedly, it would have been beneficial to also include a series of attention-check questions at regular intervals during the tasks. In doing so, it would be possible to monitor variation in participants' attention and evaluate whether the methods used to boost participants' attention were effective (Peer et al., 2022).

Given the suggested issues in regard to variable timing accuracy and insufficient control over distractors, it may be queried as to why online behavioural experiments were selected for the current research over more traditional lab-based experiments. In truth, the initial research plan did not include any online behavioural experiments. In 2019, preparations had been made to run in-person behavioural experiments exclusively. However, in March 2020, the UK government announced restrictions on all non-essential contact in response to the Covid-19 pandemic, including the immediate suspension of all face-to-face data collection (Brown & Kirk-Wade, 2021). To adapt to this change, the planned offline behavioural experiments were reprogrammed into online behavioural studies (see the Covid-19 impact statement for more information about these changes). Strikingly, few prior studies have used online behavioural tasks to assess motor optimisation (e.g., Hammerschmidt et al., 2021), or even SoA directly (e.g., Garaizar et al., 2016; Vilaza et al., 2014). Therefore, whilst the current research had not originally intended to utilise online behavioural experiments, this unexpected change provided a novel opportunity to establish the domains in which the forward model, and thereby agency, can be reliably measured within an online context.

Brief Summary

In summary, it has been argued that the results obtained from online behavioural experiments can be undermined by unintended deviations in stimulus and response timings and a lack of control over distractors. However, recent studies have reported that lab-based and web-based tasks show comparable precision in stimuli display durations and response time recording. This suggests that online behavioural tasks feature only minimal temporal delays in their stimuli and response timings. Furthermore, steps were taken to mitigate the confounding impact of distractors on participants' task performance, such as gamifying the behavioural tasks. Ergo, the results gained through the current research are unlikely to have been confounded by timing inaccuracies or participant inattention.

Parent- and Self-Report Scales

Impulsivity refers to a tendency to act prematurely without prior consideration for the consequences of one's actions (Bakhshani, 2014). In line with this definition, previous learning studies have shown that individuals with higher levels of impulsivity are less likely to use knowledge of past action outcomes to guide their subsequent choices compared to those with lower impulsivity levels (Cáceres & San Martín, 2017; Franken et al., 2008; Hogarth et al., 2015; Lim et al., 2015). This suggests that high levels of impulsivity are associated with reduced outcome learning and less premeditated action, and thus, diminished use of an appropriate forward model. Notably, impulsivity has been suggested to interact with age; declining from childhood to adulthood (Forrest et al., 2019; Harden & Tucker-Drob, 2011). For this reason, the influence of impulsivity on participants' task performance was controlled for in each of the current empirical studies. To achieve this, parents of all participants aged 4-12 completed the Strengths and Weaknesses of Attention-Deficit/Hyperactivity-symptoms and Normal-behaviours rating scale. Similarly, the UPPS-P short-form was administered to all participants aged 13-25. Both of these measures will now be described and their inclusion in the current research will be justified.

The Strengths and Weaknesses of Attention-Deficit/Hyperactivity-symptoms and Normal-Behaviours (SWAN) Rating Scale

The SWAN rating scale is an 18-item parent-rated questionnaire designed to measure ADHD symptomology in individuals aged under 18 (Swanson et al., 2001). Parents of the child participants are asked to compare their child's tendency to perform certain behaviours over the past month to other children on a 7-point Likert scale. Examples of items include, "stay seated" and "listen when spoken to directly". The SWAN scale includes an inattention subscale and a hyperactivity/impulsivity subscale. The current research analysed data only from the hyperactivity/impulsivity subscale in order to quantify each child participant's trait impulsivity. Higher SWAN-hyperactivity/impulsivity subscale scores indicate greater impulsivity.

The SWAN-hyperactivity/impulsivity subscale was used to assess impulsivity in the current research because this measure has previously been shown to have moderate internal consistency (Cronbach's α = .7) and test-retest reliability (r = .66) when used with children (Lakes et al., 2011). Additionally, the SWAN-hyperactivity/impulsivity subscale has also been found to positively correlate with similar self-report (Lakes et al., 2011) and behavioural (Figueroa-Varela et al., 2010) measures of impulsivity in developmental studies, thereby indicating good convergent validity. Taken together, this suggests that the SWAN-hyperactivity/impulsivity subscale provides a reliable and valid measure of impulsivity in children, both of which qualify this measure as ideal for use with participants aged 4-12 in the current research.

In favour of complete transparency, it should be noted that the decision to adopt the SWAN-hyperactivity/impulsivity subscale as a measure of impulsivity was not made solely on the basis of the reported reliability and validity of this tool. On the contrary, this choice was also guided by convenience. Throughout the current research, all participants aged 4-12 were recruited via an annual online event (see the online recruitment subsection of this chapter for more detail). When registering their child for the event, parents and careers were asked to complete the SWAN rating scale by the event organisers. The parent-reported data was then shared with the researchers after the event. Therefore, alongside its suggested reliability and validity, the SWAN-hyperactivity/impulsivity subscale data was also selected as an index of impulsivity due to its availability during the collection of the behavioural task data.

The UPPS-P Short-Form

The UPPS-P short-form is a 20-item self-report questionnaire used to measure self-reported trait impulsivity (Cyders et al., 2014). The UPPS-P contains 5 subscales, including negative urgency, lack of perseverance, lack of premeditation, sensation seeking, and positive urgency. Negative urgency and positive urgency refer to a tendency to engage in impulsive behaviour when experiencing negative and positive emotions, respectively. Lack of premeditation indicates an individual's tendency to act without prior planning. Lack of perseverance is often defined as a tendency to leave tasks incomplete. Sensation seeking refers to an individual's propensity to engage in thrill-seeking behaviours (Cyders & Smith, 2008). Participants are instructed to indicate the extent to which they agree with each item on a 4-point Likert scale. Example items include, "I quite enjoy taking risks" and "When I am upset I often act without thinking". Higher UPPS-P scores show greater self-reported trait impulsivity.

Similar to the SWAN-hyperactivity/impulsivity subscale, the UPPS-P short-form was used as a measure of impulsivity within the current research because it has been shown to have good internal consistency and good test-retest reliability when administered with both adult (Dugré et al., 2019; Xue et al., 2017) and adolescent (Donati et al., 2021) samples. In addition, the UPPS-P short-form has also been reported to have good convergent validity, as scores from this scale have been found to positively correlate with scores from other self-report measures of impulsivity commonly used within past literature, such as the Barratt Impulsiveness Scale-11 (BIS-11; Xue et al., 2017). This suggests that the UPPS-P short-form is a reliable and valid means through which to quantify participants' impulsivity, and thus, suitable for the present research. Furthermore, when collecting self-report data online, it is imperative to limit the time required for participants to complete the administered scales as this can prevent high participant drop-out rates (Galesic, 2006;

Hoerger, 2010). Notably, the UPPS-P short-form has been shown to be just as effective in measuring impulsivity as the original UPPS-P, yet requires less time to complete (Lozano et al. 2018). Ergo, the short length of the UPPS-P short-form further reinforced the decision to implement this scale within the current research.

Brief Summary

To summarise, the SWAN-hyperactivity/impulsivity subscale and the UPPS-P short form were used to assess participants' trait impulsivity throughout this thesis. This controlled for the influence of variation in participants' impulsivity on the findings. These two scales were selected for the current research because they have demonstrated good reliability and validity within past research. They were also chosen because they are relatively quick to complete, thus reducing the chance of high participant drop-out rates.

Online Recruitment

Moving forward, the manner in which participants were recruited and given access to the online tasks will now be outlined. Following this, the benefits and challenges that were encountered when recruiting participants online for the current research will be discussed.

In general, the current research attempted to recruit participants between the ages of 4 - 25-years-old. To achieve this, participants aged 4-12 were recruited from Summer Scientist Month (SSM) 2020 and 2021, two online events hosted throughout August by the University of Nottingham. At the event, children were encouraged to play games designed by the researchers in the School of Psychology and learn about the brain and human behaviour. The event website contained links to the current tasks hosted on Pavlovia. Participants could simply click on any of the links to participate. Participants had the opportunity to receive one point in exchange for each study that they participated in. Points could be earned throughout the online event and attendees were encouraged to earn as many points as they could. Informed consent, demographic information and the SWAN rating scale data were obtained from parents or carers by the event organisers when registering their child for the event.

For all four studies, participants aged 13-17 were recruited from two high schools in the Nottinghamshire and Derbyshire areas from January 18th 2022 - 16th March 2022. Informed consent was first gained from the headteacher of the school. After this, informed parental consent was obtained via letters distributed by each school's Head of Psychology. Amongst school pupils whose parents had consented to their participation in the research, the current studies were advertised via a poster distributed by their psychology teacher. The poster contained a link to each behavioural task and information on what each experiment involved. Additionally, the poster also contained a link to an online Qualtrics form where participants provided their informed consent and demographic information and completed the UPPS-P short form. In line with recommendations from Mackenzie et al. (2021), school pupils were also provided with an animated video which explained what each task involved and how they could participate. The video was designed to improve participants' comprehension of what the research involved, and thus, facilitate their ability to provide their own informed consent to participate in the studies. For each study that participants took part in, they were given the opportunity to enter into a prize draw to win an Amazon voucher.

Finally, participants aged 18-25 were recruited either through the University of Nottingham School of Psychology's Research Participation Scheme (RPS) or through recruitment posters published on social media (see each empirical chapter for details of the specific recruitment dates for each study). The RPS website contained a separate link for each of the current studies, alongside information on what each experiment involved. The same links and information were also presented on the recruitment posters shared via social media. Each link first directed participants to an online Qualtrics form where they provided their informed consent and demographic information. They then completed the UPPS-P via the same online form, before being redirected to the task hosted on Pavlovia. For each study that participants took part in, they were given the opportunity to enter into a separate prize draw to win an Amazon voucher. In addition to entering a prize draw, participants recruited through RPS also had the option of receiving course credit for each of the studies that they volunteered to take part in. See each empirical chapter for additional information on how the participants within each sample were recruited.

It has been argued that one benefit of conducting online experiments, relative to offline studies, is that a greater number of participants can be recruited within a smaller time frame (Sauter et al., 2020). As the research materials are stored digitally on a server in an online study, multiple participants can all access the experiment, download the necessary materials onto their own computer, and complete the study simultaneously (Grootswagers, 2020). Whereas, in a traditional offline experiment, the research materials cannot be accessed directly by participants, meaning that the researcher must test each participant face-to-face, one at a time (Anwyl-Irvine et al., 2020). Therefore, it can be said that online studies are less time-consuming to conduct compared to offline studies.

The increased time-efficiency offered by online experiments was particularly valuable for the current research, as it meant the current research was easier for school educators to accommodate. As noted earlier within this chapter, the current research was conducted during the Covid-19 pandemic. As a result of the pandemic, all UK secondary schools were closed intermittently throughout 2020, and some of 2021, as part of government-imposed restrictions on face-to-face interaction (Brown & Kirk-Wade, 2021). This meant that school educators and pupils had to navigate online teaching and learning for the first time, leading to increases in perceived workload, stress and burnout for both groups (Beames et al., 2021; Commodari & La Rosa, 2021; Kim et al., 2022). Therefore, as the current tasks could be completed by multiple students at once, and at any time, this meant that the current research caused minimal disruption to planned educational activities (Gu et al., 2016), and thereby required relatively little additional work from educators. As a result of this increased flexibility, it can be argued that teachers were more willing to allow pupils to participate in the current research, compared to if the current research had been run offline.

In opposition to the suggestion that online studies are more appealing to school educators than offline experiments due to their increased flexibility, it can be argued that participating in online research holds less educational value for school pupils than in-person studies (Alibali & Nathan, 2010). When a study is run offline, pupils are able to observe the researcher as they administer the behavioural task. Subsequently, they gain the opportunity to learn, first-hand, how scientific research is conducted (Bartlett et al., 2017). Whereas, in an online study, the researcher is often absent when the task is completed. Hence, from the educator's perspective, online experiments offer less educational benefit for their pupils, relative to offline studies (Alibali & Nathan, 2010). For this reason, it can be suggested that educators would be less willing to support the recruitment of their pupils for an online study compared to an offline study. Therefore, in order to mitigate this issue and ensure that the current research was educational, the researchers collaborated with school educators to deliver webinars to school pupils explaining how the current studies were implemented and the significance of the findings for current literature.

In addition to offering greater time-efficiency than lab-based studies, it has been argued that the results acquired through online studies are more generalisable to the wider population than the data obtained via offline studies (Grootswagers, 2020). Given that recruitment is not limited by location or office hours, it has been argued that online studies are able to recruit a larger (Adjerid & Kelley, 2018) and more diverse range of participants (Casler et al., 2013) than is often viable through offline, lab-based experiments (Mason & Suri, 2012). Therefore, this heightens the probability that the current results are reflective of young people's genuine forward model development (Berinsky et al., 2012). In support of this idea, in all of the current studies, an adequate number of participants was recruited such that the performed analyses were sufficiently powered.

Aside from providing access to a greater and more diverse sample, it has been suggested that online developmental studies overcome some of the ethical concerns associated with offline developmental experiments (Barchard & Williams, 2008). For instance, online experiments offer participants a greater level of anonymity than is feasible within an offline study. Offline data collection often involves face-to-face interaction between the researcher and participant. Inevitably, some of the participant's identifying characteristics will be perceivable to the researcher during this interaction, such as their first name, facial features, and voice (Mackenzie et al., 2021). In contrast, online studies can be completed without any direct communication between the researcher and participants, meaning that none of these characteristics need to be shared (McCabe, 2004). In the context of developmental research, it has been suggested that increased anonymity can encourage parents to allow their child to complete an online study (Dworkin et al., 2016), thus enhancing participant recruitment. Although, this lack of direct communication can also present unique ethical challenges, particularly when recruiting young people as in the current studies, as the researcher cannot know whether participants were coerced into participation by a teacher, parent or other authority figure (Friedman et al., 2016).

In addition to greater anonymity, it has been argued that it is easier for younger participants to withdraw from an online study, relative to an offline study (Mackenzie et al., 2021). In an offline study, the researcher is often present whilst the child or adolescent completes the experiment. It has been suggested that the inherent imbalance in authority status between the young participant and the adult researcher can cause participants to feel obligated to complete the full task (Birnbaum, 2004). Whereas, as the researcher is often absent during an online experiment, it has been posited that child and adolescent participants will not experience this same social pressure and thus, will feel more comfortable exerting their right to withdraw from the study (Mackenzie et al., 2021). Whilst this increased ease through which participants can exit the study does make the research more ethical, it can also lead to greater rates of attrition (Barchard & Williams, 2008). Hence, the current tasks were gamified in order to increase participant engagement and prevent withdrawals due to boredom (Looyestyn et al., 2017), as achieved in past studies (Vilaza et al., 2014).

Brief Summary

In summary, participants were recruited for the current research via an online event, through local high schools, via the university's RPS scheme, and through social media. Notably, the use of online experiments was beneficial for the current research as it meant that data could be collected from a larger and more diverse sample at a rapid rate. Additionally, online studies also offer the participant greater anonymity than offline studies, which has been suggested to enhance participant recruitment. Evidently, online experiments are likely to encounter greater rates of attrition than offline studies. They can also be perceived as holding less educational value than offline studies by gatekeepers, which can restrict researchers' access to participants. Notably, these highlighted issues were addressed in the current research by gamifying experiments and collaborating with educators to deliver educational webinars.

Data Analysis Approach

The approach taken to data analysis throughout this thesis will now be outlined and justified. More specifically, the manner in which age was conceptualised will be addressed, in addition to the statistical tests and outlier detection methods employed.

The Use of Stepwise Multiple Linear Regression Analysis

As age was treated as a continuous variable throughout this thesis, stepwise multiple linear regression analysis was used to assess the influence of age on the dependent variables within each of the current studies. More specifically, in each regression, impulsivity and sex were entered in an initial block as nuisance variables and age was entered alone in a second block. The reason for including impulsivity as a nuisance variable was described above within the parent- and self-report measures subsection. In regard to the inclusion of sex as a nuisance variable, the rate at which structural and functional maturation occurs within the brain has been shown to vary between males and females (De Bellis et al., 2001; Koolschijn & Crone, 2013; Sumich et al., 2012). Hence, sex was also included as a nuisance variable in the current research. Additional evidence regarding the effect of impulsivity and sex on participants' performance on each of the current tasks can be found within each empirical chapter.

In a stepwise multiple linear regression analysis, predictor variables are added or subtracted from the final regression model in an iterative process. At each step, a predictor variable is added to the model if it independently explains a significant proportion of the variance in the dependent variable. The order in which predictor variables are considered for addition into the model is determined by a specified criterion, such as $p < .05$ (Armstrong & Hilton, 2010). In contrast to a stepwise regression, the hierarchical regression requires that the researcher specify the order to which predictor variables are added to the model according to a theory (Leech et al., 2003). Notably, in each analysis conducted throughout this thesis, there was no theoretical reason to believe that impulsivity would explain a greater proportion of the variance in the dependent variable than sex, or vice versa. Therefore, the stepwise multiple linear regression was believed to be most suitable for the current research compared to alternative linear regression techniques, such as hierarchical multiple linear regression.

Admittedly, it is not a given that forward model development should follow a linear-shaped development trajectory. In contrast, an alternative model, such as a phase-change or quadratic model, may have been a more suitable fit for the data across the current tasks. Subsequently, in order to verify that the forward model does, indeed, follow a linear development trajectory, one would need to compare how well participants' data fit to a linear model relative to alternative development models (i.e., linear, quadratic, phase-change, sigmoid). From this, it would then be possible to draw more concrete conclusions on the true trajectory at which the forward model develops throughout childhood, adolescence and young adulthood. Notably, this raises concern over the legitimacy of the current findings and hence, the results should be interpreted with caution.

The Conversion of Linear Age to Logarithmic Age

In line with past research that has focused on understanding developmental changes in the brain (i.e., Bethlehem et al., 2022), participants' chronological age was converted to natural logarithmic age across all four empirical studies. The theoretical basis for this adjustment was grounded in the idea that time itself is logarithmic (George, 2016). It may be argued that a greater developmental difference exists between a 4-year-old and an 8-year-old compared to a 21-year-old and a 25-year-old. Compared to adulthood, childhood is marked by significant advancements in the development of cognitive and motor skills (Brocki & Bohlin, 2004; Bolger et al., 2021). As the two pairs of individuals are 4 years apart in chronological age, expressing age on a traditional linear scale would not give an accurate representation of these differing developmental discrepancies. Whereas, adjusting age to a logarithmic scale provides an effective means for capturing this phenomenon (George, 2016). The difference between a 4-year-old and an 8-year-old on a logarithmic scale ($ln(8) - ln(4) = .69$) is larger than the difference between a 21-year-old and a 25-year-old on the same scale ($ln(25) - ln(21) = .17$). Therefore, comparing participants' task performance against their logarithmic age provides a more accurate representation of the developmental trajectory of the forward model compared to if a linear age scale were used.

The Implementation of Tukey's Fences Method for Outlier Detection

To exclude any potentially confounding anomalous data, the Tukey's fences method for outlier detection (Tukey, 1977) was applied to the data collected within each empirical chapter. Using the Tukey's fences method, any data points that are more than 1.5 interquartile ranges outside of the lower and upper quartiles are identified as outliers (Tukey, 1977). This method was chosen for the current research as the interquartile range is not distorted by the presence of large outliers, unlike alternative methods of outlier detection that are based on the mean and standard deviation, such as the two standard deviation rule (Schwertman et al., 2004). Therefore, it can be argued that the Tukey's fences method provides a more reliable method of identifying outliers compared to alternative techniques, such as the two standard deviation rule.

Brief Summary

To summarise, throughout the current research, age was converted to a logarithmic scale. Given that age was treated as a continuous variable, stepwise multiple linear regression analyses were employed throughout this thesis to explore the impact of age on participants' task performance. Stepwise regression analysis was chosen over alternatives, such as hierarchical regression analysis, because neither of the two nuisance variables were suspected to be a larger predictor of task performance relative to the other. Finally, Tukey's fences method for outlier detection was implemented across all four studies because this method is less likely to be influenced by large outliers compared to alternative outlier detection tests, such as the two standard deviation rule.

Conclusion

To conclude, the current thesis explored the influence of age on participants' prediction accuracy and action-outcome learning using four online behavioural experiments. Within these experiments, four tasks were employed: a synchronisation-continuation task, a cued RT task, a goal-switching task and an outcome learning task. These tasks were chosen due to their simplicity, as this meant that they were both suitable for use with young participants and easy to implement online. In addition to these behavioural tasks, the SWAN-hyperactivity/impulsivity subscale and the UPPS-P short form were used to assess participants' trait impulsivity, and thereby, control for the effect of impulsivity on participants' task performance. These two scales were selected because they have demonstrated good reliability and validity within past research. Notably, the use of online experiments meant that data could be collected from a larger and more diverse sample at a rapid rate. Furthermore, steps were taken to mitigate the confounding impact of distractors on participants' task performance, such as gamifying any behavioural tasks.

Chapter 3: An Investigation into the Influence of Age on Sensorimotor Continuation from Childhood to Adulthood

Chapter Summary

Chapter 3 details an experiment where the aim was to investigate how sensorimotor continuation changes with age from childhood to adulthood. To achieve this, participants aged 4-25 years completed a synchronisation-continuation task where they synchronised their keypresses with a series of isochronous tones played at either a high, medium or low frequency. Participants then continued making keypresses at the same pace after the tones were removed. It was found that the accuracy and consistency of participants' sensorimotor continuation improved with age. Crucially, sensorimotor continuation is believed to rely on an individual's capacity to i) predict when to respond using a forward model and ii) adjust their forward model in light of sensory feedback. Hence, the results suggest that the forward model becomes more functionally efficient as individuals mature from childhood to adulthood.

Introduction

Sensorimotor synchronisation refers to an individual's ability to align the timing of their own actions with the occurrence of an external stimulus (Schwartze et al., 2011). Whereas, sensorimotor continuation is defined as an individual's ability to maintain this entrainment to the presented tempo after the external stimulus has been removed (McPherson et al., 2018). Both abilities are believed to rely on a forward model system (Maes, 2016). The development of sensorimotor synchronisation has previously been investigated in past literature, with most studies concluding that this ability improves with age across childhood, adolescence and adulthood (e.g., Mu et al., 2018; Thompson et al., 2015; although see Drewing et al., 2006 for an alternative account). Conversely, the developmental trajectory of sensorimotor continuation across ontogeny has received relatively less attention in past research (e.g., McAuley et al., 2006). For this reason, and in the interest of brevity, the current study will focus exclusively on understanding the developmental trajectory of sensorimotor continuation and the conclusions that can be drawn from this regarding the maturation of the forward model.

Sensorimotor continuation has often been investigated using synchronisation-continuation tasks in past research (e.g., Schwartze et al., 2011). These tasks typically involve presenting participants with sequences of isochronous auditory tones, each with a varying inter-tone-interval. Each trial tends to have an initial synchronisation phase and a subsequent continuation phase. In the synchronisation phase, participants are asked to tap their finger in time with each of the tones. During the continuation phase, participants are then instructed to carry on tapping their finger at the same pace in absence of the tones. The accuracy and consistency with which the participant can maintain the response pace adopted during the synchronisation phase is then recorded as an index of their sensorimotor continuation skill (Repp & Su, 2013).

It has previously been argued that performance on the continuation phase of the synchronisation-continuation task is reliant on the forward model (Maes, 2016). On the continuation phase of the task, it has been suggested that participants need to use their forward model to predict the precise time at which to make their next response. This prediction is believed to be formed through a weighted combination of their prior and likelihood (Berniker & Körding, 2011). In this instance, the prior is believed to include their memory of the inter-tone-interval presented during the synchronisation phase (Lewis et al., 2004; Witt & Stevens, 2013). Whereas, the likelihood refers to their perception of the time elapsed since the tactile feedback from their last tap was perceived (Narain et al. 2018). Furthermore, after each response has been made, it has been argued that the comparator mechanism within the forward model assesses whether the time interval between their taps matches with the interval that they intended to produce, which should, in turn, align with their memory of the target inter-tone-interval. If the two are incongruent, and hence their prediction of when to act was erroneous, participants then need to adjust the timing of their subsequent responses to ensure that the target response pace is achieved (Maes, 2016). Therefore, an individual's sensorimotor continuation ability demonstrates the capacity of their forward model to both predict the outcome of their action, and thereby, when best to respond, and adjust the timing of any subsequent responses in light of observed sensory feedback.

In support of the role of the forward model in determining performance on the continuation phase of the synchronisation-continuation task, previous research has suggested that responses made during this phase can be shaped by prediction errors. Participants have been found to adapt their tapping pace in response to artificial perturbations in the timing of the sensory feedback from their taps (Maes et al., 2014). This suggests that the sensory feedback accrued from executing a response is indeed used to update the forward model and guide the time at which successive responses are made. Similarly, individuals with Tourette Syndrome, who have previously been shown to have an impaired ability to update their forward model (Kim et al., 2019b), have also been shown to perform worse than neurotypical controls on the continuation phase of the synchronisation-continuation task (Graziola et al., 2020). Taken together, this reinforces the idea that participants' sensorimotor continuation ability is supported by their forward model.

In addition to the accuracy and consistency of participants' tapping pace, some past studies have also included a temporal bias measure to quantify the extent to which the participant tended to respond at a pace that was ahead of or behind the target response pace (e.g., Claassen et al., 2013). This provides extra qualitative detail regarding the direction of participants' deviation from the set response pace. A tendency to respond ahead of the target response pace might indicate that the participant was prone to underestimating the temporal interval between the presented tones. Whereas, a tendency to respond behind the target response pace might provide evidence that the participant tended to overestimate these inter-tone intervals. Crucially, as the goal of the task is to align one's response pace with the target response pace and thereby, minimise prediction error, any deviation from the target pace indicates a less precise forward model, regardless of directionality.

Notably, evidence on how sensorimotor continuation changes with age is limited within past literature. Only one previous study has claimed to have examined the development of sensorimotor continuation across the full lifespan (McAuley et al., 2006). In research by McAuley et al. (2006), participants aged 4-95 were asked to tap in time with a series of isochronous tones and then continue tapping at the same pace when the tones were removed. It was found that participants' ability to accurately and consistently reproduce the presented tone sequence followed a quadratic pattern with performance improving with age throughout childhood, peaking in adulthood, before declining as individuals transitioned into old age. This suggests that sensorimotor continuation improves with age from childhood to adulthood.

In criticism of the study conducted by McAuley et al. (2006), it should be noted that participants' ages only spanned between 4-12 and 18-95, with no participants aged between 13-17. Therefore, by concluding that an age-related improvement in sensorimotor continuation occurs from childhood to adulthood, the researchers appear to assume that a linear development of this ability must take place between the ages of 12 and 18. Yet, this assertion cannot be conclusively supported by their findings, as they do not have any data for the adolescent development period. Consequently, this raises doubt over the idea of a continuous progression in sensorimotor continuation across adolescence.

Contrary to the idea that sensorimotor continuation improves with age throughout the transition from childhood to adulthood, alternative research has suggested that this ability is already fully developed by adolescence. Witt and Stevens (2013) instructed participants aged 12-43 to complete a synchronisation-continuation task with a target finger tapping pace of 0.75Hz. No difference was found between adolescents and adults in their ability to accurately maintain the target response pace. This suggests that sensorimotor continuation, and thus, the forward model, is already sufficiently developed by adolescence.

In opposition to the conclusions drawn by Witt and Stevens (2013), it may be argued that the reported lack of an age-related change in sensorimotor continuation can be attributed to statistical error. The analysis performed by Witt and Stevens (2013) appears to have been underpowered. According to G*Power software (Faul et al., 2009), at least 72 participants would have been required to find a medium-sized effect with 80% power and a 5% alpha level in the mixed factorial ANOVA. This suggests that the absence of an age-related difference in sensorimotor continuation occurred due to statistical error, as opposed to a genuine lack of any developmental changes in sensorimotor continuation. Moreover, this suggests that the developmental trajectory of sensorimotor continuation requires further investigation.

The Current Study

The current research aimed to examine the beneficial influence of increased age on sensorimotor continuation across childhood to adulthood, as suggested by McAuley et al. (2006), against the contradictory findings from research by Witt and Stevens (2013). This would determine the influence of age on participants' ability to use the forward model to accurately predict action outcomes and update the forward model in response to observed sensory feedback. In the current experiment, participants aged 4-25 years were presented with a synchronisation-continuation task where they were instructed to make a series of keypresses in time with an external pacing stimulus. They then had to continue making keypresses at the same pace in absence of the pacing stimulus. The accuracy and consistency with which participants could maintain the set response pace was then calculated. In addition, the extent to which participants' response pace lagged behind or ran ahead of the target response pace was also recorded. In order to rectify the criticisms raised against research by Witt and Stevens (2013), a power analysis was conducted prior to participant recruitment to ensure that all statistical tests were sufficiently powered. To test the findings of Witt and Stevens (2013), three hypotheses were formed:

1. It was hypothesised that participants' temporal accuracy would be predicted by age at all tone frequencies, with a higher temporal accuracy associated with older age.

2. It was hypothesised that participants' temporal variability would be predicted by age at all tone frequencies, with lower temporal variability associated with older age.

Method

Design

The current study used a mixed factorial design. The between-subjects independent variables were age, sex and impulsivity. The within-subjects independent variable was the frequency at which the tones were presented during each condition of the synchronisation-continuation task. The key dependent variables were participants' average temporal error, temporal variability, and temporal bias on the synchronisation-continuation task. These dependent variables acted as indicators of participants' ability to predict and adjust the timing of their movements; skills which necessitate the use of a forward model.

Participants

212 participants were initially recruited (44 male, 168 female). The age of participants ranged from 4 to 25 years (*M*=16.86, *SD*=4.2).

- 156 participants were right-handed, 16 were left-handed and 35 were ambidextrous, as measured by the Edinburgh Handedness Inventory – Short Form (EHI-SF; see appendix A for a full outline of this measure). 5 participants did not complete the EHI-SF.
- 150 participants were White, 12 was Asian, 6 were Black, 5 had mixed/multiple ethnic identities, 2 had another ethnic identity that was not listed, and 37 did not report their ethnic identity.

Participants were recruited through one of four avenues: 37 participants were recruited via the SSM event in August 2020 and August 2021, 23 were recruited from two high schools in the Nottinghamshire and Derbyshire areas from January 18th 2022 - 16th March 2022, and the remaining 152 were recruited either through the School of Psychology's RPS system or through recruitment posters published on social media between 18th November 2020 - May 15th 2021. For more detailed information on how participants were recruited from each of these sources and how informed consent was obtained, please see chapter 2.

The full experimental procedure of the current study was approved by the School of Psychology ethics committee at the University of Nottingham. Seven participants were excluded because they self-reported a diagnosis of either Attention Deficit Hyperactivity Disorder (ADHD) or Autism Spectrum Disorder (ASD). The decision to remove these participants was made on the basis that (i) the goal of the current study was to establish the neurotypical development trajectory of the forward model, and (ii) it has been previously demonstrated that individuals with these conditions tend to show impaired action planning (Gvirts Probolovski & Dahan, 2021; Haswell et al., 2009). More specifically, past research has shown that autistic individuals tend to attribute disproportionate weight to the prior, relative to the likelihood (Van de Cruys et al., 2014), leading to erroneous predictions regarding the sensory consequences of their actions (van Laarhoven et al. 2019). In contrast, individuals with ADHD tend to give excessive weight to the likelihood, rather than the prior (Gonzalez-Gadea et al., 2015), which again has been found to result in reduced anticipation of their action consequences (Marzinzik et al., 2012). Therefore, the inclusion of participants with either of these conditions would have obscured the results regarding the neurotypical developmental trajectory of the forward model, and hence these participants were removed from the sample. All remaining participants were neurotypical. The demographics of the adjusted sample can be viewed in table 3.1.

Table 3.1. The demographic characteristics of the adjusted sample.

	Age (*years*)	Sex	Ethnicity*	Handedness**
Full sample (*n*=205)	*Range*= 4 to 25 *M*=16.81 *SD*=4.26	44 Male 161 Female	143 White 12 Asian 6 Black 5 Mixed/multiple ethnic identities 2 Any other ethnic identity	150 right-handed 16 left-handed 34 ambidextrous

Note. *Ethnicity information was not collected for 34 participants. **Handedness information was not collected for 5 participants.

Materials

The Synchronisation-Continuation Task. The synchronisation-continuation task was used to measure participants' ability to plan and adjust the timing of their movements. The task was designed using PsychoPy and ran online via Pavlovia (Peirce et al., 2019, 2022). Auditory stimuli included three sequences of 20 tones. The frequency at which tones were delivered on each sequence was either 1.25Hz (low), 1.6Hz (medium), or 2.5Hz (high). Accordingly, the time interval between each tone was either 400ms, 600ms, or 800ms. Past research has noted that children tend to have a shorter preferred motor tempo than adults during unpaced finger-tapping tasks (Hammerschmidt et al., 2021). Notably, the closer a prescribed response pace is to an individual's motor tempo, the easier it is for the individual to synchronise their movements with the target response pace (Monier & Droit-Volet, 2019). Therefore, the frequency of the tone sequences was varied in order to control for the effect of age-related differences in participants' preferred motor tempo on their ability to align their movements with the timing of the tones.

Additionally, previous research has suggested that children struggle to match their actions with tones delivered at frequencies faster than 400ms as a result of underdeveloped motor coordination (Repp, 2005). Similarly, it has been suggested that data obtained from tasks with tone frequencies slower than 800ms fail to maintain young children's attention (Faria et al., 2017). Therefore, with the child participants in mind, tone frequencies between 400-800ms were chosen for this task. The pitch of the 10^{th} tone in each sequence was higher than the other tones. This high-pitched tone acted as a cue, prompting participants to begin responding. Visual stimuli on the child version of the task included an image of a barn door and 9 images of cartoon farm animals. Whereas, an image of a wooden door and comical photographs of animals in costumes were included in the version of the task designed for adolescents and young adults. In both cases, the visual stimuli were chosen to create an age-appropriate narrative and maintain participants' attention.

Self-Report and Parent-Report Measures. The following self-report measures and parent-report measures were also administered in the current study:

- The SWAN-Hyperactive/Impulsive subscale was used to measure child participants' impulsivity as reported by their parents or carers.

- The UPPS-P short-form was used to measure adolescents' and young adults' self-reported impulsivity.

Please see chapter 2 for a full outline of the SWAN rating scale and UPPS-P short-form.

Procedure

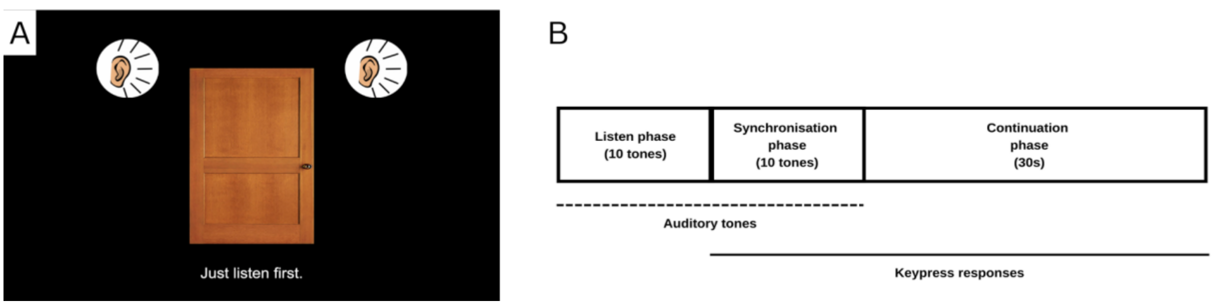
Two versions of the synchronisation-continuation task were created. One version of the task was designed with adolescents and young adults (aged approximately 13-25) in mind. This version of the task was available to participants via the two high schools, RPS, and social media posts. The second version of the task was designed for children (aged 4-12) and was administered to participants at the two SSM events. As the two task versions were created with either adolescents and young adults or children in mind, the procedures through which the two task versions were given to participants will be referred to as the adolescent and young adult procedure and the child procedure, respectfully. That being said, it should be noted that, although SSM was aimed at children aged 4-12, individuals aged between 12-17 were also permitted to take part in the event. Therefore, it was possible for a participant aged >12 to have received the child version of the task.

Adolescent and Young Adult Procedure. Participants recruited from a high school, via RPS or through social media provided their demographic information and completed the UPPS-P short form via an online survey hosted on Qualtrics before being redirected to the synchronisation-continuation task.

Upon opening the adolescent and young adult version of the synchronisation-continuation task, participants first saw a black instructions screen with details on how to complete the task. Participants were instructed to complete the task using the same hand throughout. Participants started the trials by pressing the spacebar. Participants first completed 3 practice trials to familiarise themselves with the task, before progressing to the main experiment trials. In each trial, participants were presented with a cartoon wooden door in the centre of the screen (see figure 3.1A). Each trial had three phases, a listen phase, a synchronisation phase, and a continuation phase (see figure 3.1B). After a 0.5s delay, the listen phase began and a sequence of tones was heard. In this phase, participants were instructed to listen to the first 10 tones in the sequence. The purpose of the listening phase was to give participants an opportunity to familiarise themselves with the sequence of tones. After the 10^{th} tone, the synchronisation phase began. During this phase, participants responded to the trial by pressing the space bar in time with the tones. The synchronisation phase allowed participants the chance to practice pressing the keyboard in time with the tones. After a further 10 tones, the sequence ended and the continuation phase began. In the continuation phase, participants were required to continue pressing the spacebar at the same pace. The time of each spacebar keypress made during the continuation phase was recorded. Text and additional images onscreen reminded participants when to listen, respond and then continue responding in the absence of sound. After 30s, the trial ended. If no response was made during a trial, participants were reminded of the task instructions via text presented onscreen for 2s. Each trial lasted approximately 50s.

Figure 3.1.

Image of the Synchronisation-Continuation Task and a visual representation of the trial structure.



Note. 3.1A. An image showing the listen phase of the synchronisation-continuation task. An image of a door was presented in the centre of the screen. Text displayed at the bottom of the screen and two images positioned either side of the door served as reminders to the participant that their current objective was to listen to the tones. 3.1B. A visual representation of the trial structure. The listen phase lasted for 10 tones. This was then followed by a synchronisation phase that lasted for 10 tones. The synchronisation phase was then succeeded by a continuation phase that lasted for 30s. Auditory tones were heard throughout the listen and synchronisation phases. Keypress responses were only permitted during the synchronisation and continuation phases.

Three conditions were presented during the task: a low-frequency condition, a medium-frequency condition, and a high-frequency condition. The frequency at which the tones were delivered differed per condition: 1.25Hz (low-frequency), 1.6Hz (medium-frequency), and 2.5Hz (high-frequency). A reward screen was presented for 3s after each trial. The reward screen showed a comical photograph of an animal in a costume and text praising the participant's performance. Images of costumed animals and a wooden door were used to create the narrative that participants were playing the role of a detective who had been tasked with locating a thief hidden within an apartment building. The theme of a mystery game was adopted to retain participants' attention throughout the task and increase participant recruitment. Break screens were shown for an unlimited time every 3 trials. Break screens included encouraging text to further promote participants' attention. An inter-trial interval was presented for 1s, during which a black screen with a white fixation cross was displayed. The task consisted of 12 trials that ran in a pseudo-random order, 4 for each condition. The full procedure lasted approximately 15-minutes.

Child Procedure. The child procedure was identical to the adolescent and young adult procedure with a small number of modifications. Participants' parents or carers provided their child's demographic information and completed the SWAN scale when registering their child for SSM. Once registered, participants could complete the synchronisation-continuation task at any time throughout the SSM event. Participants were encouraged to complete the left-hand vs right-hand task to measure their hand preference before completing the synchronisation-continuation task and received one point in exchange for doing so (see appendix B for a full description of the left-hand vs right-hand task).

The child version of the synchronisation-continuation task closely mirrored the adolescent and young adult version of the task with a few minor changes. Participants were presented with an image of a cartoon barn door in the centre of the screen, as opposed to the wooden door that was viewed by adolescent and young adult participants. Similarly, instead of a costumed animal, each reward screen showed a cartoon image of a sleeping farm animal. The visual stimuli were changed between the two versions of the task in order to support a new narrative; rather than searching for a thief, participants were now tasked with waking sleeping farm animals. The narrative was changed as it was believed that the concept of waking sleeping animals was more simplistic and hence, more age-appropriate. A second reward screen was shown after each trial for an unlimited time. The second reward screen informed participants that they had earned a piece of a picture and the full picture would be revealed at the end of the task. This picture-based reward was implemented as an additional means of maintaining participants' attention throughout the task. The second reward screen was not present in the adolescent and young adult version of the task as it was believed that additional reward screens were not necessary in order to maintain the attention of participants over 12-years-old. The full child procedure lasted approximately 10-minutes.

Data Analysis

Data from the self-report measures were pre-processed using MATLAB. Whereas, data obtained via the synchronisation-continuation task was pre-processed in Python using the Spyder IDE. All participants' ages were converted to natural logarithmic age, as described in chapter 2.

Self-Report Scales. All self-report and parent-report scales were summed and averaged to create an index for each of the variables of interest. In addition, SWAN-Hyperactive/impulsive subscale scores and UPPS-P short-form scores were converted to z-scores. This meant that the impact of impulsivity on task performance could be investigated across age, despite the fact that impulsivity was measured using the SWAN-Hyperactivity/impulsivity subscale for participants recruited at SSM and the UPPS-P short-form for the remaining participants.

Erroneous Data Removal. For each trial of the synchronisation-continuation task, the time intervals between each spacebar keypress were calculated. These were referred to as inter-tap-intervals (ITIs), consistent with previous finger-tapping tasks (e.g., Maes, 2016). All ITIs <200ms or >1s were removed to control for accidental keypresses and participant inattention. These cut-off values were chosen as they removed values 200ms below the lowest target response frequency (400ms) and 200ms above the highest target response frequency (800ms). On average across the participants, 2.46% of ITIs per trial fell outside of the cut-off values and were removed from further analysis (*SD* = 3.59, *range* = 0 - 34.5).

Temporal Error. For each trial, the target ITI for the condition was subtracted from each ITI to produce a set of temporal error (TE) values. To clarify, 400ms was subtracted from each ITI in a high-frequency condition trial, 600ms was subtracted from each ITI in a medium-frequency condition trial, and 800ms was subtracted from each ITI in a low-frequency condition trial. A TE value indicates the level of discrepancy between the time gap separating two of the participant's successive keypresses and the time gap between two successive tones in the presented sequence. A negative TE indicated that the participant's ITI was shorter than the target ITI, and therefore, the two successive keypresses were too close together in time. Whereas, a positive TE revealed that the participant's ITI exceeded the target ITI, i.e., the two successive keypresses were too far apart in time.

Temporal Bias. To explore the extent to which participants responded ahead of, or behind, the target response pace, the participant's TE values were first pooled for each condition. A median TE value was then calculated for each condition from the pooled TE values. This was referred to as participants' temporal bias. A negative temporal bias suggests that the participant tended to respond ahead of the set response pace. Whereas, a positive temporal bias suggests that the participant tended to respond behind the set response pace. This measure provided qualitative detail regarding the direction at which participants tended to deviate from the target response pace.

Average Temporal Error. To investigate the magnitude with which a participant deviated their response pace from the pace set by the tones, the mean TE was obtained for each condition. To achieve this, all TE values were converted to their absolute values. This step was necessary as a negative TE could cancel out the influence of a positive TE when all TEs are averaged, which would obscure the participant's true average TE. Next, the absolute TE values were pooled by condition and averaged to produce a mean TE and standard deviation for each of the three conditions. The mean TE values gave an indication of the participants' temporal accuracy on each condition. Larger mean TE values demonstrate poorer temporal accuracy and thus, a less developed ability to both predict when to respond and update their pace when errors in their timing arise.

Temporal Variability. Finally, to determine the extent to which a participant could consistently maintain the response pace set by the tones, a coefficient of variation (CV) was calculated for each condition. This was achieved by dividing the standard deviation of the TE by the mean TE for each condition. Larger CV values showed greater variability in the participant's ability to match the set response pace. Similar to low temporal accuracy, a more variable response pace was indicative of an inferior ability to plan and adjust one's response pace.

Outlier Detection. To exclude any potentially confounding anomalous data, the Tukey's fences method for outlier detection was applied to the dataset (Tukey, 1977). See chapter 2 for an explanation of why the Tukey's fences method for outlier detection was chosen over alternative outlier detection methods.

Seventeen high-frequency median TE, 9 medium-frequency median TE, 12 low-frequency median TE, 4 high-frequency mean TE, 7 medium-frequency mean TE, 14 low-frequency mean TE, 9 high-frequency CV, 1 medium-frequency CV, and 10 low-frequency CV data points were found to be more than 1.5 interquartile ranges away from the nearest quartile. Upon comparison, it was found that the removal of the anomalous data points did not affect the direction or significance of the results regarding the influence of age, sex and impulsivity on the medium-frequency median TE, the low-frequency median TE, the high-frequency CV, the medium-frequency CV, or the low-frequency CV. Hence, the identified data points for these variables were not removed in order to ensure the completeness of the data. However, excluding the extraneous data points for the high-frequency median TE and the high-frequency mean TE did have an effect on the findings. Prior to outlier removal, high-frequency median TE was significantly predicted by age (β = .24, t = 3.5, p = .001) and high-frequency mean TE was not predicted by any predictor variable. Whereas, age significantly predicted high-frequency mean TE after the outlier data points were excluded (β = -.17, t = -2.46, p = .02) and no longer predicted a significant proportion of the variation in high-frequency median TE. As a result of these changes, the 21 data points identified as outliers for these two variables were removed from the data so as not to statistically bias the results.

Statistical Analyses. All statistical analyses were run in Statistical Package for the Social Sciences (SPSS). For the reasons outlined in chapter 2, impulsivity and sex were included as nuisance variables in the current study. In further support of the inclusion of impulsivity as a nuisance variable, previous research has shown that individuals with increased impulsivity demonstrate greater variability in their capacity to maintain a set response pace on synchronisation-continuation tasks compared to those with lower levels of impulsivity (Barratt et al., 1981; Noreika et al., 2013; Valera et al., 2010). Therefore, this suggests that participants' sensorimotor continuation could have been confounded by age-related differences in their levels of impulsivity. Likewise, in the case of sex, lower response pace variability on synchronisation-continuation tasks has previously been associated with greater functional efficiency in the brain (De Guio et al., 2012). Therefore, due to the suggestion that the brain matures at different rates in males and females (De Bellis et al., 2001; Koolschijn & Crone, 2013; Sumich et al., 2012), sex was also included as a nuisance variable in the current study.

Nine stepwise multiple linear regressions were performed to investigate whether participants' average temporal error, temporal variability, or temporal bias could be predicted by age, impulsivity, or sex in each of the three conditions. In each regression analysis, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block. The purpose of these statistical tests was to reveal the extent to which an individual's age can influence their predictive motor control when the respective influences of both impulsivity and sex are taken into account. G*Power analysis revealed that 77 participants were required to obtain a medium sized-effect (f^2=.15) in a multiple linear regression with three predictor variables, 80% power, and a 5% alpha level (Faul et al., 2009). As the sample contained 205 participants, the analyses were sufficiently powered.

Results

In the current study, participants completed 12 trials of a synchronisation-continuation task. To briefly reiterate, on each trial of the task, participants were instructed to listen to, press the spacebar in time with, and finally, replicate a sequence of tones presented at either a high-, medium-, or low-frequency. When replicating a tone sequence, participants' objective was to press the spacebar in a manner such that the temporal gaps between each of their keypresses mirrored the gaps present between the tones. The level of disparity between each produced time interval relative to the genuine time intervals present in the target sequence was calculated. This revealed the accuracy and consistency at which participants could maintain the set response pace, as well as their tendency to respond ahead of or behind the set pace.

The Influence of Age on Participants' Temporal Bias in Each of the Three Tone Frequency Conditions

To explore the extent to which a participant tended to respond ahead of, or behind, the target response pace, their median TE value was obtained for each condition. This revealed the participant's temporal bias; an indicator of the directionality of participants' average deviations from the target response pace. A negative temporal bias suggested that the participant tended to respond ahead of the set response pace, i.e., press the spacebar at too quick a pace. Whereas, a positive temporal bias suggested that the participant tended to respond behind the set response pace, i.e., press the spacebar at too slow a pace.

To investigate the influence of age on participants' temporal bias, three stepwise multiple linear regressions were conducted on the high-frequency median TE, the medium-frequency median TE and the low-frequency median TE, with the factors: age, impulsivity, and sex. In each regression, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block (see table 3.2 for the final models). It was revealed that both the medium-frequency median TE and the low-frequency median TE were significantly predicted by age (both $p<.001$), and were not predicted by impulsivity or sex (all $p>.05$). As age increased, both the medium-frequency median TE and the low-frequency median TE increased. Whereas, the high-frequency median TE was not predicted by age, impulsivity or sex (all $p> .05$). To further visualise the relationship between temporal bias and age, see figure 3.2 for each median TE variable plotted against participants' unlogged age in years.

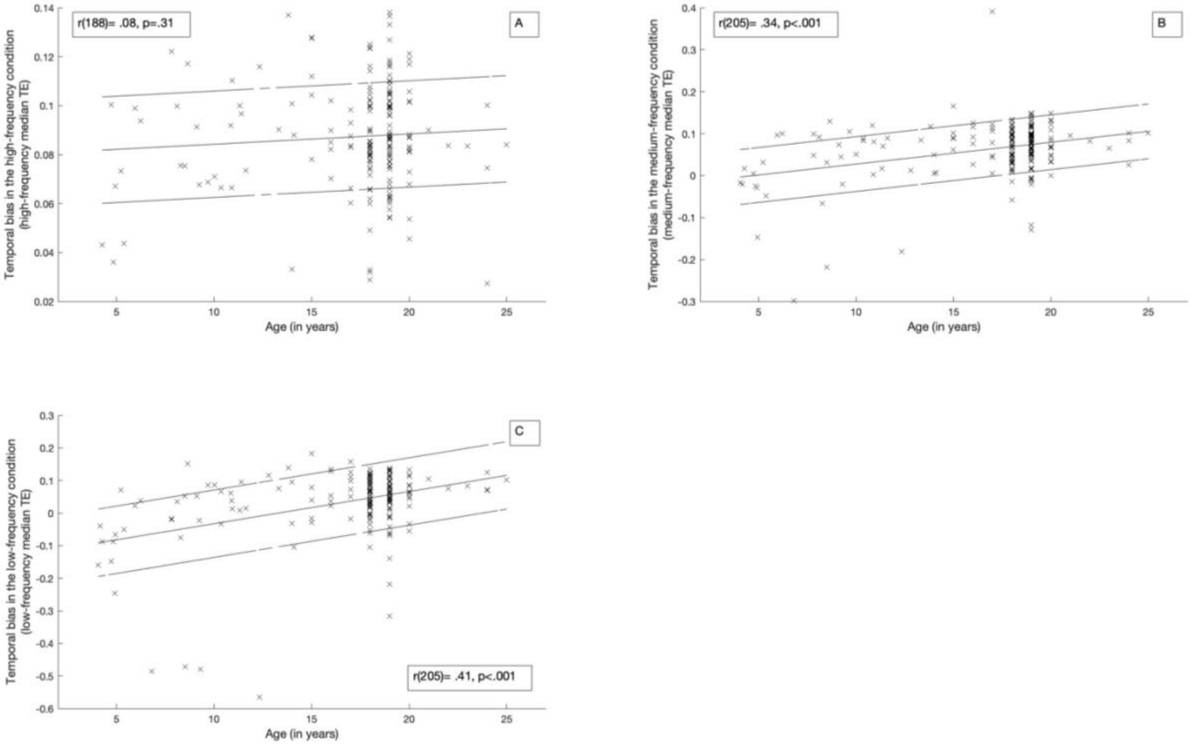
Table 3.2. The results of three stepwise multiple linear regression analyses investigating the influence of age, impulsivity and sex on the temporal bias in the high-, medium- and low-frequency conditions

Final regression model	*β*	*t*	*F*	*df*	R^2	*SE*
Temporal bias in the high frequency condition (high-frequency median TE)						
Age			No variables were entered into the model.			
Impulsivity						
Sex						
Temporal bias in the medium frequency condition (medium-frequency median TE)			28.06**	1, 202	.12	.06
Age	.35**	5.3**				
Impulsivity	.04	.6				
Sex	-.06	-.86				
Temporal bias in the low frequency condition (low-frequency median TE)			37.22**	1, 202	.16	.09
Age	.39**	6.1**				
Impulsivity	-.03	-.49				
Sex	-.02	-.31				

Note. *p<.05 **p<.001.

Figure 3.2.

Temporal bias as a function of participants' unlogged age in years and condition



Note. 3.2A. A figure showing temporal bias in the high-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 3.2B. A figure showing temporal bias in the medium-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 3.2C. A figure showing temporal bias in the low-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

The Influence of Age on Participants' Average Temporal Error in Each of the Three Tone Frequency Condition

In addition to the median TE, each participants' average TE was also calculated for each condition. This revealed the magnitude with which a participant deviated their response pace from the pace set by the tones. An average TE was obtained for each condition by averaging the absolute TE values from each trial within that condition. Mean TE values provide an indication of the participants' temporal accuracy on each condition with larger mean TE values demonstrating poorer temporal accuracy and thus, a less developed ability to both predict when to respond via the forward model and update their pace when errors in their timing arise. To investigate the extent to which age had an effect on participants' temporal error, three stepwise multiple linear regressions were conducted on the high-frequency mean TE, the medium-frequency mean TE and the low-frequency mean TE, with the factors: age, impulsivity, and sex. In each regression, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block (see table 3.3 for the final models). It was revealed that all three dependent variables were significantly predicted by age (all $p<.05$), and were not predicted by impulsivity or sex (all $p>.05$). As age increased, mean TE decreased in all three conditions. For visualisation purposes, see figure 3.3 for each mean TE variable plotted against participants' unlogged age in years.

It should be noted that, initially, a significant model was found for the medium-frequency mean TE before age was added to the model, $F(1, 202)= 4.92$, $p = .03$. Medium-frequency mean TE was significantly predicted by sex (β = -.15, t = -2.22, p = .03) and was not predicted by impulsivity (β = .06, t = .82, p = .42). The model fit was R^2= .02, SE = .05. However, after age was added to the model, the medium-frequency mean TE was significantly predicted by age (β = -.21, t = -3.03, p = .003), was not predicted by impulsivity (β = .06, t = .9, p = .37) and was no longer predicted by sex (β = -.11, t = -1.54, p = .13). Likewise, a significant model was also initially found for the low-frequency mean TE before age was added to the model, $F(1, 202)= 5.97$, p = .02. Low-frequency mean TE was significantly predicted by sex (β = -.17, t = -2.44, p = .02) and was not predicted by impulsivity (β = .07, t = 1.04, p = .3). The model fit was R^2 = .03, SE = .06. However, after age was added to the model, the low-frequency mean TE was significantly predicted by age (β = -.36, t = -5.39, p < .001), was not predicted by impulsivity (β = .08, t = 1.23, p = .22) and was no longer predicted by sex (β = -.09, t = -1.36, p = .18).

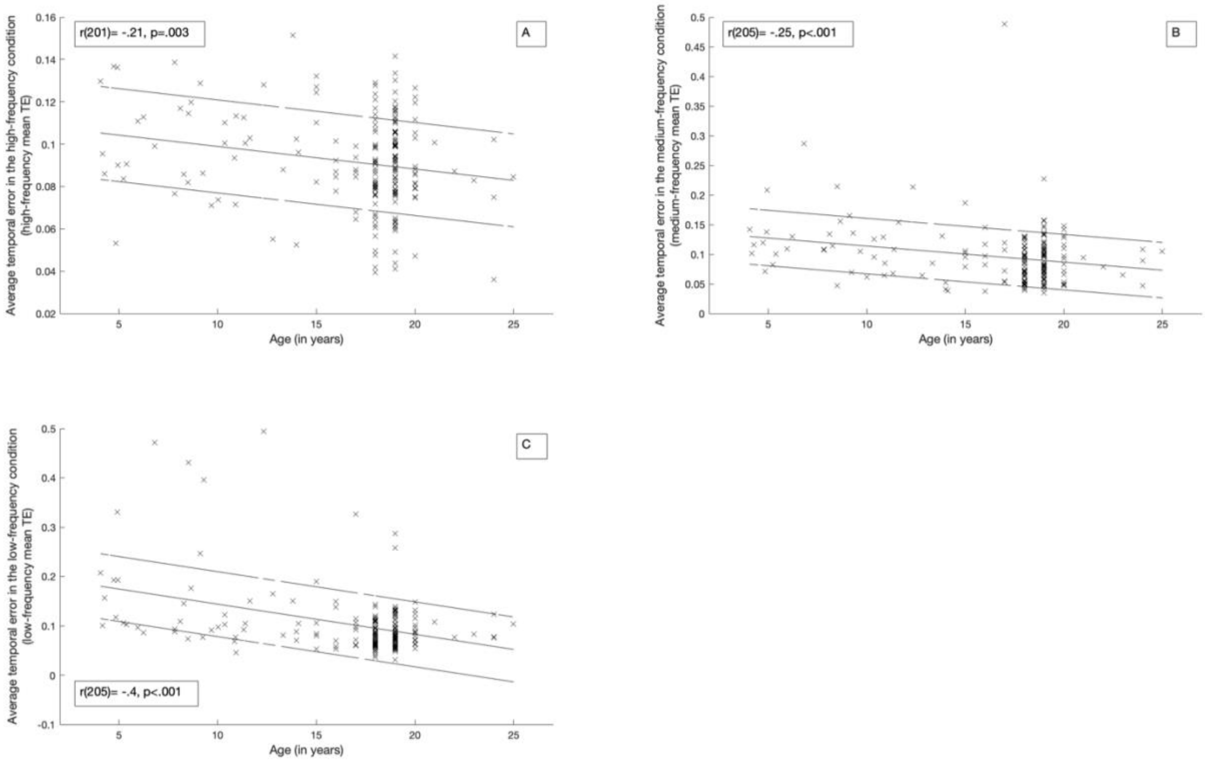
Table 3.3. The results of three stepwise multiple linear regression analyses investigating the influence of age, impulsivity and sex on the average temporal error in the high-, medium- and low-frequency conditions

Final regression model	*β*	*t*	*F*	*df*	R^2	*SE*
Average temporal error in the high frequency condition (high-frequency mean TE)			6.06*	1, 198	.03	.02
Age	-.17*	-2.46*				
Impulsivity	.06	.86				
Sex	-.03	-.35				
Average temporal error in the medium frequency condition (medium-frequency mean TE)			7.16*	2, 201	.07	.05
Age	-.21*	-3.03*				
Impulsivity	.06	.9				
Sex	-.11	-1.54				
Average temporal error in the low frequency condition (low-frequency mean TE)			17.93**	2, 201	.15	.06
Age	-.36**	-5.39**				
Impulsivity	.08	1.23				
Sex	-.09	-1.36				

Note. *p<.05 **p<.001.

Figure 3.3.

Average temporal error as a function of participants' unlogged age in years and condition



Note. 3.3A. A figure showing average temporal error in the high-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 3.3B. A figure showing average temporal error in the medium-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 3.3C. A figure showing average temporal error in the low-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

The Influence of Age on Participants' Temporal Variability in Each of the Three Tone Frequency Conditions

To explore the extent to which a participant could consistently maintain the response pace set by the tones, a CV was calculated for each condition. The larger the CV value, the greater the variability in participants' ability to match the set response pace, and hence, the less precise their forward model. To establish the extent to which age had an effect on the temporal variability in participants' response behaviour, three stepwise multiple linear regressions were conducted on the high-frequency CV, the medium-frequency CV and the low-frequency CV, with the factors: age, impulsivity, and sex. In each regression, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block (see table 3.4 for the final models). It was revealed that all three variables were significantly predicted by age (all $p<.05$), and were not predicted by impulsivity or sex (all $p>.05$). As age increased, the CV decreased in all three conditions. For visualisation purposes, see figure 3.4 for each CV variable plotted against participants' unlogged age in years.

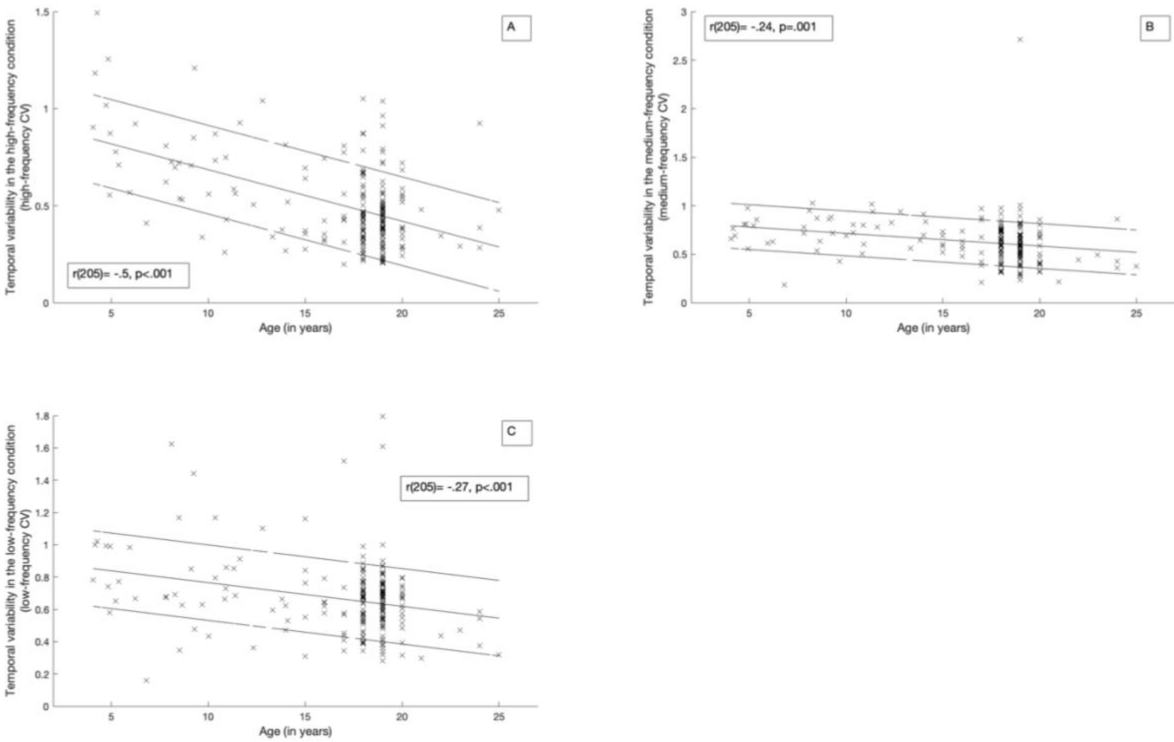
Table 3.4. The results of three stepwise multiple linear regression analyses investigating the influence of age, impulsivity and sex on the temporal variability in the high-, medium- and low-frequency conditions

Final regression model	*β*	*t*	*F*	*df*	R^2	*SE*
Temporal variability in the high frequency condition (high-frequency CV)			72.85**	1, 202	.27	.2
Age	-.52**	-8.54**				
Impulsivity	.02	.24				
Sex	.07	1.05				
Temporal variability in the medium frequency condition (medium-frequency CV)			10.66*	1, 202	.05	.23
Age	-.22*	-3.27*				
Impulsivity	.01	.19				
Sex	-.02	-.26				
Temporal variability in the low frequency condition (low-frequency CV)			14.15**	1, 202	.06	.23
Age	-.26**	-3.76**				
Impulsivity	.02	.31				
Sex	.02	.33				

Note. *p<.05 **p<.001.

Figure 3.4.

Temporal variability as a function of participants' unlogged age in years and condition



Note. 3.4A. A figure showing temporal variability in the high-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 3.4B. A figure showing temporal variability in the medium-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 3.4C. A figure showing temporal variability in the low-frequency condition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

Discussion

The purpose of the current study was to determine the influence of age on the accuracy and variability of participants' sensorimotor continuation ability, given its proposed reliance on the forward model (Maes, 2016). To achieve this, participants aged 4-25 completed a synchronisation-continuation task in which they were instructed to press a computer key in time with a series of tones presented at either a high, medium, or low-frequency. After synchronising their responses with the tones, they then had to continue making keypresses at the same pace in the absence of the tones. The accuracy and variability with which participants were able to maintain the target response pace was computed. In addition, the extent to which participants' response pace tended to run ahead of or lag behind the target response pace was also calculated. These measures acted as indicators of the participants' ability to form accurate forward model prediction of when best to respond and update their model in response to post-action feedback.

In the current study, it was found that increased age led to greater temporal accuracy and reduced temporal variability after controlling for the influence of impulsivity and sex. Furthermore, this result remained true regardless of the frequency at which the tones were delivered. These findings are consistent with both the first and second hypotheses of the current study. Similarly, the findings also align with the conclusion drawn by McAuley et al. (2006), as they suggest that the accuracy and consistency of individuals' sensorimotor continuation improves with age from childhood to adulthood. Furthermore, the findings extend our current understanding of how the forward model develops, as they suggest that both the quality of the forward model's predictions and the capacity to integrate observed action feedback into the forward model improve with increased age. That being said, it should be acknowledged that, across the frequency conditions, only a small proportion of the variance in participants' temporal accuracy and temporal variability was explained by the final models (R^2 = .03 - .27). Hence, the results should be interpreted with caution.

Despite being consistent with the conclusion of research by McAuley et al. (2006), the findings of increased temporal accuracy and reduced temporal variability with age are not in agreement with the results of research by Witt and Stevens (2013). As posited earlier within this chapter, this disparity may be explained by statistical error. More specifically, Witt and Stevens

(2013) failed to achieve sufficient statistical power in their analyses. In contrast, all analyses were sufficiently powered in the current study. Therefore, this reinforces the idea that the effect of age on sensorimotor continuation was merely obscured by statistical error in research by Witt and Stevens (2013).

As an alternative explanation, differences in the method through which temporal accuracy was measured may have given rise to the discrepancy between the current findings and those obtained by Witt and Stevens (2013). In their study, Witt and Stevens (2013) quantified temporal accuracy by measuring the average asynchrony between a participant's tap and the time at which the participant should have tapped according to the target response pace. Evidently, this measure assumes that preserving a perfect alignment between the onset of each response and the timestamp of each tone is necessary to maintain the presented response pace. However, this is not always the case. For instance, it is plausible that a participant could have successfully reproduced the target ITI, but have shifted the onset of each of their responses slightly later than was expected. This would result in a constant offset between the time of each tone and the time of each response, thus falsely degrading the participant's recorded temporal accuracy. Therefore, this suggests that the findings obtained by Witt and Stevens (2013) were not reflective of participants' genuine sensorimotor continuation skill. In contrast, the current study measured poor temporal accuracy as the average extent to which participants' inter-response-intervals deviated from the target inter-response-interval. Hence, the current findings are more representative of participants' sensorimotor continuation ability, as they were not confounded by deviations in the specific time during the trial at which a response was made.

The present study also found that, when the target ITI was set to 600ms or 800ms, a greater tendency to respond ahead of the target response pace was linked to younger age. Whereas, there was no effect of age on the temporal bias at the high-frequency (400ms) target response pace. These findings appear to suggest that younger individuals' poorer temporal accuracy relative to older individuals arises due to a tendency to overestimate the time at which they need to respond. Alternatively, the association between younger age and tending to respond ahead of the target response pace can also be explained by age-related differences in participants' propensity to shift their response pace to their preferred response tempo. During the continuation phase of synchronisation-continuation tasks, previous research has shown that children's response pace drifted to their preferred response tempo to a greater extent than that of adults (McAuley et al., 2010). In addition, children's preferred response tempo has previously been shown to range between 400-500ms and gradually slow over time before averaging between 600-800ms in adulthood (Baruch et al., 2004; Hammerschmidt et al., 2021; Monier & Droit-Volet, 2019; Provasi & Bobin-Bègue, 2003). Therefore, this suggests that younger participants' tendency to respond faster than the 600ms and 800ms target pace occurred as a result of their response pace drifting to their preferred response tempo of 400-500ms. In support of this idea, the tendency to response ahead of the target pace was found to decline with age, in line with the suggestion that children's propensity to move their response pace to their preferred motor tempo reduces with age. However, this interpretation does not explain the observed link between older age and a greater tendency to respond *behind* the target response pace. Further research is needed to elucidate why this occurred.

During the synchronisation-continuation task, it is believed that individuals predict when to make each response via a weighted combination of a prior and a likelihood (Berniker & Körding, 2011). The likelihood refers to the individual's perception of the time elapsed since their last response (Narain et al. 2018). Whereas, the prior refers to their memory of the average ITI presented during the synchronisation phase of the task (Lewis et al., 2004; Witt & Stevens, 2013). For this reason, one limitation of the current study is that the reported age-related improvement in sensorimotor continuation may have been modulated by developmental changes in participants' capacity to retain the target ITI in working memory. In agreement with this suggestion, previous research has demonstrated that, during the continuation phase of a synchronisation-continuation task, children aged 11-12 were able to maintain an isochronous ITI in working memory for a longer timeframe compared to children aged 6-10 (Gomes et al., 1999). This suggests that the rate at which nonverbal, auditory information degrades over time within working memory declines with age in childhood. Therefore, future research should control for the impact of age-related variability in young people's working memory capacity on the results in order to produce more reliable conclusions on how sensorimotor continuation changes with age from childhood to adulthood.

Aside from age-related deficits in participants' ability to recall the target ITI, it can be argued that, in comparison to older participants, younger participants were less capable of perceiving the target ITI. Past research has shown that children perform worse than adults on tasks where they are asked to discriminate between different auditory and visual stimuli based on their duration (Droit-Volet et al., 2006; Droit-Volet et al., 2007; Zélanti & Droit-Volet, 2012). This suggests that the capacity to accurately perceive presented time intervals matures with age from childhood to adulthood. Therefore, the current results can be partly attributed to age-related improvements in participants' ability to accurately judge the temporal interval presented between the tones, as opposed to being solely the result of maturation in the quality of their forward model.

In addition to age-related differences in time perception, it may be argued that the results could have also been affected by varying years of music training amongst participants. Music training refers to the act of learning to play a music instrument, either formally with an instructor or informally via self-instruction (Braun Janzen et al., 2014). Past research has shown that possessing a greater number of years of music training was associated with more accurate and less variable synchronised tapping performance across participants aged 8-80 (Thompson et al., 2015). This is because experience of learning to play a musical instrument is believed to provide individuals with the opportunity to develop their ability to predict when the next beat will occur within a sequence (Slater et al., 2018) and practice in adjusting the timing of their movements according to auditory cues (Krause et al., 2010). Intuitively, older participants will have had more time in which to gain music training than younger participants. Hence, the observed age-related change in sensorimotor continuation skill could be, at least partially, explained by variation in individuals' years of music training. Therefore, future research aimed at determining the influence of age on sensorimotor continuation should control for the impact of variability in individuals' years of music training.

The results of the current study demonstrated that sensorimotor continuation develops with age from childhood through to adulthood. However, it has been argued that effective sensorimotor continuation is reliant on two distinct mechanisms: i) the individual's ability to form accurate forward model predictions about when to make each response, and ii) their ability to correct their forward model in instances where their response pace has deviated from the target pace (Maes, 2016). Consequently, to gain a holistic understanding of how the forward model system matures with age, future research is needed to separate the relative developmental trajectories of these two mechanisms. This could be accomplished by modifying the current task to include tone sequences with unexpected shifts in the length of the inter-tone-interval, as implemented in research by Hove et al. (2017). By measuring the speed and accuracy with which the individual can alter their response pace during the synchronisation phase to match these unexpected changes, it would be possible to assess how the ability to adjust one's forward model develops with age. Similarly, the extent to which individuals can anticipate when these shifts in inter-tone-interval will occur after repeated exposure could also be quantified, as achieved in research by Mills et al. (2015). Through this measure, it would then be possible to isolate the impact of age on the quality of individuals' forward model predictions on when best to respond.

To conclude, the purpose of the present study was to determine the influence of age on young people's sensorimotor continuation. It was found that both the accuracy and consistency with which the set response pace could be replicated improves with age from childhood to adulthood. As sensorimotor continuation is believed to rely on a forward model, the current results suggest that the forward model becomes more functionally efficient with age. To solidify these interpretations, future studies should elucidate whether the current findings can be replicated after controlling for the impact of variation in individuals' working memory capacity, time perception and years of music training. Moving forward, it would also be beneficial for future research to examine the relative development trajectories of both the ability to form forward model predictions and update the forward model in response to observed sensory feedback. This would provide a more in-depth understanding of how precisely the forward model develops in functional efficiency from childhood to adulthood.

Chapter 4: Exploring the Impact of Age on Predictive Motor Timing from Childhood to Adulthood

Chapter Summary

The purpose of chapter 4 was to explore how predictive motor timing changes from childhood to adulthood, given the suggested role of the forward model in facilitating anticipatory motor action. This was achieved by administering a cued RT task to participants aged 4-25 years. Participants were first presented with a cue stimulus, followed by a target stimulus after a variable interval. Their objective was to respond as soon as a target stimulus appeared. It was found that both the ratio of anticipatory to reactive responses made by the participant and the average speed and consistency of their anticipatory decision process increased with age. This suggests that the ability to form accurate forward model predictions develops with age from childhood to adulthood.

Introduction

The study reported in chapter 3 was successful in demonstrating that individuals become more effective at using their forward model to control the accuracy and consistency of their motor actions with age. Notably, it remains unclear as to whether the age-related improvements described in chapter 3 arose due to the development of participants' prediction abilities, their ability to learn from action feedback, or some combination of the two. By disentangling the developmental trajectories of these two abilities, it will be possible to gain a more in-depth understanding of how the ability to use a forward model improves with age. Therefore, chapter 4 focused exclusively on determining how the rate at which individuals use their forward model to predict when to respond develops with age.

Predictive motor timing refers to an individual's ability to manipulate the timing of an intended action such that its occurrence aligns with the predicted onset of an imminent stimulus (Tanaka et al., 2021). This ability has previously been measured using a cued reaction time task (Brown, 2019). In a cued RT task, participants are first presented with a cue stimulus. The cue stimulus is then succeeded by a target stimulus after a given time interval. Participants' objective is to make a response as soon as the target stimulus has been perceived (Debrabant et al., 2012). Crucially, two distinct cognitive control processes can be employed to achieve this objective: a proactive control process and a reactive control process (Braver, 2012). The proactive control process is prompted by the onset of the cue stimulus. The individual's forward model is used to predict when the target stimulus will most likely occur and hence, when best to respond such that their keypress is temporally-contingent with the target stimulus' onset. Relevant muscles then remain primed to execute the planned, anticipatory response throughout the duration of the cue (Burnett Heyes et al., 2012). Whereas, when the reactive control process is implemented, actions are selected retrospectively in response to the perception of the target stimulus. Consequently, executing a reactive response does not require any internal action preparation via the forward model in advance of the target stimulus' onset, and is instead, triggered solely by external events (Lucenet & Blaye, 2014).

Whilst it may be argued that a reactive response is indeed sufficient for completing a cued RT task, this approach is suboptimal given that the time needed to perceive the target stimulus and produce an appropriate response will contribute to the participant's RT (Burnett Heyes et al., 2012). In comparison, these time costs are reduced when an anticipatory response is made. By using the forward model to predict when the target stimulus will onset, participants have primed their sensory and motor systems to perceive the stimulus and perform an appropriate response before the target has even appeared (Braver, 2012). Therefore, effective predictive motor timing necessitates an anticipatory response strategy, as opposed to a reactive strategy, in order to minimise the individual's RT relative to the target stimulus' onset. For this reason, past studies have examined both the proportion of anticipatory responses produced by the individual (Debrabant et al., 2012), and the speed and consistency at which the decision to make an anticipatory response was reached (Adam et al., 2012; Burnett Heyes et al., 2012) as indices of predictive motor timing.

Notably, few prior studies have attempted to determine the full trajectory at which predictive motor timing develops from childhood to young adulthood (Debrabant et al., 2012). Instead, most studies have merely compared the performance of children and young adults on cued RT tasks (Iselin & DeCoster, 2009), often reporting that the latter demonstrate greater anticipatory response behaviour than the former (Brown, 2019; Perchet & Garcia-Larrea, 2005). This suggests that young adults show a greater tendency to prepare motor responses to anticipated stimuli than children, and thus, are more likely to form predictions about the outcome of their action. However, it remains unclear as to whether these results can be extended across adolescence. A limited number of studies have shown that the ability to predict both the location (Van Gerven et al., 2016) and the identity (Iselin & DeCoster, 2009) of expected stimuli and prepare appropriate responses improves with age across adolescence. This increase in anticipatory behaviour has been argued to result from maturational changes occurring within the prefrontal cortex during adolescence (Smith et al., 2011). Therefore, it has been posited that predictive motor timing, or in other words, the tendency to use the forward model to prepare actions in advance, should also improve as individuals develop throughout adolescence. However, further research is needed in order to examine this idea empirically.

The Current Study

The purpose of the current study was to establish the extent to which the tendency to form predictions about when to respond in anticipation of a sensory event improves with age from childhood to adulthood. This would reveal how the capacity to generate predictions about the outcomes of planned actions using a forward model changes with age across this period. To achieve this, participants completed a cued RT task where they were presented with an amber cue stimulus, followed by a target green stimulus after a variable interval. Participants were instructed to click the screen as soon as the target stimulus became visible. Responses were classified as anticipatory or reactive based on the participant's RT. From this, the ratio of anticipatory to reactive responses was calculated. The higher ratio, the greater the extent to which the participant tended to use their forward model to form predictions about the outcome of their action. In addition, participants' RT data was also fitted to a two-horse linear rise-to-threshold model to calculate the speed and variability at which the decision to make an anticipatory response was reached. The faster and more consistent the rate of rise in their anticipatory decision process, the greater their tendency to form forward model predictions and prepare response in advance. To extend the findings of past research (e.g., Brown, 2019), two hypotheses were made:

1) It was hypothesised that the ratio of anticipatory to reactive responses would be predicted by age, with greater age associated with a larger ratio.

2) It was hypothesised that the speed and variability in the rate of rise in the anticipatory decision process would be predicted by age, with greater age associated with a faster and more consistent rate of rise.

Method

Design

An independent measures design was used in the current study. The independent variables were age, sex and impulsivity. The key dependent variables were the ratio of anticipatory to reactive responses, the mean rate of rise in the anticipatory decision process, and the variability in the rate of rise in the anticipatory decision process. These dependent variables acted as indicators of participants' ability to prepare actions in advance; a skill that demands the use of the forward model.

Participants

323 participants were initially recruited (83 male, 239 female, and 1 preferred not to say). The age of participants ranged from 4 to 25 years (*M*=16.36, *SD*=4.14).

- 220 participants were right-handed, 23 was left-handed and 52 were ambidextrous, as measured by the EHI-SF (see appendix A for a full outline of this measure). 28 participants did not complete the EHI-SF.
- 212 participants were White, 19 were Asian, 7 were Black, 10 had mixed/multiple ethnic identities, 3 had another ethnic identity that was not listed, 8 preferred not to say, and 64 did not report their ethnic identity.

Participants were recruited through one of four avenues: 70 participants were recruited via SSM in August 2020 and August 2021, 58 were recruited from two high schools in the Nottinghamshire and Derbyshire areas from January 18th 2022 - 16th March 2022, and the remaining 195 were recruited either through RPS or through recruitment posters published on social media between 18th November 2020 - May 15th 2021. For more detailed information on how participants were recruited from each of these sources and how informed consent was obtained, please see chapter 2. The full experimental procedure of the current study was approved by the School of Psychology ethics committee at the University of Nottingham. Seventeen participants were excluded because they self-reported a diagnosis of either ADHD or ASD. The rationale for removing participants with these two conditions was outlined in chapter 3. All remaining participants were neurotypical. One participant was also removed because they responded during the red light presentation on >50% of the trials on the cued RT task. The demographics of the adjusted sample can be viewed in table 4.1.

Table 4.1. The demographic characteristics of the adjusted sample.

	Age (*years*)	Sex	Ethnicity*	Handedness**
Full sample (*n*=305)	*Range*= 4 to 25 *M*=16.39 *SD*=4.17	78 Male 226 Female 1 Preferred not to say	199 White 19 Asian 7 Black 10 Mixed/multiple ethnic identities 2 Any other ethnic identity 7 Preferred not to say	209 right-handed 22 left-handed 49 ambidextrous

Note. *Ethnicity information was not collected for 61 participants. **Handedness information was not collected for 25 participants.

Materials

The Cued Reaction Time Task. A cued RT task was used to measure participants' predictive motor control. The task closely mirrored the tasks used in research by Brown (2019) and Burnett Heyes et al. (2012). The task was designed using PsychoPy and ran online via Pavlovia (Peirce, 2019, 2022). Stimuli consisted of two 119x178 pixel images of cartoon race cars. The cars were coloured red and blue to differentiate between the participant-controlled car and the computer-controlled car. Stimuli also included a 500x200 pixel image of a traffic light. The traffic light changed colour sequentially from red to amber to green during each trial.

Self-Report and Parent-Report Measures. The following self-report measures and parent-report measures were administered in the current study:

- The SWAN-Hyperactive/Impulsive subscale was used to measure child participants' impulsivity as reported by their parents or carers.
- The UPPS-P short-form was used to measure adolescents' and young adults' self-reported impulsivity.

Please see chapter 2 for a full outline of the SWAN rating scale and UPPS-P short-form.

Procedure

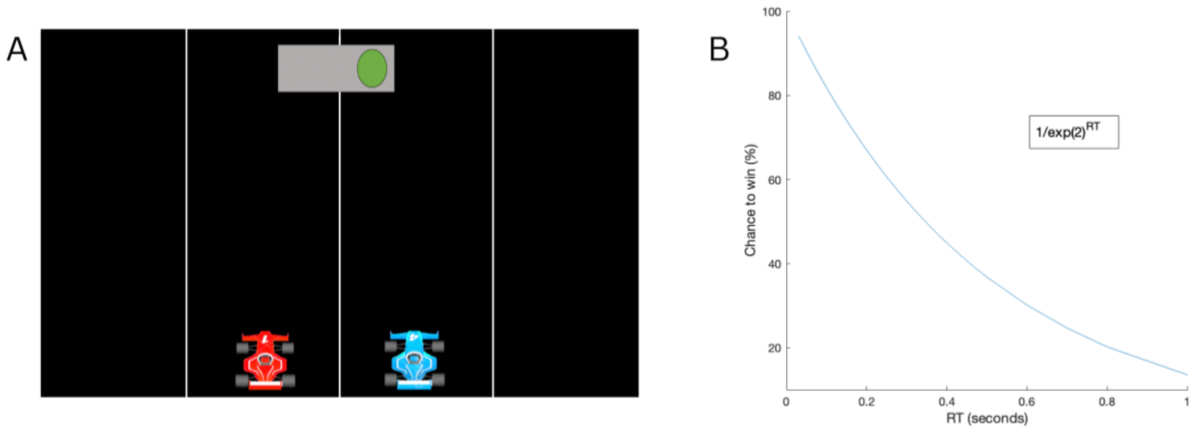
Similar to the synchronisation-continuation task reported in chapter 3, two versions of the cued RT task were created with the same age groups in mind: one version for adolescents and young adults and a second version for children. The adolescent and young adult version was available to participants via the two high schools, RPS, and social media posts. Whereas, the child version was administered at the two SSM events. As noted in chapter 3, although SSM was aimed at children aged 4-12, individuals aged between 12-17 were also permitted to take part. Therefore, it was possible for a participant aged >12 to have received the child version of the task.

Adolescent and Young Adult Procedure. Adolescent and young adult participants provided their demographic information and completed the UPPS-P short-form via a survey hosted on Qualtrics before being redirected to the cued RT task.

Upon opening the cued RT task, participants first saw a black instructions screen with details about how to complete the task. Participants started the trials by pressing the spacebar. Participants were instructed to complete the task using the same hand throughout. Participants first completed 5 practice trials to familiarise themselves with the task, before progressing to the main experiment trials. In each trial, participants were presented with a red car and a blue car positioned at the bottom of the screen (see figure 4.1A). The red car and blue car were each controlled by the participant and the computer, respectfully. A set of traffic lights was positioned at the top of the screen with only the red light visible. After 1s, the red light was replaced by an amber light. The duration of the amber light varied on each trial via a Gaussian distribution with a mean of 750ms and a standard deviation of 125ms. When the amber light duration terminated, the green light was shown for 1s. Participants responded by clicking the screen as soon as the green light was visible. Time taken to click the screen was recorded relative to both the amber light's onset and the green light's onset. On each trial, the chance of winning was determined via an inverse exponential temporal discount function based on the participant's reaction time relative to the green light's onset ($1/\exp(2)^{RT}$; see figure 4.1B). Faster responses resulted in a higher chance of winning the trial. However, any responses that occurred before the green light's onset or after the green light's duration had ended resulted in the participant losing the trial. This time pressure was designed to encourage participants to prepare their responses in anticipation of the green light's onset.

Figure 4.1.

Image of the cued RT task and the inverse exponential temporal discount function.



Note. 4.1A. An image showing the cued RT task. The participant's and computer's cars are positioned at the bottom of the screen. The traffic light is shown at the top of the screen with the green light visible. 4.1B. A figure showing the inverse exponential temporal discount function used to calculate a participant's chance of winning a trial based on their reaction time relative to the green light's onset.

Making a response caused the cars to travel up the screen to simulate a car drag race. If the participant won the trial, then the participant's car moved at a faster pace than the computer's car. The opposite occurred when the participant lost the trial. As the cars moved, the sound of a car engine was heard. Each trial lasted for approximately 3s. After each trial, a feedback screen was presented for 1s. The feedback screen informed participants of whether they had won, lost, responded during the red light presentation, or failed to respond on the previous trial. The feedback screen also showed the current number of trials that both the participant and the computer had won throughout the experiment. If either no response was made or a response was made during the red light presentation during a trial, the feedback screen reminded participants of the task instructions. Break screens were shown every 25 trials for 2s each. The break screens included encouraging text to promote participants' attention. Different encouraging messages were shown depending on whether the current number of trials won by the participant was greater or lower than the number of trials won by the computer. A reward screen was shown every 25 trials for an unlimited time. The reward screen featured an image of either a bronze, silver, or gold trophy and an encouraging message. An inter-trial interval was presented for 1s, during which a black screen was displayed.

There were 100 trials in total. The full adolescent and young adult procedure lasted approximately 15-minutes.

Child Procedure. The procedure completed by children mirrored that of the adolescent and young adult participants with a few minor exceptions. Child participants' parents or carers provided their child's demographic information and completed the SWAN rating scale when registering their child for SSM. Once registered, participants could complete the cued RT task at any time throughout the SSM event. Participants were encouraged to complete the left-hand vs right-hand task before completing the cued-RT task to measure their hand preference during daily tasks. For a full description of the left-hand vs right-hand task, please see appendix B. On the cued RT task, break and reward screens were presented every 10 trials rather than every 25 trials. Reward screens contained an image of a sticker and an encouraging message to promote participants' attention. The full child procedure lasted approximately 10-minutes.

Data Analysis

All data pre-processing procedures were conducted in MATLAB. All participants' ages were converted to natural logarithmic age, as described in chapter 2.

Self-Report Scales. All self-report and parent-report scales were summed and averaged to create an index for each of the variables of interest. In addition, SWAN-Hyperactive/Impulsive subscale scores and UPPS-P short-form scores were converted to z-scores. This meant that the impact of impulsivity on task performance could be investigated across age, despite the fact that impulsivity was measured using the SWAN-Hyperactivity/impulsivity subscale for child participants and the UPPS-P short-form for adolescent and young adult participants.

The Recorded Average Amber Light Duration. On each trial of the cued RT task, the duration of the amber light varied around a Gaussian distribution with a mean of 750ms and an SD of 125ms. This means that it should not have been possible for the amber light duration to be less than 500ms or greater than 1-second as this would have been over 2SDs away from the mean amber light duration. An amber light duration of this magnitude would likely have been caused by a technical error, such as a temporal delay in the updating of visual stimuli by the participant's internet browser. For this reason, the total number of trials with a recorded amber light duration that was less than 500ms or greater than 1-second was recorded and these trials were removed from the analysis. This provided a means through which to retrospectively verify whether the task ran as intended online and remove the statistical influence of any trials where technical errors may have occurred. On average, 0.36% of trials displayed an amber duration outside of the expected range of 500ms-1s (*SD* = 1.35, *range* = 0-17).

Percentage of Non-Response Trials. On the cued RT task, participants were instructed to respond by clicking the screen as soon as the green light became visible. They then had 1-second in which to make their response. Hence, a lack of a response on a given trial indicated either poor task comprehension or participant inattention. On average, 1.47% of trials received no response (*SD*= 3.77, *range*= 0 - 33). All trials without a response were removed from further analysis in order to remove the influence of these occurrences on the results.

Percentage of Responses Made During the Red Light. On each trial of the cued RT task, participants were presented with a set of traffic lights which moved from red to amber to green across the course of the trial. To reiterate, the red light and green light were both on screen for 1-second each. Whereas, the amber light's duration was variable. Due to the uncertainty around the exact duration of the amber light on any given trial, it is feasible that a participant who has prepared their motor response in anticipation of the green light's onset may accidentally respond whilst the amber light is still onscreen. However, unlike the amber light, the red light's presence does not signify that the green light's onset is imminent. Therefore, any responses made during the red light's presentation are unlikely to have been caused by the mistimed execution of a prepared motor response and are instead, more likely to have resulted from either poor task comprehension or participant error. On average, 1.28% of trials were responded to during the presentation of the red light (*SD*= 2.73, *range*= 0 - 20.48). All trials for which this type of response was made were removed from further analysis in order to remove the influence of these occurrences on the results. Taken together, 2.76% of trials were removed on average due to either: an amber duration outside of the expected range, a failure to respond or a response made during the red light presentation (*SD* = 5.31, *range* = 0 - 45.16).

Reaction Time. Reaction time was calculated relative to both the onset of the amber light (RT_A) and the green light onset (RT_G). It is useful to examine both RT_A and RT_G, as an RT_A demonstrates precisely where the participant's response fell within the time window available for a valid response to be made, which overlaps both the amber and green light durations. This information is helpful for constructing the two-horse linear rise-to-threshold model, as outlined below. Whereas, an RT_G provides a more direct indication of how successful the participant was in achieving the task goal, i.e. the temporal closeness of their response to the green light's onset. A negative mean RT_G shows that the participant had a tendency to respond before the green light's onset. Whereas, a positive RT_G indicates that the participant tended to respond after the green light's onset.

The Anticipatory to Reactive Response Ratio. Anticipatory responses were defined as any response with an RT_A that was less than the cut-off criterion of 850ms. This value was selected as it combined the average amber light duration (750ms) and 100ms of the green light's duration. This is in line with past research which has defined an anticipatory response as any response made within 100ms of the onset of the target stimulus (e.g., Debrabant et al., 2012). Therefore, any response that was made between the amber light onset and 100ms into the green light's duration was classified as anticipatory. Whereas, all responses made after the cut-off criterion were counted as reactive responses. The total number of anticipatory responses was divided by the total number of reactive responses to form the anticipatory to reactive response ratio. This ratio indicated whether a participant tended to respond in an anticipatory or reactive manner. Higher values demonstrated a larger tendency to make anticipatory responses, and thus, showed more frequent use of the forward model to plan motor responses ahead of their execution, as opposed to merely responding reactively.

The Two-Horse Linear Rise-to-Threshold Model Parameters. All valid responses made on the cued RT task can be classified as either the product of an anticipatory or a reactive decision process. As each trial progresses, these two decision processes compete in a two-horse race toward the threshold level required for a response to be executed (Carpenter et al., 2009). The decision process which reaches the action execution threshold first determines the identity of the performed action. The anticipatory decision process rises slowly from the onset of the amber light toward the action execution threshold; a response is prepared in anticipation of the green light's arrival. Whereas, the reactive decision process is triggered by the presence of the green light and rises sharply to the response initiation threshold; here, no prior response preparation transpires (Burnett Heyes et al., 2012).

To statistically capture this race, a two-horse linear rise-to-threshold model outlined by Burnett Heyes et al. (2012) was applied to each participant's RT_A data (see equation 4.1). $Pr(T \leq t)$ denotes the cumulative probability function of a participant's RT_A data, i.e., the probability that a response had occurred (T) by time t following the amber light onset. ΨA and ΨR signify the anticipatory and reactive cumulative recinormal distributions. The notation, t0, refers to the cut-off criterion, 850ms. Hence, the ΨA was constructed on the probability that the participant responded at a time after the amber light onset (t). Whereas, the ΨR was based on the probability that the participant responded at a time following the cut-off criterion (t-t0).

$$Pr(T \leq t) = \psi A(t) + \psi R(t - t0) - \psi A(t)\psi R(t - t0) \tag{4.1}$$

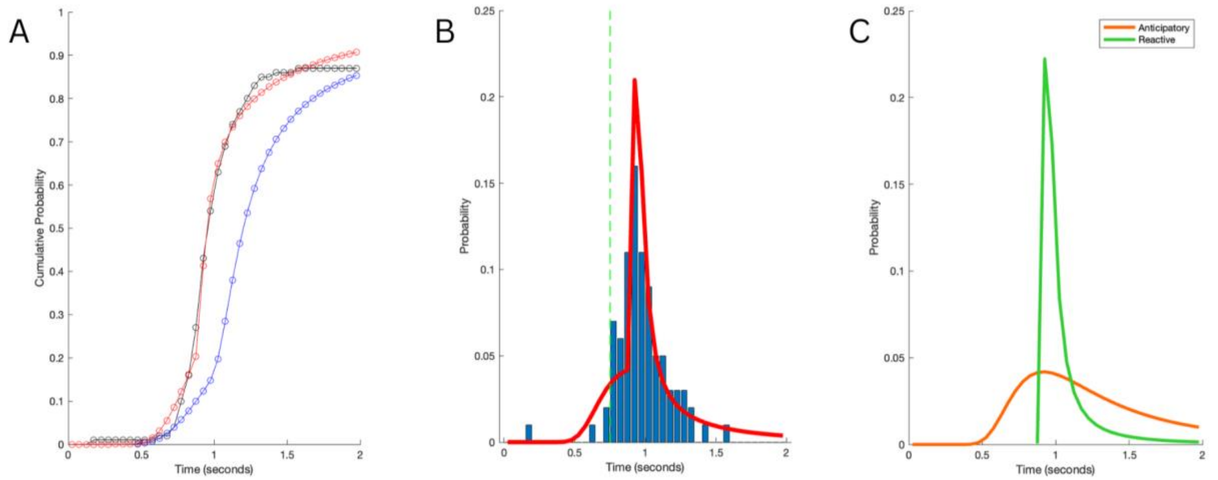
In order to fit the two-horse linear rise-to-threshold model to the data, a cumulative probability function (CPF) of the participant's RT_A data was first calculated (see the black line on figure 4.2A). The CPF describes the probability that a response had occurred (T) by time t following the amber light onset. To construct the CPF, both the average duration of the amber light (750ms) and the full duration of the green light were segmented into equally spaced intervals of 50ms, spanning 1.75 seconds in total. For each time interval, the probability that the participant had made a response by the onset of that specific time interval. In addition to the CPF, a probability density function (PDF) was also produced for visualisation purposes using a similar procedure; the probability that the participant had made a response during each individual time interval was recorded (see the blue bars on figure 4.2B).

Next, the two-horse linear rise-to-threshold model was used to fit two cumulative recinormal distributions to the participant's CPF, each representing the anticipatory (ΨA) and reactive (ΨR) decision processes, respectfully (see the red line on figure 4.2A for the overall model fit). A cut-off criterion (t0) of 850ms was used to define the time windows for which an anticipatory or reactive response would fall. As described above in the case of the anticipatory to reactive response ratio, this cut-off value was selected as it incorporated the average amber light duration (750ms) and 100ms of the green light's duration. Hence, the ΨA was constructed based on the probability that the participant responded at any time after the amber light onset (t). Whereas, the ΨR was based on the probability that the participant responded at any time following the cut-off criterion (t-t0).

Through maximum likelihood estimation, the mean rate of rise for the anticipatory (μA) and reactive (μR) distributions and their corresponding SDs (σA, σR) were calculated. The mean rate of rise and standard deviation of each distribution demonstrates, on average, how quickly and variably each of the two decision processes rose toward the required threshold for executing a response. For a visual example of these decision processes, please refer to figure 4.2C. This shows one participant's PDFs for the anticipatory and reactive decision processes constructed using the mean and SD parameters obtained through fitting the two-horse linear rise-to-threshold model to their RT_A data. Finally, a coefficient of variation was calculated for each decision process in order to obtain a more precise measure of variability than is afforded by the standard deviation (CV_A, CV_R). If a participant tended to prepare a response using their forward model, then the ΨA should have risen sharply towards the action execution threshold, resulting in a higher μA and lower CV_A compared to those who tended to make only reactive, stimulus-driven responses.

Figure 4.2.

The cumulative probability function, probability density function, and anticipatory and reactive distributions for one participant.



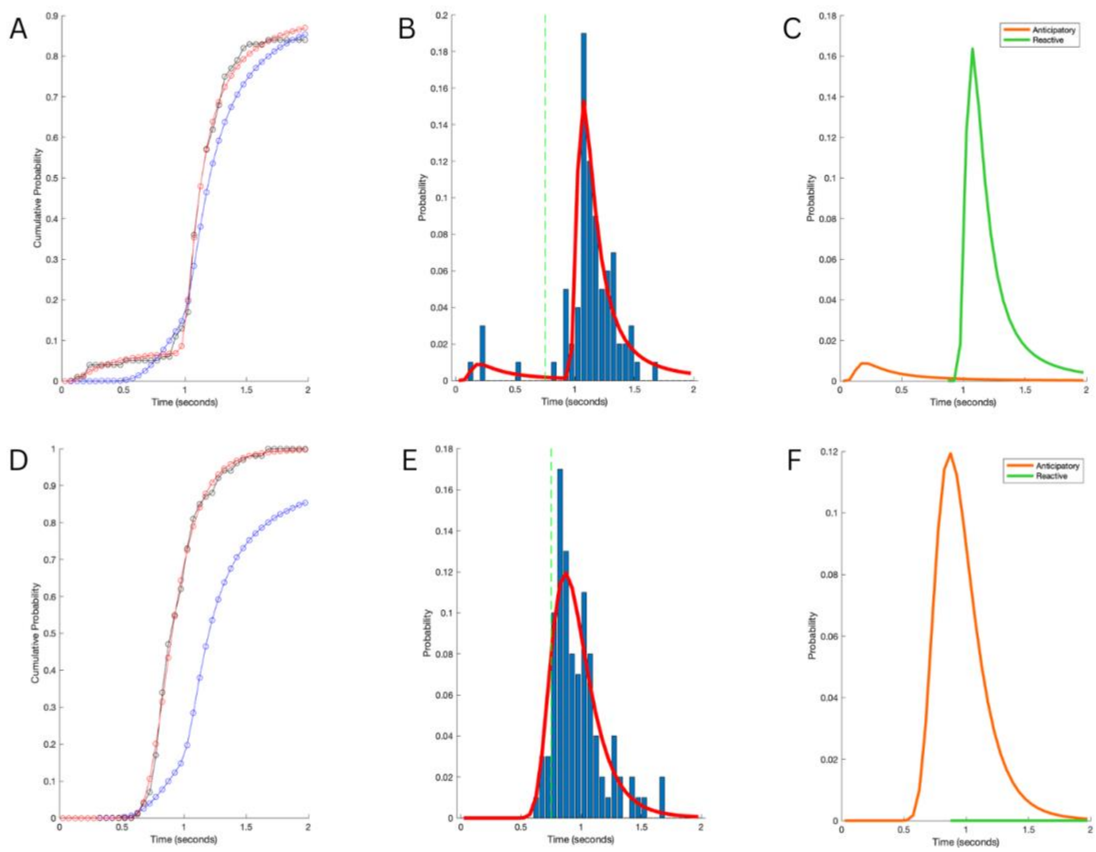
Note. 4.2A. A figure showing the cumulative probability function (CPF) for participant 9. This shows the cumulative probability that the participant had made a response at each incremental time interval throughout the combined timecourse of the amber and green light durations. The black line shows the participant's CPF constructed from their RT_A data. The blue line shows the two-horse linear rise-to-threshold model (see equation 4.1) prior to being fitted to the participant's CPF through maximum likelihood estimation. The ΨA and ΨR distributions in this default model have the following arbitrary parameters: $\mu A = 0.5$, $\sigma A = 0.5$, $\mu R = 2$, $\sigma R = 2$. The red line then shows the two-horse linear rise-to-threshold model after being fitted to the participant's CPF. The fitted model parameters for this participant were $\mu A = 0.78$, $\sigma A = 0.44$, $\mu R = 5.18$, $\sigma R = 10.69$. 4.2B. A figure showing the probability density function (PDF) for participant 9's RT_A data. This reveals the probability that the participant had made a response by any given timepoint within the amber and green light durations. Each blue bar represents the probability that the participant responded during the specific time interval to which the blue bar spans. The dashed green vertical line indicates the average onset of the green light, i.e., 750ms after the amber light's onset. The red line shows the probability density function constructed using the participant's fitted model parameters. 4.2C. A figure showing the separate probability density functions for the anticipatory and reactive decision processes constructed using the fitted model parameters for participant 9. The orange line shows the anticipatory distribution. Whereas, the green line shows the reactive distribution.

Despite both defining 850ms as a cut-off criterion during their formation, it is important to clarify that the μA is not synonymous with the anticipatory to reactive response ratio. During the calculation of the anticipatory to reactive response ratio, each RT_A is labelled as either anticipatory or reactive according to an absolute binary classification system. Any responses that are made before the cut-off criterion are classified as anticipatory, i.e., prepared before the green stimulus onset. Whereas, all those that fall after this value are reactive. The specific time at which a response occurred within the anticipatory or reactive temporal window is irrelevant. Whereas, as the two-horse linear rise-to-threshold model's parameters capture the speed and variability at which each decision process triggered a motor response, all four variables are dependent on precisely when during a trial the response was made. For instance, if a participant responded at approximately the same time point within the anticipatory time window on multiple trials then this would lead to a higher μA and a lower CV_A compared to if their responses were more widely distributed throughout this period across the trials. In the former scenario, the slope of the PDF curve for the ΨA (as visualised in figure 4.2C) would rise sharply to a peak. The same is also true for the μR and the CV_R; the greater the temporal clustering of the reactive responses within the reactive time window, the greater the μR and the lower the CV_R.

In some circumstances, a negative μA or μR could occur. These negative values indicated that the two cumulative recinormal distributions could not be fit to the participants' response probability data. For instance, in the case of participant 10, a μA of -8.81 was observed. The participant's RT_A data appeared to show a binary pattern of anticipatory responses; there was a high probability that the participant would respond both at the very beginning of the amber light's duration and later towards the end of the anticipatory response window (see figures 4.3A-C). An optimal PDF curve for this dataset would have two peaks. Hence, this binary response pattern cannot be adequately captured by a single μA and CV_A parameter. As a result, the produced μA value is unlikely to be representative of the participant's predictive motor timing behaviour. Similarly, participant 60 demonstrated a large negative μR of -202.34. The data showed that this participant had a high probability of responding both immediately after the cut-off criterion and at a later stage in the green light's duration (see figures 4.3D-F). Again, this binary pattern of reactive responses cannot not be reliably captured using a single μR and CV_R parameter. As these values can be attributed to statistical error, any negative μA, μR, CV_A, or CV_R values were excluded from the statistical analyses. For this reason, 9 μA, 9 CV_A, 122 μR, and 122 CV_R data points were removed from the data.

Figure 4.3.

The cumulative probability functions, probability density functions, and anticipatory and reactive distributions for two participants with erroneous model fits.



Note. 4.3A-C. Three figures showing the cumulative probability function (4.3A), probability density function (4.3B), and anticipatory and reactive distributions (4.3C) for participant 10. See figure 4.2 for a detailed description of how the data is presented in each figure. In particular, on 4.3B, note the relatively tall bars close to time = 0 on the x-axis and just after the green dashed line. This indicates that this participant made a high proportion of responses both early into the amber light duration and within 100ms of the green light's onset, leading to an erroneous, negative μA value (μA = -8.81). Notably, the constructed PDF curves (4.3C) fail to capture this binary response pattern. 4.3D-F. Three figures showing the cumulative probability function (4.3D), probability density function (4.3E), and anticipatory and reactive distributions (4.3F) for participant 60. See figure 4.2 for a detailed description of how the data is presented in each figure. On 4.3D, it can be seen that this participant made a high proportion of responses both immediately after the cut-off criterion (850ms) and later in the green light's duration, leading to an erroneous, negative μR value (μR = -202.34). Notably, the constructed PDF curves (4.3F) were a poor fit for the participant's data; most reactive responses appear to have been falsely classified as anticipatory.

Bayesian Learning Model Parameters. In addition to determining how the ability to form predictions via the forward model changes with age, the current research had also aimed to examine age-related changes in the ability to update the forward model in light of the goal-related information gained via past action experience. To achieve this aim, a Bayesian learning model was fitted to each participant's RT_A data. For each trial of the cued RT task, the fitted model parameters revealed the weight that the participant had attributed to an average estimate of all past amber light durations (i.e., the prior) relative to the amber light duration shown on the most recent trial (i.e., the likelihood) when predicting the most probable time at which the green light would onset (i.e., the posterior), and therefore, when best to respond. In accordance with typical learning tasks (Jacobs & Kruschke, 2011; Yu & Dayan, 2004), it was expected that the weight on prior would increase over time as more amber durations were observed. Hence, the higher the average weight on the prior, the greater the participant's ability to successfully update their forward model in light of new information. Unfortunately, it was found that all participants failed to construct a reliable representation of the prior. This suggests that either the cued RT task was not suitable for this type of analysis, or the Bayesian learning model used was incorrect. For this reason, the influence of age on the weight attributed to the prior and the likelihood during the cued RT task was not examined in this thesis (see appendix C for more information).

Outlier Detection. To exclude any potentially confounding anomalous data, the Tukey's fences method for outlier detection was applied to the dataset (Tukey, 1977). Fifteen anticipatory to reactive response ratio, 10 μA, 22 CV_A, 15 μR, 15 CV_R data points were found to be more than 1.5 interquartile ranges away from the nearest quartile. Upon comparison, it was found that the removal of the anomalous data points did not affect the direction or significance of the results regarding the influence of age, sex and impulsivity on the ratio of anticipatory to reactive responses, the μA, or the CV_A. Hence, the identified data points for these variables were not removed in order to ensure the completeness of the data. However, excluding the extraneous data points for μR and CV_R did have an effect on the findings. Prior to outlier removal, μR was significantly predicted by impulsivity (β = .18, t= 2.47, p = .01) and CV_R was not predicted by any predictor variable. Whereas, age significantly predicted both μR (β = .21, t= 2.76, p= .01) and CV_R (β = -.17, t= -2.23, p= .03) after the outlier data points were excluded. In addition to age, CV_R was also significantly predicted by impulsivity (β = .16, t= 2.12, p= .04) after outliers were removed. Hence, the 30 data points identified as outliers for these two variables were removed from the data so as not to statistically bias the results.

Statistical Analyses. All statistical analyses were run in SPSS. For the reasons outlined in chapter 2, impulsivity and sex were included as nuisance variables in the current study. In further support of the inclusion of impulsivity as a nuisance variable in the current study, it can be noted that, in order to perform well on the cued RT task, participants needed to refrain from executing their response until a moment close to the target visual stimulus' onset time. Past studies have demonstrated that individuals with a high level of impulsivity struggle to withhold their responses in comparison to those with lower levels of impulsivity (Leshem & Yefet, 2019). Therefore, impulsivity was included as a nuisance variable to control for the impact of age-related variation in impulsivity on the current findings. Similarly, in terms of sex, past research has suggested that females demonstrate greater efficiency in deploying proactive control over their actions compared to males (Bianco et al., 2020; Yücel et al., 2012; Smittenaar et al., 2015). Furthermore, it has been argued that this is due to sex differences in the rate at which regions of the brain implicated in cognitive control develop (Christakou et al., 2009). Hence, sex was also included as a nuisance variable in the analyses.

Five stepwise multiple linear regressions were performed to investigate whether the anticipatory to reactive response ratio, the μA, the CV_A, the μR, or the CV_R could be predicted by age, impulsivity, or sex. In each stepwise multiple linear regression analysis, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block. The purpose of these statistical tests was to reveal the extent to which an individual's age can influence their predictive motor control when the respective influences of both impulsivity and sex are taken into account. G*Power analysis revealed that 77 participants were required to obtain a medium sized-effect (f^2=.15) in a multiple linear regression with three predictor variables, 80% power, and a 5% alpha level (Faul et al., 2009). Therefore, as the sample contained 305 participants, the performed analyses were sufficiently powered.

Results

In the current study, participants completed 100 trials of the cued RT task, which involved making a mouse click response as soon as a green light stimulus became visible. On average, participants responded 194.59ms after the green light's onset (*SD* = 76.97, *range* = 30.76 - 500.55).

The Influence of Age on Participants' Anticipatory and Reactive Response Behaviour

On each trial of the cued RT task, a participant could respond in either an anticipatory or a reactive manner. Anticipatory responses necessitate the use of the forward model as they have to be prepared during the amber light's presentation in order to ensure that a response can be given as soon as the green light becomes visible. Whereas, a reactive response was triggered by the onset of the green light and therefore, did not require any predictive motor planning using the forward model (Burnett Heyes et al., 2012). On average, 30.4% of participants' responses were anticipatory (*SD* = 15.12, *range* = 1 - 82). Notably, this demonstrates that every participant made at least one anticipatory response. Additionally, this average anticipatory response rate was significantly greater than 0, $t(304) = 35.14$, $p<.001$, $d= 15.11$, meaning that a sufficient level of anticipatory responses were made across participants to warrant further analysis. That being said, it should be noted that this figure was significantly lower than the 47% rate of anticipatory behaviour reported by Brown (2019), $t(304) = -19.19$, $p<.001$, $d = 15.12$. The anticipatory to reactive response ratio was calculated to provide an indication of a participant's tendency to respond in an anticipatory manner relative to their tendency to make a reactive response. Higher values demonstrate a larger tendency to make anticipatory responses over reactive responses, and thus, show more frequent use of the forward model to plan motor responses ahead of their execution.

The μA and the CV_A were obtained from participants' reaction time data using the two-horse linear rise-to-threshold model. Similar to the anticipatory to reactive response ratio, the μA and the CV_A also provide information regarding the participant's anticipatory response behaviour;

they demonstrate the average speed and variability with which a participant's anticipatory decision process rose toward the action execution threshold. The larger the μA, the sharper the rise in their anticipatory decision process and thus, the greater their tendency to plan their response ahead of time using the forward model. Likewise, the lower the CV_A, the narrower the rise in their anticipatory decision process and therefore, the more consistent their tendency to plan a response, as opposed to responding reactively.

In addition to the μA and the CV_A, the μR and the CV_R were also calculated from participants' reaction time data via the two-horse linear rise-to-threshold model. The μR and the CV_R show how quickly and variably a participant's reactive decision process rose toward the action execution threshold on average. Much like the μA, the greater the μR, the sharper the rise in the participant's reactive decision process and the greater their tendency to respond reactively. The lower the CV_R, the narrower the rise in the reactive decision process and the more consistent their tendency to respond reactively, rather than plan their responses. It was important to examine the μR and the CV_R as they act as the foil to the both the μA and the CV_A, respectively. Therefore, both parameters reveal the extent to which a participant failed to recruit the forward model and engage in predictive motor planning, and instead produced reflexive responses.

Five stepwise multiple linear regressions were conducted on the anticipatory to reactive response ratio, the μA, the CV_A, the μR and the CV_R, with the factors: age, impulsivity, and sex. In each regression analysis, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block (see table 4.2 for the final models). It was revealed that all five variables were significantly predicted by age (all $p<.05$). The anticipatory to reactive response ratio, the μA, and the μR were found to increase with age. Whereas, the CV_A and the CV_R decrease with age. None of the dependent variables were predicted by sex (all $p>.05$). Impulsivity was only a significant predictor of the CV_R ($p<.05$) and did not significantly predict any of the other four variables (all $p>.05$). As impulsivity score increased, CV_R increased. For visualisation purposes, see figure 4.4 and 4.5 for each of the independent variables plotted against participants' unlogged age in years.

It should be noted that, initially, a significant model was found for CV_A before age was added to the model, $F(1, 289) = 4.96$, $p = .03$. CV_A was significantly predicted by sex ($\beta = -.13$, $t = -2.23$, $p = .03$) and was not predicted by impulsivity ($\beta = .08$, $t = 1.28$, $p = .2$). The model fit was $R^2 = .02$, $SE = .58$. However, after age was added to the model, CV_A was significantly predicted by age ($\beta = -.2$, $t = -3.42$, $p = .001$), was not predicted by impulsivity ($\beta = .07$, $t = 1.26$, $p = .21$) and was no longer predicted by sex ($\beta = -.07$, $t = -1.2$, $p = .23$). Likewise, a significant model was also initially found for CV_R before age was added to the model, $F(1, 163) = 4.31$, $p = .04$. CV_R was significantly predicted by impulsivity ($\beta = .16$, $t = 2.08$ $p = .04$) and was not predicted by sex ($\beta = -.03$, $t = -.44$, $p = .66$). The model fit was $R^2 = .03$, $SE = .96$. However, after age was added to the model, both age ($\beta = -.16$, $t = -2.13$, $p = .04$) and impulsivity ($\beta = .15$, $t = 2.01$, $p = .05$) were revealed to be significant predictors of CV_R, whereas sex was not a significant predictor ($\beta = .03$, $t = .38$, $p = .7$).

Table 4.2. The results of five stepwise multiple linear regression analyses investigating the influence of age, impulsivity and sex on the temporal bias in the anticipatory to reactive response ratio, μA, CV_A, μR, and CV_R

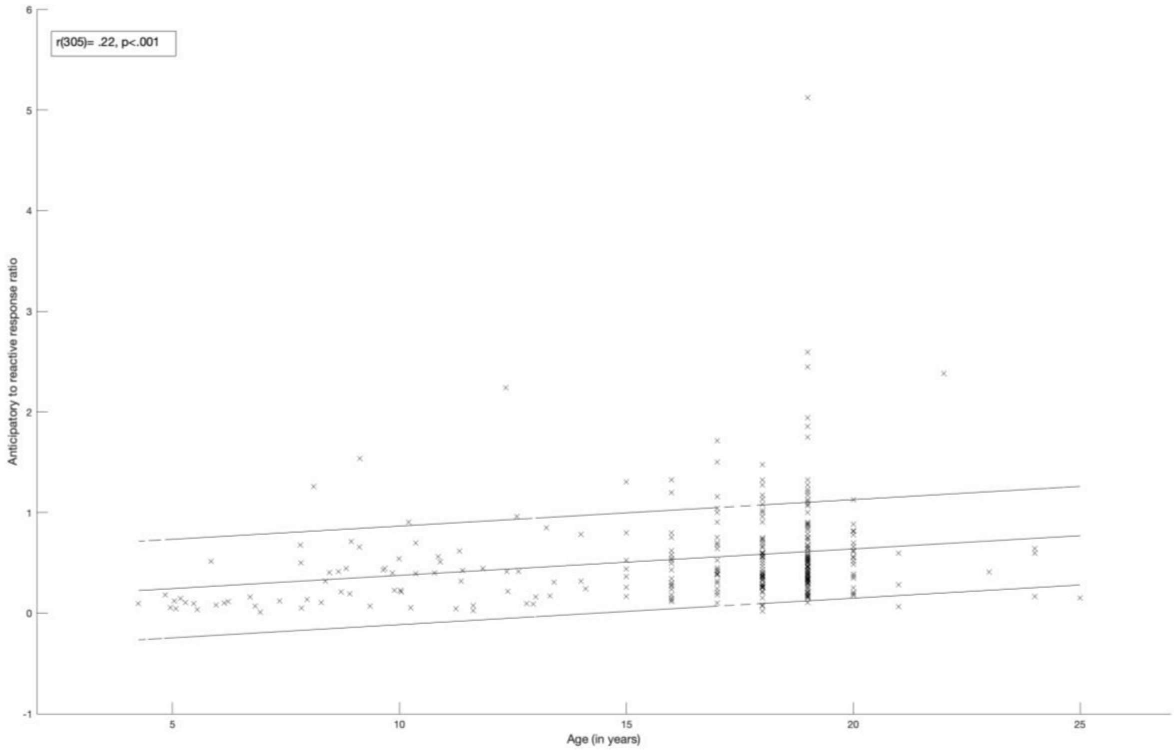
Final regression model	*β*	*t*	*F*	*df*	R^2	*SE*
The anticipatory to reactive response ratio			15.56**	1, 298	.05	.48
Age	.22**	3.94**				
Impulsivity	.05	.96				
Sex	-.07	-1.09				
The mean rate of rise in the anticipatory decision process (μA)			95.8**	1, 289	.25	.12
Age	.5**	9.79**				
Impulsivity	-.06	-1.21				
Sex	-.05	-.88				
The variability in the rate of rise in the anticipatory decision process (CV_A)			8.4**	2, 288	.06	.57
Age	-.2*	-3.42*				
Impulsivity	.07	1.26				
Sex	-.07	-1.2				

Final regression model	β	t	F	df	R^2	SE
The mean rate of rise in the reactive decision process (μR)			7.05*	1, 164	.04	1.65
Age	.2*	2.66*				
Impulsivity	-.01	-.16				
Sex	-.08	-1.02				
The variability in the rate of rise in the reactive decision process (CV_R)			4.47*	2, 162	.05	.95
Age	-.16*	-2.13*				
Impulsivity	.15*	2.01*				
Sex	.03	.38				

Note. *p<.05 **p<.001

Figure 4.4.

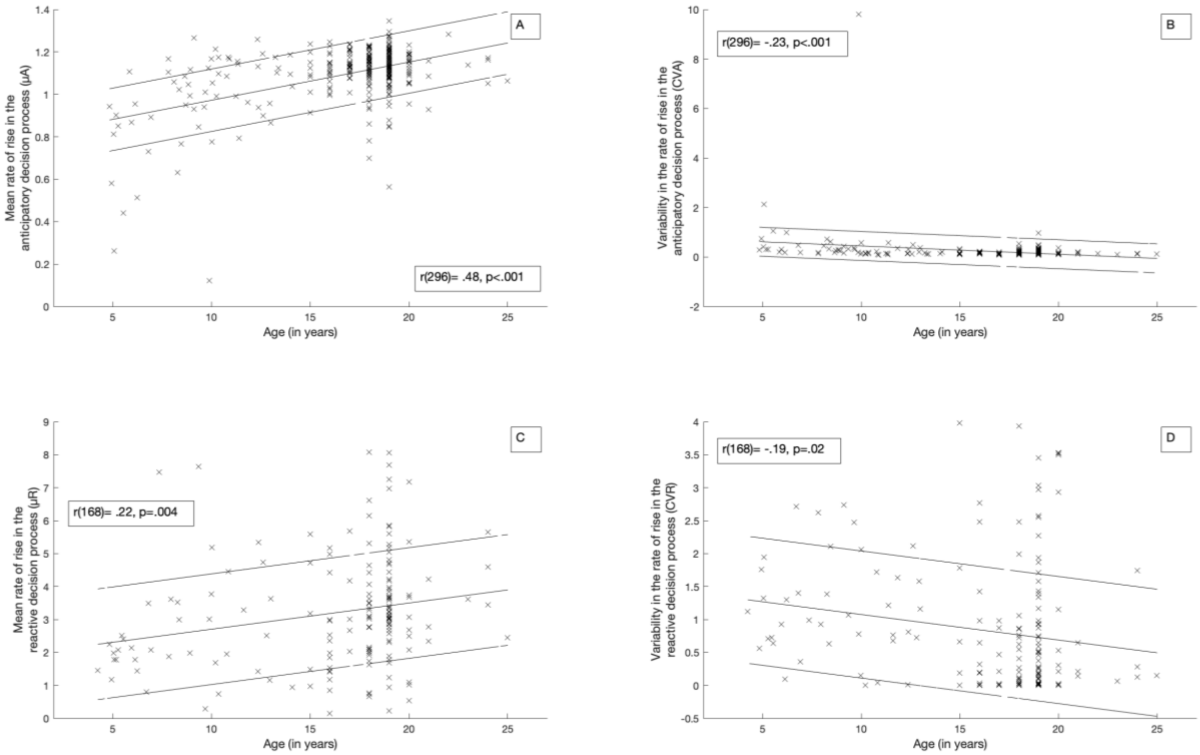
Anticipatory to reactive response ratio as a function of participants' unlogged age



Note. A figure showing the anticipatory to reactive response ratio as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

Figure 4.5.

The μA, CV_A, μR, and CV_R as a function of participants' unlogged age in years



Note. 4.5A. A figure showing the mean rate of rise in the anticipatory response process (μA) as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 4.5B. A figure showing the variability in the rate of rise in the anticipatory decision process (CV_A) as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 4.5C. A figure showing the mean rate of rise in the reactive response process (μR) as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 4.5D. A figure showing the variability in the rate of rise in the reactive decision process (CV_R) as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

Discussion

The purpose of the current study was to determine how predictive motor timing changes with age from childhood through to adulthood, and thereby, extend the findings of past research (e.g., Brown, 2019). To briefly recap, participants aged 4-25 completed 100 trials of a cued RT task. During the task, participants had to respond as soon as a target stimulus appeared. The ratio of anticipatory to reactive responses made by the participant and the average speed and variability of their anticipatory decision process were recorded. These measures showed participants' tendency to use their forward model to predict when to respond in order for their action to coincide with the onset of a target sensory event.

In the current study, it was found that as age increased, the tendency to make anticipatory responses over reactive responses also increased. Similarly, greater age was also associated with a faster and more consistent rate of rise in the anticipatory decision process towards the action execution threshold. These findings support both hypotheses of the current study. Furthermore, the results of the present study extend both the results of past research (Brown, 2019; Perchet & Garcia-Larrea, 2005) and our current understanding of how the forward model develops, as they suggest that the tendency to form forward model predictions about when to execute a motor response increases with age from childhood to adulthood. The current findings also mirror the results of previous literature, which have argued that the rate at which individuals attempt to predict the properties of anticipated stimuli and prepare actions in advance improves with age from childhood to adulthood (e.g., Van Gerven et al., 2016).

In parallel to results obtained for the anticipatory decision process, the present study also found that older age predicted a faster and more consistent rate of rise in the reactive decision process towards the action execution threshold. This suggests that the average speed and consistency at which individuals decide to make a reactive response increases with age. This finding is consistent with past research, which has shown that the average response time to unanticipated stimuli becomes faster with age across ontogeny (Dykiert et al., 2012). Taken together, this suggests that the average speed at which individuals can proactively prepare responses in advance of expected stimuli and reactively formulate appropriate actions in response to unexpected stimuli both improve with age. However, it should be noted that only a small proportion of the variance in participants' anticipatory to reactive response ratio, as well as the speed and consistency of the rate of rise in both their anticipatory and reactive decision processes, was explained by the final models (R^2 = .04 - .25). Hence, the current findings should be interpreted with caution.

In spite of the noted age-related increase in anticipatory response behaviour, it should be acknowledged that, on average, less than a third of the responses made by participants were anticipatory. This figure was significantly lower than the average percentage of anticipatory responses reported by Brown (2019). Strikingly, the task administered in the current study was near identical to the task administered by Brown (2019). The only difference between the two tasks was that the present task ran online, whereas the task reported by Brown (2019) was completed in-person. This raises concern regarding the legitimacy of the current findings, as this indicates that an aspect of the current task may have artificially deterred some participants from proactively preparing their responses. Admittedly, the precise factor that could have manipulated participants' behaviour in this manner is difficult to identify retrospectively. Previous literature has found that participants tend to expect a delay in the speed at which experimental stimuli are presented online compared to offline (Garaizar et al., 2016). Therefore, this expectancy may have caused participants to be slower to respond in the online context than the offline context, resulting in fewer anticipatory responses. Moreover, determining the exact causal precursor to this low rate of anticipatory response behaviour will require future research.

It can also be posited that the results were confounded by the way in which anticipatory responses were defined. In the current study, anticipatory responses were determined in relation to a cut-off criterion, which was equal to the average amber light's duration plus 100ms of the green light's duration. This meant that any responses which occurred during the amber light presentation were classed as anticipatory; they had been prepared in advance of their execution using a forward model. However, it should be acknowledged that a response made whilst the amber light was visible demonstrates a less effective use of the forward model compared to a response executed at the moment when the green light onset. Therefore, to provide a more precise measure of participants' forward model functioning, future research should use stricter criteria to define anticipatory responses, such as classifying only responses made within 100ms of the green light's onset as anticipatory.

Admittedly, it can also be argued that the current results were likely to have been confounded by age-related differences in participants' temporal perception. In order to perform the cued RT task well, it can be argued that participants need to use the amber duration information presented on past trials, as well as the time elapsed since the current amber light onset, to guide their prediction of when to respond (Berniker & Körding, 2011). As noted in chapter 3, prior literature has shown that the acuity with which children can perceive the duration of both visual and auditory stimuli is inferior compared to that of adults (Droit-Volet et al., 2006; Droit-Volet et al., 2007; Zélanti & Droit-Volet, 2012). This indicates that the capacity to judge the temporal duration of a stimulus matures with age from childhood to adulthood. Therefore, it can be argued that the current results were confounded by age-related variation participants' ability to accurately perceive the amber light duration observed on each trial. To improve, the impact of age on temporal perception could be measured, and thus controlled for, in future studies using temporal bisection task as administered in research by Droit-Volet et al. (2007).

To conclude, the current study found that the tendency to prepare a motor response in anticipation of a sensory event increases with age from childhood to adulthood. This extends the findings of past literature, as it suggests that the extent to which individuals use their forward model to guide their actions develops over this period. In order to solidify this conclusion, future research is needed to determine the factor, or factors, that could have hindered the proportion of anticipatory responses that were produced by participants in the current study. Likewise, future research is also warranted in order to elucidate whether the current results can be replicated after controlling for the influence of age-related variation in temporal perception on the findings.

Chapter 5: Exploring the Influence of Age on Task-Switching from Childhood to Adulthood

Chapter Summary

Chapter 5 describes an experiment where the aim was to investigate how task-switching changes with age from childhood to adulthood. To accomplish this, participants aged 5-21 years completed a goal-switching task where they had to switch between performing pro-saccades and anti-saccades according to the colour of the presented stimuli. Switch costs and mixing costs were calculated separately for pro-saccade and anti-saccade trials based on participants' response accuracy. To execute the correct response, participants needed to use the presented stimulus colour to then select the correct action-outcome pairing from prior knowledge. Hence, switch costs demonstrated the cost to accuracy of having to switch between action-outcome pairings when determining how to respond. Whereas, mixing costs showed the cost to accuracy of having to maintain, and select between, these action-outcome associations. Unfortunately, the pro-saccade switch costs, anti-saccade mixing costs, and pro-saccade mixing costs did not significantly differ from zero. Hence, the influence of age on these variables could not be examined. However, it was found that the anti-saccade switch costs reduced with age. This provides further support for the idea that the ability to combine context cues with prior knowledge to form accurate forward model predictions develops with age from childhood to adulthood.

Introduction

The experiment described in chapter 4 revealed that the tendency to form forward model predictions develops with age from childhood to adulthood. Notably, the cued RT task presented in chapter 4 required participants to time their responses to coincide with the onset of a target stimulus. Arguably, this meant that the findings were confounded by age-related variation in participants' temporal perception. Hence, the objective of the current chapter was to overcome the noted limitation of the study outlined in chapter 4 by administering a task where the temporal characteristics of the presented stimuli were irrelevant to the task objective. Ergo, the current chapter extends our knowledge on how the accuracy of one's forward model predictions changes with age.

Task-switching refers to an individual's ability to flexibly shift between two or more different task objectives (Barcelo et al., 2006). Effective task-switching requires that the individual maintain numerous action-outcome pairings within memory, each outlining how a particular stimulus, or category of stimuli, can be interacted with, and to what effect (Koch et al., 2010). Relevant prior knowledge must then be combined with available context cues in order to select the most appropriate action for the current task (Berniker & Körding, 2011). Hence, effective task-switching necessitates intentional, goal-directed behaviour executed via a forward model, as opposed to automatic, habitual responses made without appropriate reference to this stored action-outcome knowledge (Hogarth et al., 2015). Consequently, by examining how the propensity to switch between different tasks changes with age, it is possible to investigate how the ability to form accurate forward model predictions matures.

Task-switching has typically been measured using behavioural tasks where participants are required to switch between two or more conflicting task objectives within a single block of trials (Karayanidis & McKewen, 2021). The trials within these mixed-task blocks are categorised as either repeat trials or switch trials (Kiesel et al., 2010). A repeat trial refers to a trial where the current task objective matches with the objective of the last trial. Whereas, a switch trial refers to instances for which these two objectives differ (Cepeda et al., 2001). Switch costs are then calculated for each participant by finding the difference in error rate, or the time taken to respond, on switch trials relative to repeat trials in mixed-task blocks. Switch costs have been suggested to demonstrate the damage to performance of having to shift between the action-outcome associations stored within memory (Manzi et al., 2011). The lower the switch cost, the greater the efficiency with which an individual can switch between known action-outcome pairings, and thus, act intentionally, avoiding an incorrect, habitual response.

In addition to mixed-task blocks, previous studies have also often included single-task blocks, where the task objective remains static across the trials (Kiesel et al., 2010). Mixing costs can then be obtained by calculating the difference in performance on repeat trials in the mixed-task blocks compared to performance on single-task block trials. In contrast to a switch cost, a mixing cost demonstrates the detrimental impact to performance of having to maintain, and select between, two or more action-outcome pairings (Manzi et al., 2011). The lower the mixing cost, the greater the individual’s ability to select appropriate prior action-outcome knowledge from memory to combine with contextual evidence to determine how best to respond. Moreover, it may be argued that both switch costs and mixing costs provide an effective means through which to quantify an individual’s task-switching skill, and thereby, their ability to supress automatic, habitual responses and author appropriate, intentional actions via the forward model.

Traditionally, past studies have tended to calculate switch costs and mixing costs for both participants’ error rate and their saccade duration (e.g., Reimers & Maylor, 2005). However, in line with previous research (e.g., Papoutsaki et al., 2018; Semmelmann & Weigelt, 2018; Slim & Hartsuiker, 2022), pilot data revealed that the eye-saccade data obtained via the WebGazer software lacked sufficient temporal precision (see discussion for further detail). Hence, saccade duration was not recorded in the final version of the current task. As a result, all switch costs and mixing costs were computed on the basis of participants’ error rates in the present study. For simplicity, the terms “switch cost” and “mixing cost” will now be used exclusively to refer to the costs calculated on the basis of error rates.

Previous studies have predominantly found that children demonstrate higher switch costs and mixing costs than young adults on task-switching paradigms (e.g., Davidson et al., 2006; Kray et al., 2004; Kray et al., 2008). However, only two previous studies have included adolescents within their sample, both of which failed to find conclusive evidence to support the existence of an age-related decline in either switch costs or mixing costs (Manzi et al., 2011; Reimers & Maylor, 2005). Manzi et al. (2011) presented children (9-10), adolescents (13-14), and young adults (20-27) with either three 1s, three 3s, a single 1, or a single 3. Participants had to switch between reporting the quantity and identity of the numbers displayed. Consistent with the idea of an age-related reduction in mixing costs, it was found that adolescents showed higher mixing costs than young adults. However, mixing costs did not differ between children and adolescents, nor children and young adults. Whereas, switch costs did not differ between the three age groups. Similarly, Reimers and Maylor (2005) presented participants aged 10-66 with a series of photographs of faces. Participants had to switch between reporting the gender and emotional expression of the target face. Contrary to past literature, no age-related changes in switch costs or mixing costs were found. Taken together, these findings suggest that the current consensus on the full developmental trajectory of both switch costs and mixing costs remains ambiguous and would benefit from further research.

The Current Study

The purpose of the current study was to determine the impact of age on participants' ability to flexibly switch between different task objectives. To achieve this aim, participants aged 5-21 were instructed to complete the goal-switching task, inspired by the task used by Jung et al. (2015). During the task, participants moved their fixation point to the left or right side of their computer screen in response to the colour and position of a visual stimulus. More specifically, participants made pro-saccades in response to green stimuli; moving their fixation point towards the presented stimulus. Whereas, red stimuli warranted anti-saccades, where participants shifted their gaze away from the stimulus and towards the opposite side of the screen. In two single-task conditions, stimuli colour remained fixed across trials. Whereas, in a mixed condition, both red and green stimuli were presented in a random order. Switch costs and mixing costs were then computed based on participants' error rate in the same manner as described in previous studies (e.g., Manzi et al., 2011). Based on previous literature, two hypotheses were formed, which have been outlined below. If supported by the data, both hypotheses would indicate that task-switching, and thereby, the ability to shift between action-outcome pairings to facilitate effective goal-directed action, improves with age from childhood to adulthood.

1. To test the findings of both Manzi et al. (2011) and Reimers and Maylor (2005) against the contradictory findings from past literature (e.g., Davidson et al., 2006), it was hypothesised that:
    a. Pro-saccade switch costs would be predicted by age, with lower switch costs associated with older age.
    b. Anti-saccade switch costs would be predicted by age, with lower switch costs associated with older age.

2. To test the results of Manzi et al. (2011) and Reimers and Maylor (2005) against contradictory findings from past literature (e.g., Kray et al., 2004; Kray et al., 2008), it was hypothesised that:

a. Pro-saccade mixing costs would be predicted by age, with lower mixing costs associated with older age.
b. Anti-saccade mixing costs would be predicted by age, with lower mixing costs associated with older age.

Method

Design

A mixed factorial design was used in the current study. The between-subjects independent variables were age, sex, and impulsivity. The within-subjects independent variables were trial type (pro-saccade trial or anti-saccade trial) and task type (switch task or repeat task). The four dependent variables were the pro-saccade and anti-saccade switch costs and the pro-saccade and anti-saccade mixing costs obtained from the goal-switching task. Switch costs demonstrate the participant's ability to make appropriate shifts between the different action-outcome pairings stored within memory when predicting how best to respond. Whereas, mixing costs provide an indication of participants' ability to maintain, and select between, two or more competing action-outcome associations. Four additional dependent variables were also recorded to explore the suitability of the goal-switching task as a method of measuring task-switching online in participants of varying ages. These dependent variables were the total number of position errors, the average time to reposition, the average number of fixation errors, and the average time to first fixate.

Participants

100 participants were initially recruited (15 male, 85 female). The age of participants ranged from 5 to 21 years (*M*= 17.7, *SD*= 2.69).

- 75 participants were right-handed, 4 were left-handed and 18 were ambidextrous, as measured by the EHI-SF (see appendix A for a full outline of this measure). 3 participant did not complete the EHI-SF.
- 74 participants were White, 11 were Asian, 2 were Black, and 7 had a mixed/multiple ethnic identity, and 6 did not report their ethnic identity.

Participants were recruited through one of four avenues: 7 participants were recruited through SSM in August 2021, 12 were recruited from two high schools in the Nottinghamshire and Derbyshire areas from January 18th 2022 - 16th March 2022, and the remaining 81 participants were recruited either through RPS or through recruitment posters published on social media from the 9th November 2021 - 18th March 2022. For more detailed information on how participants were recruited from each of these sources and how informed consent was obtained, please see chapter 2. The full experimental procedure of the current study was approved by the School of Psychology ethics committee at the University of Nottingham.

Five participants were excluded because they self-reported a diagnosis of either ADHD, ASD or Obsessive Compulsive Disorder (OCD). The rationale for removing participants with ASD or ADHD was outlined in chapter 3. In terms of participants who reported a diagnosis of OCD, prior research has suggested that those with OCD tend to perceive top-down, interoceptive signals, such as their memory of past events, as unreliable, and thus, underweight the prior relative to the likelihood (Fradkin et al., 2020; Schultchen et al., 2019), resulting in a failure to accurately predict the sensory consequences of their own actions (Gentsch et al., 2012). Therefore, to avoid obscuring the results regarding the neurotypical developmental trajectory of the forward model, these participants were removed from the sample. All remaining participants were neurotypical. A further 11 participants were also excluded because they failed to respond on over 50% of the goal-switching task trials. All participants reported normal vision or agreed to wear contact lenses to correct their vision throughout the task. The demographics of the adjusted sample can be viewed in table 5.1.

Table 5.1. The demographic characteristics of the adjusted sample.

	Age (*years*)	Sex	Ethnicity*	Handedness**
Full sample (*n*=84)	*Range*= 5 to 21 *M*=17.63 *SD*=2.89	13 Male 71 Female	61 White 10 Asian 2 Black 5 Mixed/multiple ethnic identities	65 right-handed 4 left-handed 12 ambidextrous

Note. *Ethnicity information was not collected for 6 participants. **Handedness information was not collected for 3 participants.

Materials

The Goal-Switching Task. The goal-switching task was used to measure participants' ability to adapt their movements in response to changes in the task objective. The design of the task was inspired by the task used in research by Jung et al. (2015). The task was designed using PsychoPy and ran online via Pavlovia (Peirce, 2019). Stimuli consisted of a 200x200 pixel red square and a 200x200 pixel green square. In the child version of the task, each square contained an identical cartoon face. Faces were included in order to promote children's attention.

WebGazer Software. WebGazer is a JavaScript-based library designed to track and record participants' eye movements in real-time via their device's webcam (Papoutsaki et al., 2017). WebGazer is comprised of two elements: a pupil detector and a gaze estimator. The pupil detector is used to identify the location of the participant's pupils within the webcam feed. This is achieved by pinpointing two circular regions which each possess a higher contrast relative to their surrounding area (Papoutsaki et al., 2017). The gaze estimator was used to approximate participants' point of fixation via a regression model. Mouse click events made during a series of calibration trials were used to guide the model's predictions (Papoutsaki et al., 2017). Mouse click events were used based on the assumption that participants' cursor potion and gaze location should align when a mouse click is made (Hauger et al., 2011). Previous studies have reportedly obtained similar results using WebGazer to measure eye-movements online compared to lab-based eye-tracking methods, albeit with greater variance in the spatial and temporal resolution of the collected data (Papoutsaki et al., 2018; Slim & Hartsuiker, 2022).

Self-Report and Parent-Report Measures. The following self-report measures and parent-report measures were administered in the current study:

- The SWAN-Hyperactive/Impulsive subscale was used to measure child participants' impulsivity as reported by their parents or carers.
- The UPPS-P short-form was used to measure self-reported impulsivity in adolescents and young adults.

Please see chapter 2 for a full outline of these two measures.

Procedure

Similar to the synchronisation-continuation task reported in chapter 3, two versions of the goal-switching task were created with the same age groups in mind: one version for adolescents and young adults and a second version for children. The adolescent and young adult version was available to participants via the two high schools, RPS, and social media posts. Whereas, the child version was administered at the two SSM events. As was noted in chapter 3, although SSM was aimed at children aged 4-12, individuals aged between 12-17 were also permitted to take part. Therefore, it was possible for a participant aged 12+ to have received the child version of the task.

Adolescent and Young Adult Procedure. Adolescent and young adult participants provided their demographic information, reported whether or not they needed to wear glasses to correct their vision, and completed the UPPS-P short-form via a survey hosted on Qualtrics before being redirected to the goal-switching task. If participants reported that they needed to wear glasses to correct their vision, they were instructed to wear contact lenses whilst completing the task, and report whether or not it was possible for them to do this. This controlled for impact of light glare from the lenses of the glasses on the recorded point of fixation during the task.

Upon opening the goal-switching task, participants were first presented with an instructions screen with details about how to complete the task. The instructions also asked that participants sit in a brightly lit room, preferably in front of a window or a lamp, and that they remove anything that may be obscuring their eyes, such as their hair. This maximised the likelihood that participants' point of fixation would be recorded accurately. Throughout the task, all written instructions were accompanied by a voice-over which read the instructions to participants. Participants also saw a 320x240 pixel video feed taken from their device's built-in webcam using the WebGazer JavaScript library (Papoutsaki et al., 2017). This was positioned in the top centre of the screen and contained a 200x160 pixel outline of a green square, two thirds the size of the video feed. Throughout the experiment, the position of the participant's face in relation to the square outline was continuously monitored by the WebGazer program. When the participant's face was within the square outline, the outline turned green. If their face left the square outline, the outline's colour changed to red. This helped participants to position themselves in a way such that their eye movements could be detected by WebGazer. Once the participant was satisfied that their face was positioned within the square outline, they could press the spacebar to progress to the next screen. In addition to the position of their face, participants' eye movements were also accessed by the WebGazer program via their device's webcam.

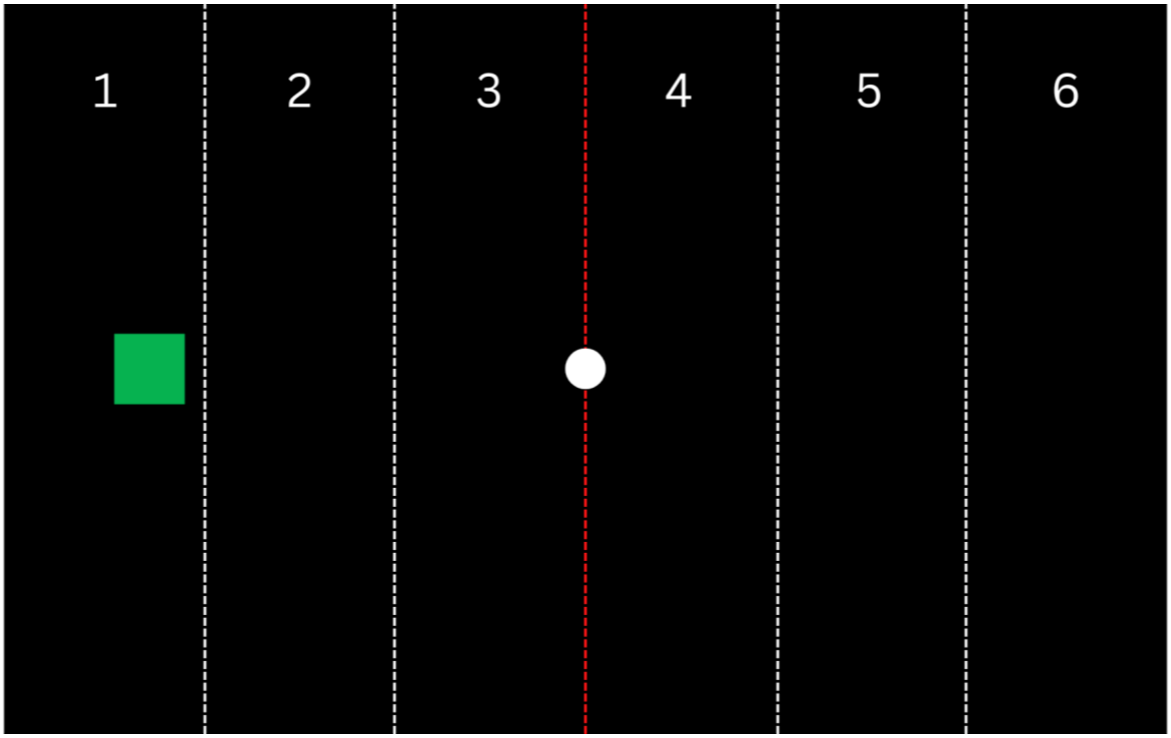
Next, participants completed 34 calibration trials. At the beginning of the calibration trials, the video feed was removed from the screen so as not to distract the participant during the task. In the event that the participant's face exited the square outline, the video feed was revealed again. The video feed then disappeared 3s after the participant's face had re-entered and remained within the square outline. Occasions for which the participant's face left the square outline were defined as position errors and were recorded. In addition, the time taken for the participant's face to re-enter the square outline was also recorded. Both the number of position errors and the time taken to reposition the face assessed how difficult it was for participants to maintain their position in the square outline and reposition themselves after moving outside of the outline. On each calibration trial, a 30x30 pixel white square appeared at a random location on the screen. Participants were instructed to click on each square. After this, participants were presented with a demonstration of the main task and completed 5 practice trials to familiarise themselves with the task. Participants were then asked to report via a keypress whether they believed that their eye-movements had been tracked correctly during the practice trials. If the participant responded no, then the calibration trials were repeated once more before the main experiment trials began. Whereas, if a yes response was given, then the participant progressed immediately to the main experiment trials.

Throughout the task, the screen was divided vertically into 6 invisible zones (see figure 5.1). At the beginning of each trial, a 40x40 pixel white circle was visible in the centre of the screen, on the border of the 3rd and 4th zone. The circle flashed for 300ms. The circle remained onscreen until the participant's fixation point had entered into either the 3rd or 4th zone. The time taken for first fixation to be achieved, i.e., the time taken for the participant's fixation point to first enter the 3rd or 4th zone on each trial relative to the first circle flash, was recorded. In addition, any instance for which the participant's fixation point then left the 3rd or 4th zone again before the target square had appeared was defined as a fixation error and was also recorded. Both the time taken to first fixate and the number of fixation errors serve as an indication of the participant's difficulty in achieving and maintaining fixation during the goal-switching task. If the time to first fixate exceeded 2 seconds, text was displayed below the white circle to encourage the participant to move their point of fixation into the 3rd or 4th zone. After first fixation had been achieved, participants were presented with a target red or green square positioned in either the 1st or 6th zone (see the square shown in figure 5.1 for an example). The probability that the target square would appear in either the 1st or 6th zone was equal and randomised across the trials.

Participants responded by moving their fixation point into either the 1st or 6th zone as quickly as possible. If the target square was green, then participants responded by moving their fixation point into the same zone (1st or 6th zone) as the target square. Whereas, if the target square was red, then participants moved their fixation point into the end zone that did not contain the target square (again, either the 1st or 6th zone). Participants' responses were recorded as either correct or incorrect accordingly. In keeping with previous research (Alahyane et al., 2014), for occasions in which the participant's fixation point initially entered into the 2nd or 5th zone in the incorrect direction, changed direction and then ultimately entered into the 1st or 6th zone in the correct direction, this type of response was logged as incorrect. The reasons for this are two-fold. First, this initial error in the direction of their saccade demonstrates an automatic, habitual response, and thus, an absence of deliberate motor control guided by appropriate prior knowledge (Tanaka et al., 2021). Secondly, although this would show that the participant was successful in correcting the trajectory of their saccade, this correction occurs at a relatively late stage in the saccade's full movement path (i.e., right before the 1st or 6th zone is reached). Therefore, this demonstrates a failure to combine appropriate prior knowledge with the current contextual evidence when predicting how best to respond. All correct responses were awarded one point. The participant's total earned points were displayed in the top right-hand corner of the screen throughout the task. Participants were encouraged to score as many points as possible in order to promote their attention.

Figure 5.1.

Image of the goal-switching task with additional notations.



Note. An image showing the goal-switching task. A green target square is positioned at the left side of the screen within the 1st zone. A white circle is positioned at the centre of the screen. This was used to attract participants' fixation point towards the centre of the screen at the beginning of each trial. The dashed lines and numbers 1-6 indicate each of the 6 invisible zones. The dashed red line represents the centre of the screen. Notably, the dashed lines and numbers were not visible to the participant.

Two types of trial were presented during the task: pro-saccade trials and anti-saccade trials. In pro-saccade trials, only green squares appeared. Whereas, in anti-saccade trials, only red squares were shown. In addition to this, the task contained 3 conditions, a pro-saccade condition, an anti-saccade condition, and a mixed condition. Each condition was presented in two separate blocks with the order of the blocks counterbalanced between participants. The pro-saccade condition featured only pro-saccade trials. The anti-saccade condition contained only anti-saccade trials. Whereas, the mixed condition included an equal number of both pro-saccade and anti-saccade trials and the presentation order of each trial type was randomised. All trials in the mixed condition were further divided into switch trials and repeat trials. A switch trial referred to a trial in which the colour of the target square on the current trial differed from the colour of the target square on the previous trial. Whereas, a repeat trial referred to a trial in which the current colour of the target square matched the colour of the target square on the previous trial. To clarify, this resulted in four possible trial-task type combinations for trials in the mixed condition: pro-saccade switch, pro-saccade repeat, anti-saccade switch, and anti-saccade repeat. There were equal numbers of each trial-task type combination in the mixed condition. Each trial lasted for approximately 1s. If no response was made 2s after the target had appeared, text was displayed that reminded participants of the task instructions. Break screens were shown every 80 trials for an unlimited amount of time. Participants pressed the spacebar when they were ready to continue with the task. The break screens included encouraging text to promote participants' attention. An inter-trial interval was presented for 1s, during which a black screen was displayed. There were 240 trials in total, 80 per condition, and 40 per block. The full procedure lasted approximately 15-minutes.

Child Procedure. The procedure completed by children mirrored that of the adolescent and young adult participants with a few minor exceptions. Child participants' parents or carers provided their child's demographic information and completed the SWAN when registering their child for SSM. Once registered, participants could complete the goal-switching task at any time throughout the SSM event. Participants were encouraged to complete the left-hand vs right-hand task before completing the goal-switching task to measure their hand preference during daily tasks. For a full description of the left-hand vs right-hand task, please see appendix B. At the beginning of the goal-switching task, participants were asked to report via a keypress whether or not they needed to wear glasses to correct their vision. If the participants gave a yes response, they were asked to wear contact lenses throughout the task and report whether or not it was possible for them to do so. During the task, an image of a sticker was presented for an unlimited time every 10 trials as a reward to promote children's attention. Break screens were presented every 20 trials rather than every 80 trials. There were 120 trials in total, 40 per condition, and 20 per block. The full child procedure lasted approximately 5-minutes.

Data Analysis

All data pre-processing procedures were conducted in MATLAB and all statistical analyses were run in SPSS. Participants' ages were converted to natural logarithmic age, as described in chapter

2. Any trials on the goal-switching task for which no response was recorded were removed from the analysis. On average 17% of trials were removed per participant (*SD* = 12.27, *range* = 0.42 - 48.33). Among participants who reported the need to wear glasses to correct their vision (61% of the sample), all stated that they would wear contact lenses whilst completing the study.

Self-Report and Parent-Report Scales. All self-report and parent-report scales were summed and averaged to create an index for each of the variables of interest. In addition, SWAN-Hyperactive/Impulsive subscale scores and UPPS-P Short-form scores were converted to z-scores.

Error Rate. An error rate was calculated for all trials in the pro-saccade condition, all trials in the anti-saccade condition, and all trials of each trial-task type combination in the mixed condition. This was achieved by dividing the number of incorrect responses by the total number of trials within each condition, or within each trial-task combination type in the case of the mixed condition trials. This provided an indication of the extent to which the participant struggled to appropriately adjust, or maintain, the congruency between their response and the current trial's objective.

Switch Costs. To measure how changes in the task objective between consecutive trials influenced participants' ability to direct their gaze correctly, switch costs were calculated for each trial type resulting in a pro-saccade switch cost and an anti-saccade switch cost for each participant. The pro-saccade switch cost was produced by finding the difference in error rate between the pro-saccade repeat trials and the pro-saccade switch trials in the mixed condition. Whereas, the anti-saccade switch cost was obtained by subtracting the error rate for the anti-saccade repeat trials from the error rate for the anti-saccade switch trials in the mixed condition. For clarification, on switch trials, the target gaze location altered between successive trials. Whereas, on a repeat trial the target gaze location remained unchanged. Hence, by comparing the error rate on switch vs repeat trials, a switch cost effectively demonstrates the cost to response accuracy that occurred when the task objective was changed between trials. Lower switch costs demonstrate a better ability to switch between known action-outcome pairings, and hence, use their forward model to act intentionally and avoid making an erroneous, habitual response.

Mixing Costs. In addition to switch costs, mixing costs were also calculated for each trial type. The pro-saccade mixing cost was calculated by subtracting the error rate for the repeat trials on the pro-saccade condition from the error rate for the pro-saccade repeat trials on the mixed condition. Whereas, the anti-saccade mixing cost was obtained by subtracting the error rate for the repeat trials on the anti-saccade condition from the error rate for the anti-saccade repeat trials on the mixed condition. To clarify, in the pro-saccade condition and anti-saccade condition, only one colour of square was presented, and hence, only one type of response was required throughout the condition (i.e., always look towards the square, or always look away from the square). Whereas, in the mixed condition, both green and red squares were shown, meaning that participants had to frequently change their method of response. Hence, a mixing cost provides an indication of the participant's ability to maintain, and select between, two known action-outcome associations within a single condition.

Total Number of Position Errors. The total number of occasions for which the participant's face left the square outline throughout the experiment was recorded. This measure shows how difficult it was for participants to maintain a position where they were in view of the webcam. It is important to note that position errors were not bound by the temporal constraints of a single trial as it was possible for several trials to pass before the participant's face returned to the square outline. Hence, the total number of position errors was analysed as averaging across trials would not produce a meaningful result.

Average Time to Reposition. The average time taken for a participant to re-enter the square outline was found by dividing the total time taken for the participant to reposition their face within the square outline by the total number of position errors made. Similar to the total number of position errors, the average time taken to reposition also shows how difficult it was for participants to position themselves within view of the webcam.

Average Number of Fixation Errors. The average number of fixation errors made by each participant was calculated by dividing the total number of fixation errors by the total number of trials completed. This measure shows how difficult it was for participants to hold their fixation point at the centre of the screen at the start of each trial.

Average Time to First Fixate. The average time to first fixate on the white circle cue at the beginning of each trial was calculated for each participant by dividing the total time taken to first fixate by the total number of trials. This measure is useful as an excessive average time to first fixate may indicate an issue with the calibration of the eye-tracking and therefore, provide an insight into the suitability of the task as an online measure of task-switching. It should be noted that, due to limitations in WebGazer's spatial resolution (Slim & Hartsuiker, 2022), the precise point of fixation cannot be known. As a result, the recorded time taken to move the fixation point to the 3rd or 4th zone of the screen may be somewhat inaccurate and any results based on this measure should be interpreted with caution.

Peak Velocity and Saccade Duration. It may have been informative to measure the peak velocity and saccade duration of participants' saccades, as achieved in prior task-switching studies which utilised eye-tracking techniques in a lab-based setting (e.g., Jung et al., 2015). Through this, it would have been possible to gain an additional measure of the extent to which changes in the task objective hindered participants' task-switching performance. Initially, an attempt was made to measure both the peak velocity and duration of participants' saccades on each trial. However, there was an observable spatial and temporal offset between the fixations and saccades approximated by WebGazer and participants' genuine eye-movements, both of which are detrimental to the accurate calculation of peak velocity and saccade duration.

This observation of poor temporal and spatial acuity was consistent with previous studies that also employed WebGazer to record participants' saccades online (e.g., Papoutsaki et al., 2018; Semmelmann & Weigelt, 2018; Slim & Hartsuiker, 2022). Whilst it has been argued that WebGazer can determine the approximate area of the screen to which the fixation point is located (Semmelmann & Weigelt, 2018), it has also been noted that it lacks the spatial precision required to track the distance travelled by a participant's saccade (Slim & Hartsuiker, 2022). For instance, Slim & Hartsuiker (2022) reported an average offset between participants' actual and estimated fixation point of approximately 30% of their computer screen size. Similarly, whilst lab-based experiments have reported that a saccade to a target stimulus should take approximately 200ms to perform (e.g., Matin et al., 1993), Semmelmann and Weigelt (2018) found that saccades recorded via WebGazer lasted 450–750ms on average. Therefore, it was concluded that the data acquired through WebGazer was not suitable for calculating peak velocity and saccade duration. Consequently, both measures were ultimately removed from the final versions of the task.

Outlier Detection. To exclude any potentially confounding anomalous data, the Tukey's fences method for outlier detection was applied to the dependent variables (Tukey, 1977). It was revealed that 1 pro-saccade switch cost, 2 anti-saccade switch cost, 1 pro-saccade mixing cost, 1 anti-saccade mixing cost, 5 total number of position errors, 7 average time to reposition, and 2 average time to first fixate data points were more than 1.5 interquartile ranges away from the nearest quartile. Upon comparison, it was found that the removal of the anomalous data points did not affect the direction or significance of the results regarding the influence of the predictor variables on the anti-saccade switch cost, the anti-saccade mixing cost, the total number of position errors or the average time to reposition. Hence, the identified data points for these variables were not removed in order to ensure the completeness of the data. However, excluding the extraneous data points for the pro-saccade switch cost, the pro-saccade mixing cost, and the average time to first fixate did have an effect on the findings.

Prior to outlier removal, the pro-saccade mixing cost was significantly predicted by both age (β = .24, t = 2.3, p = .02) and sex (β = .24, t = 2.28, p = .03). Whereas, the variance in the pro-saccade switch cost and the average time to first fixate was not significantly explained by any of the predictor variables. After the outliers were removed, age significantly predicted the pro-saccade switch cost (β = .27, t = 2.54, p= .01) and the average time to first fixate (β = -.23, t = -2.1, p= .04). However, the pro-saccade mixing cost was not predicted by any of the predictor variables after outliers were excluded. As a result of the observed changes, the 4 data points identified as outliers for these three variables were removed from the data so as not to statistically bias the results.

Results

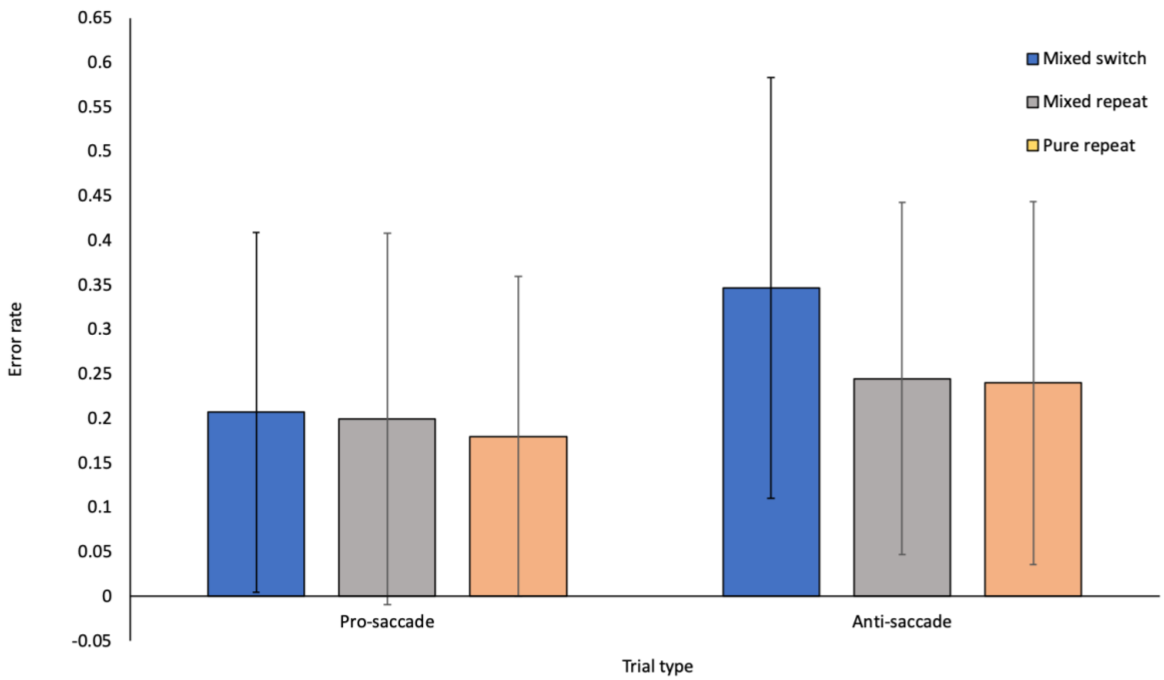
In the current study, participants completed a goal-switching task to measure their task-switching ability. On each trial of the task, participants were presented with a red or a green square on the left or right of the screen. When a green square was presented, participants were instructed to look towards the side of the screen that contained the square. Whereas, when a red square was visible, participants were required to look towards the side of the screen that did not contain the square.

Task Performance

On average, participants successfully looked towards the correct side of the screen on 77.54% of trials (*SD* = 17.74, *range* = 37.23 - 98.2). This performance level was significantly different from chance ($t(83) = 14.23$, $p<.001$, $d= .18$), suggesting that participants understood the task instructions. To verify that a sufficient number of errors were made across each trial-task type combination in the mixed condition, as well as within each of the two pure conditions, six one sample t-tests were conducted. G*Power analysis revealed that 27 participants were required to obtain a medium sized-effect ($d= .5$) in a one sample t-test with 80% power and a 5% alpha level (Faul et al., 2009). Therefore, the analyses were sufficiently powered. Figure 5.2 shows average error rate as a function of trial-task type and condition.

Figure 5.2.

Average error rate as a function of trial-task type and condition



Note. A figure showing the average error rate for each trial-task type combination in the mixed condition, in addition to the error rate for each pure condition. Error bars represent +/- 1 standard deviation. The blue and grey bars indicate the average error rate for the switch and repeat trials within the mixed condition, separated by trial type. Whereas, the orange bars show the average error rate for the trials in the two pure conditions, i.e., the pure pro-saccade condition and pure anti-saccade condition.

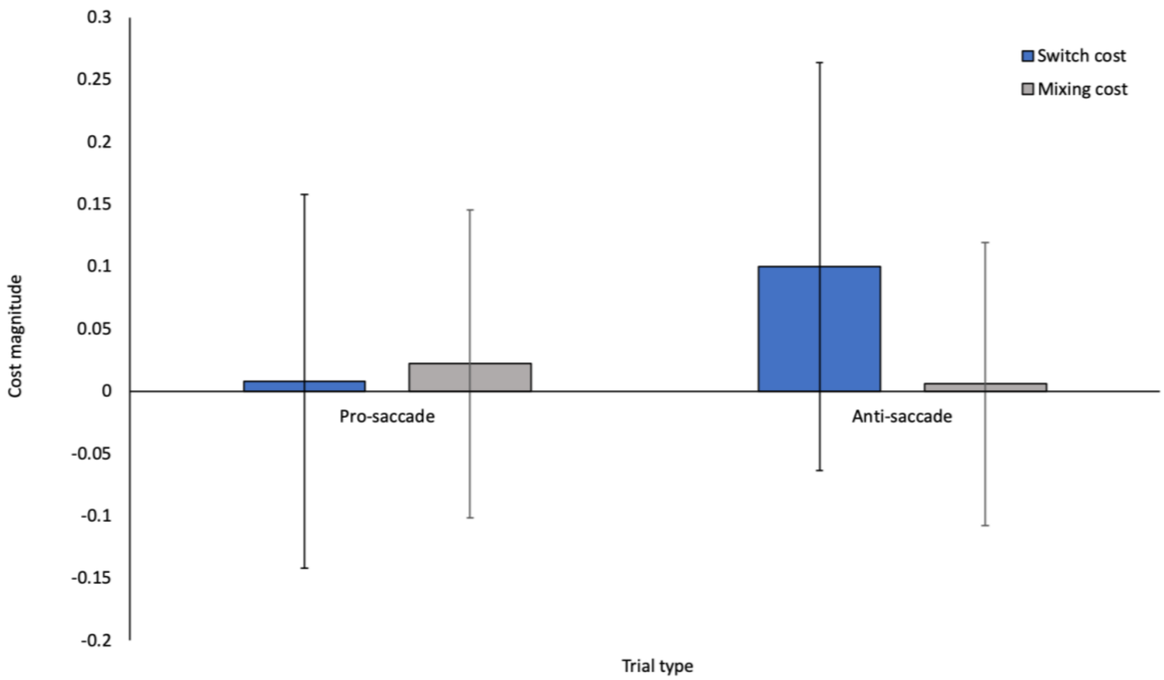
Within the mixed condition, error rates for the pro-saccade switch trials ($t(83)= 9.39$, $p<.001$, $d= .2$), pro-saccade repeat trials ($t(83)= 8.75$, $p<.001$, $d= .21$), anti-saccade switch trials ($t(83)= 13.46$, $p<.001$, $d= .24$), and the anti-saccade repeat trials ($t(83)= 11.33$, $p<.001$, $d= .2$) were all found to significantly differ from zero. Similarly, the error rate for trials in the pure pro-saccade condition ($t(83)= 9.17$, $p<.001$, $d= .18$) and pure anti-saccade condition ($t(83)= 10.8$, $p<.001$, $d= .2$) also significantly differed from zero. As a brief reminder, switch costs were calculated for each trial type by subtracting the error rate for the repeat trials from the error rate for the switch trials within the mixed condition. Whereas, mixing costs were found for each trial type by subtracting the error rate for the repeat trials on a pure condition from the error rate for the repeat trials on the mixed condition. Taken together, this suggests that there were a sufficient number of errors made to progress with the analysis.

Switch Costs and Mixing Costs as a Function of Trial Type

To determine whether there was a significant difference in switch costs and mixing costs between the two trial types, two paired sample t-tests were conducted. This revealed the extent to which the current task adequately replicated the goal-switching tasks implemented within previous research (e.g., Reimers & Maylor, 2005). As a reminder, switch costs quantified the extent to which participants struggled to shift between the two action-outcome pairings when predicting how best to respond, and thereby, suppress an incorrect, habitual response. Whereas, mixing costs revealed participants' ability to select between, and maintain, two action-outcome associations. It was expected that the anti-saccade switch cost and anti-saccade mixing cost would be higher than the pro-saccade switch cost and pro-saccade mixing cost. This is because individuals tend to reflexively direct their case towards incoming stimuli. Hence, an automatic pro-saccade response must be supressed in favour of a more intentional anti-saccade response on the anti-saccade trials, thus creating greater opportunity for error (Jung et al., 2015). In addition to this, the anti-saccade trials in the current task featured a spatial incongruity between the side of the screen to which the target stimulus was positioned and the side of the screen to which participants needed to look. As a result, the anti-saccade trials were more cognitively demanding than pro-saccade trials, where this spatial incongruity was absent. G*Power analysis revealed that 27 participants were required to obtain a medium sized-effect (d=.5) in a paired samples t-test with 80% power and a 5% alpha level (Faul et al., 2009). Therefore, the analyses were sufficiently powered. Figure 5.3 shows switch costs and mixing costs as a function of trial type.

Figure 5.3.

Switch costs and mixing costs as a function of trial type



Note. A figure showing switch costs and mixing as a function of trial type. Error bars represent +/- 1 standard deviation.

It was found that the anti-saccade switch cost was significantly higher than the pro-saccade switch cost, $t(82)= -3.9$, $p<.001$, $d= .22$. However, there was no significant difference between the pro-saccade mixing cost and the anti-saccade mixing cost, $t(82)= 1.08$, $p= .29$, $d= .15$.

The Extent to Which the Switch Costs and Mixing Costs for Each Trial Type Significantly Differed from Zero

As seen in figure 5.3, the means for both switch costs and mixing costs appeared to be relatively small. To explore this observation further, four one-sample t-tests were conducted. These tests revealed the extent to which the pro-saccade switch cost, anti-saccade switch cost, pro-saccade mixing cost and anti-saccade mixing cost differed from zero. Anti-saccade switch costs were found to be significantly different from zero, $t(83)= 5.62$, $p<.001$, $d= .16$. This suggests that a greater number of errors were made on anti-saccade switch trials compared to anti-saccade repeat trials within the mixed condition. Whereas, pro-saccade switch costs ($t(82)= .49$, $p= .63$, $d= .15$), anti-saccade mixing costs ($t(83)= .48$, $p= .63$, $d= .11$) and pro-saccade mixing costs ($t(82)= 1.63$, $p= .11$, $d= .12$) did not significantly differ from zero. This suggests that there was no difference in the number of errors made on pro-saccade switch trials and pro-saccade repeat trials in the mixed condition, nor any difference in the number of errors made on mixed condition trials and pure condition trials for either trial type. Given that the pro-saccade switch costs, anti-saccade mixing costs, and pro-saccade mixing costs did not significantly differ from zero, the influence of age on these variables will not be examined in the current study.

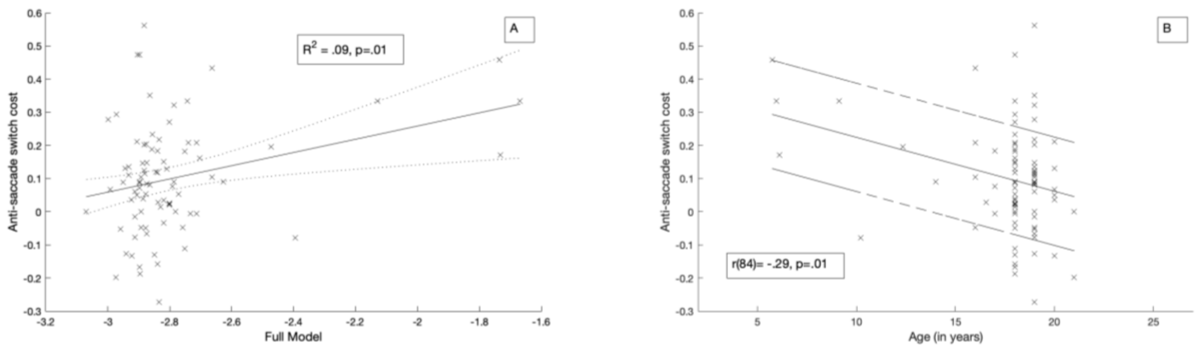
The Influence of Age on Participants' Anti-Saccade Switch Costs

To investigate the extent to which the variance in anti-saccade switch costs could be explained by age, a stepwise multiple linear regression was conducted. Impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block. The purpose of this test was to reveal the extent to which an individual's age influences their ability to combine appropriate action-outcome knowledge with available context cues when the respective influences of both impulsivity and sex are taken into account. G*Power analysis revealed that 77 participants were required to obtain a medium sized-effect (f^2=.15) in a multiple linear regression with three predictor variables, 80% power, and a 5% alpha level (Faul et al., 2009). Therefore, as the sample contained 84 participants, the analyses were sufficiently powered.

The rationale for including impulsivity and sex as nuisance variables in the current regression analysis was outlined in chapter 2. In further support of the inclusion of impulsivity as a nuisance variable, previous research has argued that higher levels of impulsivity are associated with a reduced ability to inhibit automatic, habitual responses (Bari & Robbins, 2013). Hence, impulsivity was included as a nuisance variable to control for the potential impact of age-related variation in impulsivity on the current findings. In regard to sex, past research has suggested that females demonstrate greater efficiency in deploying proactive control over their actions compared to males (Bianco et al., 2020; Yücel et al., 2012; Smittenaar et al., 2015). Furthermore, it has been argued that this is due to sex differences in the rate at which regions of the brain implicated in cognitive control develop (Christakou et al., 2009). Therefore, sex was also included as a nuisance variable within the analyses. Figure 5.4A shows anti-saccade switch cost plotted against a model of age, impulsivity and sex, whereas, figure 5.4B shows anti-saccade switch cost plotted against participants' unlogged age in years for visualisation purposes.

Figure 5.4.

Anti-saccade switch cost plotted against both a model of age, impulsivity and sex and participants' unlogged age in years.



Note. 5.4A. A figure showing anti-saccade switch cost plotted against a model of age, impulsivity and sex. Error bars represent +/- 1 confidence interval. 5.4B. A figure showing anti-saccade switch cost as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

A significant model was found, $F(1,81) = 7.64$, $p = .01$. It was revealed that the anti-saccade switch cost was significantly predicted by age ($\beta = -.29$, $t = -2.76$, $p = .01$), and was not predicted by impulsivity ($\beta = .06$, $t = .54$, $p = .55$) or sex ($\beta = .03$, $t = .23$, $p = .82$). As age increased, the anti-saccade switch cost decreased. The overall model fit was $R^2 = .09$ ($SE = .16$).

The Influence of Age on the Total Number of Position Errors, the Average Time Taken to Reposition, the Average Number of Fixation Errors, and the Average Time Taken to First Fixate

To establish the suitability of the goal-switching task for use with participants of different ages, the total number of position errors, the average time taken to reposition, the average number of fixation errors, and the average time taken to first fixate were recorded. More specifically, these variables examined whether participants could maintain their position in view of the webcam and hold their fixation point at the start of a trial without the presence of a researcher to give continued instruction. Four simple linear regressions were then conducted to investigate the extent to which age could explain the variance in each of these four variables (see table 5.2 for the final models). G*Power analysis revealed that 55 participants were required to obtain a medium sized-effect (f^2=.15) in a simple linear regression with 80% power and a 5% alpha level (Faul et al., 2009). As the sample contained 84 participants, the analyses were sufficiently powered. It was revealed that the total number of position errors, the average number of fixation errors, and the average time taken to first fixate were significantly predicted by age (all p<.05). As age increased, the total number of position errors, the average number of fixation errors, and the average time taken to first fixate decreased. Unlike the other three variables, the average time taken to reposition was not significantly predicted by age (p>.05). For visualisation purposes, see figure 5.5 for each variable plotted against participants' unlogged age in years.

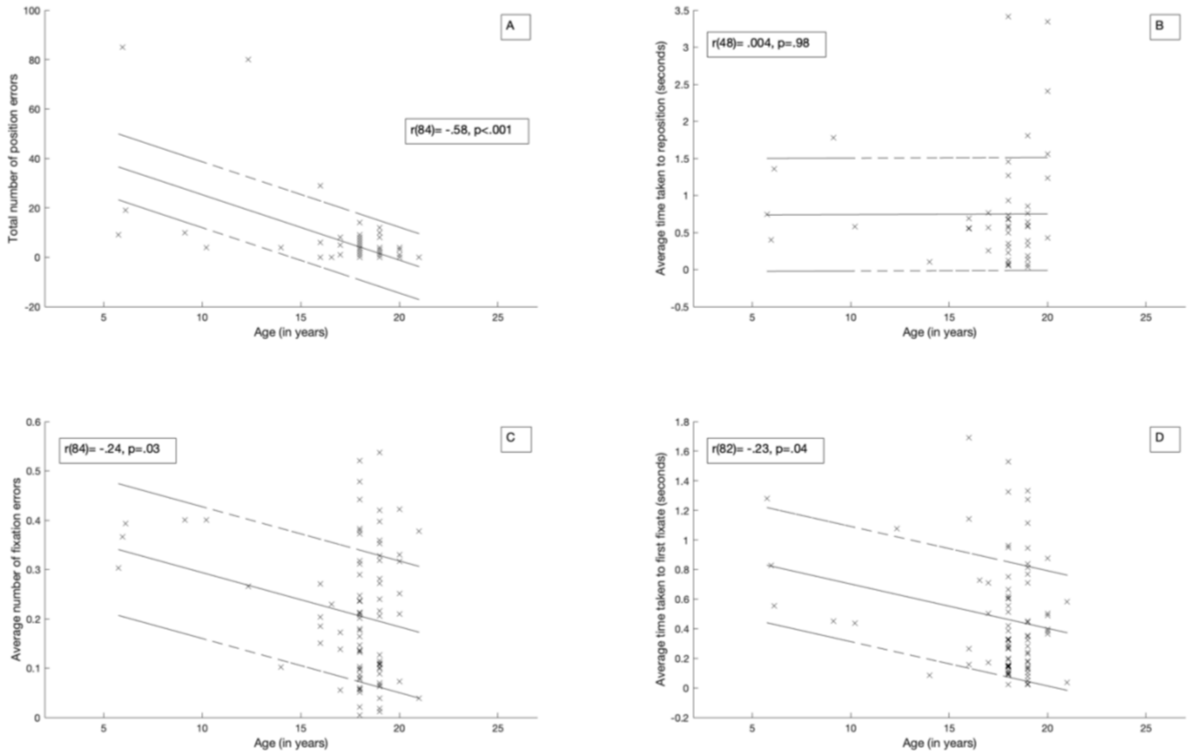
Table 5.2. The results of four simple linear regression analyses investigating the influence of age on the total number of position errors, the average time taken to reposition, the average number of fixation errors, and the average time taken to first fixate

Final regression model	*β*	*t*	*F*	*df*	R^2	*SE*
The total number of position errors			39.02**	1, 82	.32	11.05
Age	-.57**	-6.25**				
The average time taken to reposition						
Age		No variables were entered into the model.				
The average number of fixation errors			5.71*	1, 82	.07	.13
Age	-.26*	-2.39*				
The average time taken to first fixate			4.41*	1, 80	.05	.38
Age	-.23*	-2.1*				

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

Figure 5.5.

The total number of position errors, the average time taken to reposition, the average number of fixation errors, and the average time taken to first fixate as a function of participants' unlogged age in years



Note. 5.5A. A figure showing the total number of position errors as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 4.5B. A figure showing the average time taken to reposition as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 4.5C. A figure showing the average number of fixation errors as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 4.5D. A figure showing the average number of fixation errors, and the average time taken to first fixate as a function of participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

Discussion

The purpose of the current study was to determine how the ability to maintain, and switch between, different action-outcome pairings according to the current task objective changes with age from childhood to adulthood. To briefly recap, participants aged 5-21 completed a goal-switching task where they switched between looking towards green square stimuli and looking away from red square stimuli. To execute the correct response, participants needed to use the presented stimulus colour to then select the correct action-outcome pairing from prior knowledge. To that end, switch costs and mixing costs were recorded for both the pro-saccade and anti-saccade trial types. Switch costs measured the extent to which participants failed to shift between action-outcome pairings when selecting their responses. Whereas, mixing costs provided an index of the cost to response accuracy of having to maintain, and select between, different action-outcome associations within memory.

In agreement with both the second sub-hypothesis (1b) of the current study and previous literature (e.g., Davidson et al., 2006), it was found that the variance in the anti-saccade switch cost was explained by participants' age, even after accounting for impulsivity and sex. More specifically, older age was associated with a lower anti-saccade switch cost, and therefore, a better ability to flexibly switch from a pro-saccade response method to an anti-saccade response method. This result is also consistent with past developmental literature. Alahyane et al. (2014) demonstrated that the number of errors accrued when switching between pro- and anti-saccades declined with age across children (8-12) and adolescents (13-17), before plateauing in young adulthood (18-25). This age-related change was associated with developmental improvements in the activity of the fronto-parietal network (Alahyane et al., 2014); regions of the brain implicated in proactive motor control and the inhibition of interference from irrelevant action-outcome pairings (Cooper et al., 2015; Jamadar et al., 2010). Therefore, this further supports the idea that task-switching improves with age as children transition to young adulthood. Moreover, the current findings extend our prior understanding of forward model development, as they suggest that the capacity to form accurate forward model predictions improves with age from childhood to adulthood. That being said, it should be acknowledged that only a small proportion of the variance in participants' anti-saccade switch cost was explained by the final model (R^2 = .09). Hence, the results should be interpreted with caution.

Despite the noted utility of the current results in extending our knowledge of how the forward model develops, it should be acknowledged that the findings do not align with the results of previous task-switching studies which also included adolescents within their sample (e.g., Manzi et al., 2011; Reimers & Maylor, 2005). In contrast to the current results, both Manzi et al. (2011) and Reimers and Maylor (2005) reported that there was no difference in the switch costs accrued by children, adolescents and young adults when asked to shift between two competing tasks. It is difficult to specify the precise reason for this disparity without engaging in mere speculation. Notably, the tasks used by Manzi et al. (2011) and Reimers and Maylor (2005) differed from the current task in two key ways. First, in research by Manzi et al. (2011) and Reimers and Maylor (2005), participants switched between performing different actions in response to the same set of stimuli. Whereas, in the current task, green stimuli were always responded to with pro-saccades and red stimuli always warranted anti-saccades. Second, in the tasks used by Manzi et al. (2011) and Reimers and Maylor (2005) participants responded via keypresses. Whereas, participants in the current study responded by directing their gaze to the left and right of the screen. Speculatively, it may be argued that these methodological differences could have led to inconsistencies in how cognitively demanding each task was to complete; thereby, resulting in the noted disparity in the effect of age across the studies.

Unfortunately, the pro-saccade switch costs, pro-saccade mixing costs and anti-saccade mixing costs were found to be statistically close to zero in the present study. The low pro-saccade switch costs suggest that switching from an anti-saccade to a pro-saccade task objective incurred no significant cost to response accuracy relative to instances where the pro-saccade objective remained active over consecutive trials. Similarly, the low mixing costs suggest that response accuracy was not hindered by having to maintain, and select between, two different response options within a single condition. As a result of these low costs, it was not possible to reliably test the influence of age on these three variables (hypotheses 1a, 2a, or 2b). Thus, the present study cannot draw any concrete conclusions on how the ability to select between different action-outcome pairings when predicting how to best respond changes with age.

It is unclear as to why the pro-saccade switch costs and mixing costs were low, especially given that previous lab-based eye-tracking studies have reported sizeable pro-saccade and anti-saccade switch costs and mixing costs in neurotypical adults (Jung et al., 2015). Speculatively, one reason for this may be that the 2-second time window for responses allowed participants sufficient time to override any incorrect, automatic responses before executing their actual response. This may then have deflated switch costs and mixing costs across participants. Future research is needed to empirically assess whether the response time limit or any other specific aspect of the current task could have caused these low switch and mixing costs to occur.

Admittedly, the present study possessed a number of limitations. For instance, due to the method in which gaze locations were approximated by WebGazer, it is likely that the accuracy of participants' recorded fixation positions reduced over time as the trials progressed. In order to match pupil locations with their corresponding screen coordinates, WebGazer functions under the assumption that the position of the cursor and location of an individual's gaze will align when an intentional cursor click is made (Papoutsaki et al., 2018). Therefore, successive clicks are used to continuously calibrate WebGazer's estimations against the participant's genuine gaze positions. However, as the time elapsed since the last mouse click grows, the likelihood that a spatial offset will arise between the record screen coordinates and genuine gaze location increases (Papoutsaki et al., 2018). In the current study, cursor clicks were only recorded during the calibration trials presented at the beginning of the goal-switching task. This undermines the reliability of the current findings, as it suggests that the accuracy with which the direction of participants' saccades were recorded deteriorated over time. To ensure that a precise calibration is maintained, future iterations of the goal-switching task should include additional calibration trials presented at regular intervals between the main experimental trials, as employed by Slim and Hartsuiker (2022). Alternatively, the webcam feed could be recorded and later reviewed by two independent coders in order to verify that the direction of participants' saccades was measured correctly, as achieved in research by Scott and Schulz (2017).

Aside from technical issues introduced by the manner in which WebGazer was utilised, it may be argued that participants' ability to adhere to the task instructions also distorted the robustness of the eye-tracking data. The present study revealed that participants' ability to maintain their position in view of the webcam and hold their gaze at the centre of the screen at the start of each trial improved with age. Removing one's gaze from the screen, or one's body from the webcam view, would have disrupted the calibration between the true gaze locations and the recorded screen coordinates (Papoutsaki et al., 2018). Therefore, this raises further concern regarding the legitimacy of the collected error rate data. This then further supports the idea that future research should include more frequent calibration trial blocks to ensure that the alignment between genuine and measured eye-positions is maintained. Furthermore, to improve younger participants' engagement with the task and thereby, prevent position and fixation errors from occurring, future research should modify the presented stimuli to be more appealing to children. For example, Semmelman et al. (2017) repeatedly flashed an image of a cartoon monkey in the centre of the screen at the beginning of each trial in order to attract children's attention to that location. Hypothetically, this would likely have been a more effective approach to maintaining young participants' attention compared to the white dot used in the current study.

In addition to concerns related to measurement accuracy, the composition of the sample should also be acknowledged as a limitation of the current study. The recruited participants were fairly homogenous in age (see appendix D for a visualisation of the age spread within the final sample). Over two thirds of the sample (72.62%) were aged between 18-19, whilst only 17.86% of participants were aged under 18, including nine adolescents (aged 13-17), and six children (aged 5-12). This lack of variance in participant age raises concern over the robustness of the present findings, as it may be argued that the height of the slope in the regression model was driven by only a small number of younger participants (see figure 5.4). Therefore, the findings should be interpreted with caution. As noted in the Covid-19 impact statement, the government imposed restrictions on face-to-face research and teaching introduced in response to the Covid-19 pandemic meant that there were few opportunities available for recruiting child and adolescent participants. That being said, in comparison to the other three studies reported in this thesis, the present study recruited the lowest number of participants aged under 18. This suggests that a specific aspect of the current task deterred younger participants from taking part.

Speculatively, parents may have felt discouraged from allowing their child to complete the goal-switching task due to privacy concerns. Indeed, previous developmental studies which have used webcams to collect data online have cited parents' privacy concerns as a potential reason for low participant recruitment relative to similar lab-based studies (e.g., Semmelmann et al., 2017). Given that parents' privacy concerns were not measured in the current study, it was not possible to test whether this had any sizeable effect on child and adolescent engagement with the research. Notably, mitigations were made to prevent this occurrence, such as emphasising to parents and participants via written correspondence that no video recordings would be taken during the task. However, it would also be beneficial for future researchers to arrange school-based talks designed specifically to answer any ethical and data-privacy concerns that participants and their parents may have, as implemented in previous studies (e.g., Rait et al., 2015).

To conclude, the purpose of the current study was to determine how task-switching develops from childhood to adulthood. This provided an indication of how the capacity to combine relevant prior knowledge with current contextual information to form veridical outcome predictions matures with age. Whilst it was not possible to examine the impact of age on participants' mixing costs, it was found that switch costs declined with age for the anti-saccade trials. This suggests that the ability to flexibly shift between learned action-outcome associations when predicting how best to respond improves from childhood to adulthood. Notably, confidence in the reliability of the current findings was undermined by the suggested lack of spatial acuity in the recorded gaze locations and the heterogeneity of participants' ages within the recruited sample. Therefore, future research is needed to verify whether the similar findings can be obtained using a task with better spatial precision and a wider range of participant ages.

Chapter 6: Establishing How the Ability to Make Appropriate Modifications to Action-Outcome Knowledge Changes from Adolescence to Adulthood

Chapter Summary

The purpose of chapter 6 was to establish how the ability to appropriately update action-outcome knowledge in response to post-action feedback changes with age from adolescence to adulthood. To achieve this, participants aged 14-24 completed an outcome learning task where they were instructed to select between two prize boxes with the goal of obtaining a reward. In a stable condition, the probability that each box would deliver a reward remained fixed. Whereas, in a volatile condition, the reward probabilities shifted between the two boxes over time. The extent to which participants' updated their choice behaviour in response to recent trial outcomes was indexed via a learning rate. Notably, it is advantageous to have a lower learning rate within a stable context than in a volatile context. Hence, participants needed to alter their learning rate according to the volatility of the current context. Strikingly, it was revealed that learning rate was not influenced by age in either condition. However, the difference in learning rate between conditions was found to increase with age. This suggests that the magnitude to which individuals can tailor their action-outcome knowledge appropriately for the current context refines with age from adolescence to adulthood.

Introduction

The studies described in chapters 4 and 5 demonstrated that the accuracy with which individuals can predict the most probable outcome of their actions via a forward model improves with age from childhood to adulthood. Taken together, the findings from both studies address the first goal of this thesis. The current chapter will then focus on the second thesis goal; how the ability to make appropriate modifications to current action-outcome knowledge in response to post-action feedback develops with age.

Outcome learning refers to an individual's ability to update their knowledge of the action-outcome contingencies present within their environment in response to post-action feedback (Kawato & Wolpert, 2007). The external world is an inherently dynamic space with multiple entities interacting with, and thereby modifying, their surroundings simultaneously (Wolpert &

Flanagan, 2001). As a result, the exact outcome of an action can rarely be predicted with absolute certitude (van Beers et al., 2002). Therefore, in order to achieve desired action goals, an individual must maintain an up-to-date understanding of the probabilistic associations present between actions and their effects and use this knowledge to guide their behaviour (Wolpert & Ghahramani, 2000). To ensure that this knowledge is maintained, it is believed that an estimate of the feedback observed after a self-generated action is incorporated into their prior estimate (Körding & Wolpert, 2006). Subsequent actions are then guided by this revised prior knowledge of the probabilistic relationship between the performed action and the expected effect (Berniker & Körding, 2011). Evidently, the ability to update to one's constructed forward model is imperative to effective goal-directed action.

Crucially, to interact appropriately with the environment, it is not sufficient to indiscriminately update the prior to the same extent in response each action outcome (Hohwy, 2017). Instead, the rate at which amendments are made to learned action-outcome associations must be modulated according to the relative volatility of current context (Gershman, 2015). A stable context refers to a scenario in which the probabilistic relationships between actions and effects remain fixed over time. Any unexpected outcomes can then be attributed to chance within this context (Behrens et al., 2007). For instance, if there was a fixed 90% probability that action A would result in outcome A, there would be a 10% chance that the unexpected outcome B could occur instead. In comparison to a stable environment, the probabilistic links tying specific actions to particular effects are subject to frequent change within a volatile context (Browning et al., 2015). For instance, in a volatile environment where action A was first predominantly associated with outcome A, an abrupt alteration to this probabilistic relationship could occur such that, action A would be more strongly related to outcome B rather than outcome A. Therefore, an unexpected outcome observed within a volatile context should be interpreted as evidence that the underlying action-outcome contingencies have changed and a behaviour change is required.

The extent to which the prior estimate is modified in response to an observed outcome is determined by the individual's current learning rate (Eckstein et al., 2022). From a Bayesian perspective, the learning rate refers to the time scale for which past outcomes are integrated to form the prior and thereby, used to determine the individual's next action (Hohwy, 2017). Learning rates can vary between 0 and 1. The closer the learning rate is to 1, the greater the extent to which the prior was constructed from only the most recent outcomes, as opposed to a wider history of observed feedback. Whereas, a comparably lower learning rate indicates that the opposite is true (Eckstein et al., 2019).

When a low learning rate is employed within a stable context, each new outcome is incorporated into the prior incrementally over time. As a result, the prior gradually becomes a more reliable estimate of the stable action-outcome mappings (Jacobs & Kruschke, 2011). Hence, the ability of surprising, yet rare, outcomes to cause substantial changes in current action-outcome knowledge and subsequent choice behaviour decreases over time (Berniker & Körding, 2011). In other words, the effect of a detected discrepancy between the expected and observed outcome on current action-outcome knowledge diminishes as the volume of past feedback accumulated increases. Moreover, using a relatively low learning rate is advantageous within a stable context as successive priors, i.e., conceptualisations of the fixed action-outcome structure, will become more veridical over time, leading to more optimal choice behaviour (Browning et al., 2015).

Whilst it has been argued that incorporating a range of past outcomes into one's prior is beneficial within a stable context, it has also been suggested that this practice can be detrimental to performance within a volatile context (Behrens et al., 2007). This is because recent outcomes are more informative of the dynamic action-outcome contingencies active within the volatile context compared to outcomes that occurred further in the past (Browning et al., 2015). Therefore, updating the prior incrementally via a low learning rate can impair an individual's ability to correctly perceive these rapid-changing action-outcome pairings. Whereas, possessing a high learning rate decreases the amount of past outcome evidence integrated within the prior (Hohwy, 2017). This then enables unexpected outcomes observed in a volatile context to trigger larger, and more prompt, updates to learned action-outcome mappings than those seen in a stable context. Ergo, this suggests that a high learning rate is more beneficial than a low learning rate within a volatile environment (Browning et al., 2015). Moreover, an individual's ability to appropriately update their forward model can be determined by examining the degree to which their learning rate alters according to the volatility of current context.

Evidently, the ability to tailor the rate at which surprising outcomes shape behaviour in accordance with the volatility of the current context is crucial for daily life (Hohwy, 2017). Despite this, a paucity of studies have examined how this ability develops from adolescence to adulthood (DePasque & Galván 2017). In addition, of the studies which have explored how this ability develops, all have focused exclusively on determining how behaviour is modified either within a stable context (e.g., van den Bos et al., 2012) or within a volatile context (e.g., Hauser et al., 2015) alone. Thus far, no studies have compared the degree to which individuals of varying ages are able to adjust their learning rate in stable contexts relative to volatile contexts. It is crucial to examine how outcome learning develops differently between the two contexts, as this goes beyond merely demonstrating that the individual can inflate or reduce the volume of past outcomes used to construct the prior, but instead compares the degree to which they can adaptively modulate their learning rate appropriately for each context.

In comparison to adults, it has been argued that adolescents are less able to appropriately update their behaviour in response to the action-feedback observed within stable contexts (Decker et al., 2016; Xia et al., 2021). This assertion has been made on the basis that adolescents have previously been shown overestimate contextual volatility on probabilistic learning tasks with stable action-outcome contingencies, resulting in both elevated learning rates and poorer choice accuracy relative to adults (van den Bos et al., 2012; Jepma et al., 2020). This suggests that, within a stable context, adults update their action-outcome knowledge at a more gradual pace than adolescents, thus resulting in superior performance. Consistent with this idea, it has also been shown that unexpected outcomes observed in a stable context have a greater influence on adolescents' choice behaviour than that of adults (Barash et al., 2019; Hartley & Somerville, 2015; van Duijvenvoorde et al. 2013). Therefore, these findings suggest that the ability to maintain an accurate understanding of the action-outcome contingencies present within a stable context and adjust choice behaviour accordingly improves with age from adolescence to adulthood.

Although adolescents' reported tendency to overestimate contextual validity may not be ideal for a stable environment, a tendency to employ a high learning rate would be beneficial for a volatile environment (Behrens et al., 2007). Consequently, it has been argued that adolescents' outcome learning skills surpass those of adults in volatile contexts (Gopnik et al., 2017).

Adolescence has previously been characterised as a developmental period uniquely equipped to tolerate action-outcome uncertainty (Lourenco & Casey, 2013) due to heightened plasticity in the brain (Larsen and Luna, 2018) and more frequent exposure to novel contexts (Somerville et al., 2017) compared to adults. In line with this idea, Hauser et al. (2015) administered a probabilistic reversal learning task to adolescents aged 12-16 and adults. During the task, action-outcome contingencies covertly switched each time a target number of correct choices had been made. It was found that adolescents tended to have a higher learning rate than adults, meaning that they learned from unexpected outcomes at a faster rate. This suggests that adolescents' ability to appropriately adapt their behaviour for a volatile context declines with age as they transition to adulthood.

When directly compared, the results of past research, such as van den Bos et al. (2012) and Hauser et al. (2015), imply that adolescents and adults do not alter their approach to updating the prior between contexts of varying levels of volatility. In both stable and volatile contexts, adolescents appear to show a tendency to integrate only the most recent outcomes into their prior. Whereas, adults appear to gradually integrate a wider range of past outcomes into their prior over a longer timeframe, irrespective of contextual volatility. Evidently, this suggests that an age-related dissociation in outcome learning exists between the different contexts. Adults appear to be better than adolescents at updating their forward model within stable contexts, where a lower learning rate is optimal. Whereas, adolescents are better able to update their internal model within volatile contexts, where a higher learning rate is more beneficial.

Contrary to the argument that adolescents employ a higher learning rate than adults in volatile contexts, more recent research has argued that the ability to respond to sudden changes in action-outcome contingencies actually improves with age from adolescence to adulthood (Eckstein et al., 2022). In contrast to the findings of Hauser et al. (2015), Eckstein et al. (2022) reported that learning rate increased from adolescence to adulthood on a probabilistic learning task with dynamic action-outcome mappings. This suggests that the degree to which individuals can reduce the volume of past outcomes integrated into the prior in response to frequent action-outcome association changes improves across adolescence. Furthermore, when combined with the results from other learning studies (e.g., van den Bos et al., 2012), this suggests that the ability to both lower one’s learning rate in stable contexts and inflate one’s learning rate in volatile contexts improves as adolescents age to adulthood. Hence, when comparing learning rate between the two contexts, one would expect the difference in learning rate to increase with age.

The Current Study

The purpose of the current study was to establish the impact of age on participants’ ability to adapt the extent to which their decisions are guided by past action-outcomes according to the volatility of the current context. To fulfil this aim, participants aged 14-24 completed a rewarded learning task, based on the task administered by Browning et al. (2015). On each trial of the task, participants were presented with two boxes and were instructed to select the box that they believed to contain a reward. In the stable condition, the relative probability that each choice would lead to a reward outcome was fixed. Whereas, in the volatile condition, these probabilities shifted between the two boxes over time. Participants’ learning rate was calculated for each condition to quantify the extent to which their choice behaviour was modified in response to the most recent trial outcomes. Subsequently, the difference in learning rate between the two conditions demonstrates the degree to which participants could adapt their learning in response to fluctuations in the contextual volatility. Based previous literature, three hypotheses were formed, as outlined below.

1) To test adults’ proposed superiority over adolescents in adapting their behaviour to stable action-outcome contingencies (e.g., van den Bos et al., 2012; Jepma et al., 2020), it was hypothesised that learning rate would be predicted by age within the stable condition, with greater age associated with a lower learning rate.

2) To test the findings of Eckstein et al. (2022) against the contradictory results of Hauser et al. (2015), it was hypothesised that learning rate would be predicted by age within the volatile condition, with greater age associated with a higher learning rate.

3) To test the idea that adults are better able to adjust their behaviour in response to changes in environmental volatility than adolescents (van den Bos et al., 2012; Eckstein et al., 2022), it was hypothesised that the difference in learning rate would be predicted by age, with greater age associated with a larger difference in learning rate between conditions.

Method

Design

A mixed factorial design was used in the current study. The independent variables were age, sex and impulsivity. The within-subject independent variable was the relative volatility of the two outcome probabilities presented on each condition of the outcome learning task. For both the stable and volatile conditions, four dependent variables were obtained: the learning rate, the variability in the learning rate, the inverse decision temperature, and the variability in the inverse decision temperature. Additionally, the difference in the each of these four dependent variables between the two conditions was also calculated. To clarify, this resulted in 12 dependent variables in total. The learning rate demonstrates the extent to which the participant's choices were influenced by more recent action outcomes compared to the outcomes of actions made further in the past. Thus, the learning rate provided an indication of the extent to which the most recent trial outcome caused a substantial change in the current prior estimate.

In comparison to the learning rate, the inverse decision temperature indicated the extent to which a participant's choices were guided by an understanding of the relative advantage of selecting one choice option over its alternative. Hence, the inverse decision temperature demonstrated participants' ability to maintain an up-to-date forward model regarding the probabilistic relationships linking each option to each outcome and apply this knowledge when forming decisions on which option to select. A relatively low inverse decision temperature indicated that a participant's actions were more akin a series of random choices rather than considered selections based on the knowledge learned from past outcomes. Moreover, the findings gained in regard to the impact of age on the inverse decision temperature will attempt to replicate the results gained from the previous empirical chapters presented in the current thesis, i.e., the extent to which the ability to use and sustain a forward model improves with age. In contrast, the results obtained in regard to age-related changes in the learning rate will move a step beyond this by revealing the extent to which participants correctly modified the weight attributed to recent over past action outcomes when formulating their choice decisions.

Participants

339 participants were initially recruited (49 male, 290 female). The age of participants ranged from 14 to 24 years (*M*=18.4, *SD*=1.04).

- 285 participants were right-handed, 20 were left-handed and 34 were ambidextrous, as measured by the EHI-SF (see appendix A for a full outline of this measure).
- 273 participants were White, 34 were Asian, 12 were Black, 17 had mixed/multiple ethnic identities, and 3 had another ethnic identity that was not listed.

Thirty six participants were recruited from two high schools in the Nottinghamshire and Derbyshire areas between January 18th 2022 - 16th March 2022. The rest of the sample were recruited either through RPS or through recruitment posters published on social media from the 26th April 2021 - November 17th 2021. For more detailed information on how participants were recruited from each of these sources and how informed consent was obtained, please see chapter 2. In exchange for volunteering to take part in the experiment, participants had the opportunity to be entered into a prize draw to win an Amazon voucher. In addition, in order to motivate participants to fully engage with the task goals, participants were informed that an additional Amazon voucher would be awarded to the participants who scored the highest, second highest and third highest number of points during the outcome learning task. The full experimental procedure of the current study was approved by the School of Psychology ethics committee at the University of Nottingham.

Eighteen participants were excluded because they self-reported a diagnosis of either ASD, ADHD or OCD. The rationale for removing participants with ASD or ADHD was outlined in chapter 3, and the rationale for excluding those with OCD was detailed in chapter 5. All remaining participants were neurotypical. One further participant was also excluded for failing to make a response on over 50% of the trials on the outcome learning task. The demographics of the adjusted sample can be viewed in table 6.1.

Table 6.1. The demographic characteristics of the adjusted sample.

	Age (*years*)	Gender	Ethnicity	Handedness
Full sample	*Range*=	48 Male	258 White	270 right-handed
(*n*=320)	14 to 24	272 Female	32 Asian	19 left-handed
	M= 18.4		11 Black	31 ambidextrous
	SD= 1.02		16 Mixed/multiple ethnic identities	
			3 Any other ethnic identity	

Materials

The Outcome Learning Task. The outcome learning task was used to measure participants' ability to update their forward model in response to observed trial outcomes. The task was designed using PsychoPy and ran online via Pavlovia (Peirce, 2019). Stimuli consisted of a 300x300 pixel image of a red prize box and a 300x300 pixel image of a blue prize box.

The UPPS-P Short-Form. The UPPS-P short-form was used to measure self-reported impulsivity. Please see chapter 2 for a full outline of this measure.

Procedure

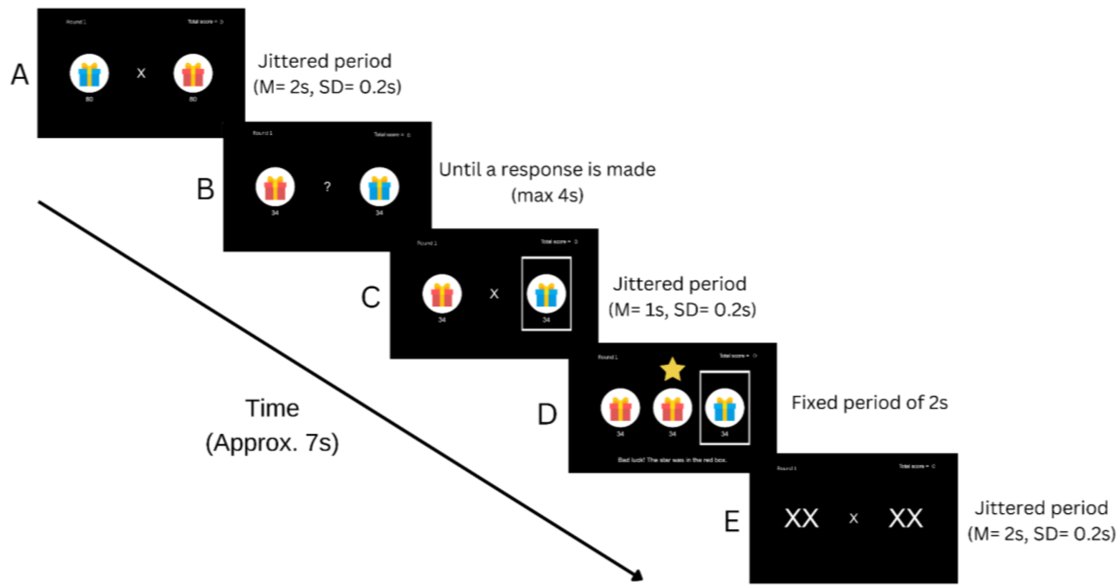
Participants provided their demographic information and completed the UPPS-P short-form via a survey hosted on Qualtrics. Following this, participants were redirected to the outcome learning task. Upon opening the task, participants saw a black instructions screen with details on how to complete the task. Participants were also instructed to complete the task using the same hand throughout. Written instructions were also accompanied by a voiceover which read the instructions aloud for the participant. On the next screen, participants viewed a demonstration of the task and listened to a voiceover which explained each stage of the demonstration. Each screen of the instructions and the demonstration contained a red square which changed to green after the voiceover audio clip had ended. Only after pressing the green squares would the task progress to the next screen. This ensured that participants could not skip the instructions or the demonstration. Participants then completed 5 practice trials to familiarise themselves with the task, before progressing to the main experiment trials.

On each trial, a red prize box and a blue prize box were presented on the left and right of the screen (see figure 6.1). The probability that each coloured box would appear on the left or right of the screen was equal and randomised across the trials. Each box also had a number of points displayed directly beneath it. On each trial, one of the two coloured boxes was the correct box to choose. In exchange for selecting the correct box, participants received the number of points that was displayed below that box. However, if the incorrect box was chosen, then participants were awarded zero points on that trial. The number of points below each box were equal on each trial and were selected randomly with replacement from an array of values ranging from 1-99.

At the beginning of each trial, a white fixation cross was visible in the centre of the screen for a jittered time period (*M*= 2s, *SD*= 0.2s). The fixation cross was then replaced by a white question mark, which signified that the participant could now respond. The question mark remained onscreen until either the participant had made a response or 4s had passed. Participants responded by clicking on one of the two prize boxes as quickly as possible. Participants' responses were recorded as either correct or incorrect accordingly. Participants were encouraged to score as many points as possible in order to promote participants' attention throughout the task. If participants failed to respond within 4s of the question mark cue's onset, text was displayed which reminded participants of the task instructions. After the question mark had disappeared, a white box appeared around the prize box that the participant had selected and the question mark was replaced by a fixation cross. After a jittered time period (*M*= 1s, *SD*= 0.2s), the correct prize box was displayed for 2s alongside its corresponding points value. During this time, an image of a star moved upwards from the correct prize box to create the illusion that the star had been contained within the box. An inter-trial interval was then presented for a jittered time period (*M*= 2s, *SD*= 0.2s), during which two white capital Xs and the white fixation cross were displayed. Each trial lasted for approximately 7s. Break screens were shown every 60 trials for an unlimited amount of time. Participants could end the break and continue with the task at any time by clicking a green square displayed onscreen. The break screens included encouraging text to promote participants' attention and information about how far they had progressed through the task.

Figure 6.1.

Image of the trial structure for the outcome learning task.



Note. An image showing the trial structure for the outcome learning task. Each black box depicts a specific moment of the trial. The diagonal arrow represents the order in which each moment was experienced, i.e., beginning with 6.1A and ending with 6.1E. 6.1A. The red and blue prize boxes are presented to the left and right of the screen. Points values are positioned below each box. The white fixation cross shown in the centre of the screen indicates that no responses can be made. 6.1B. The question mark shown in the centre of the screen indicates that responses can now be made. 6.1C. A white square outline surround the box that was selected by the participant. 6.1D. The correct box is shown in the centre of the screen. Text at the bottom of the screen informs the participant of the trial outcome. 6.1E. Two masks are shown to the left and right of the screen, alongside the fixation cross.

The task contained two conditions, a stable condition and a volatile condition. In the stable condition, the probability that one of the two coloured prize boxes would be the correct box remained constant at 75%. For instance, there may have been a 75% chance of the red prize box being the correct box and a 25% chance of the blue prize box being the correct box throughout the condition. Whereas, in the volatile condition, the probability that one of the two coloured prize boxes would be the correct box shifted from 80% to 20% every 20 trials. For example, there may have been an 80% chance of the red prize box being the correct box and a 20% chance of the blue prize box being the correct box for the first 20 trials. For the next 20 trials, this would then switch such that selecting the blue prize box had an 80% chance of resulting in a correct response and selecting the red prize box had an 20% chance of resulting in a correct response. The order in which the conditions were presented was counterbalanced between participants. There were 180 trials in total, 90 per condition. The full procedure lasted approximately 30 minutes.

Data Analysis

All data pre-processing procedures were conducted in MATLAB. All participants' ages were converted to natural logarithmic age, as described in chapter 2. Any trials for which no response was made were removed from the analysis. On average, 2.19% of trials were removed per participant (*SD*= 3.44, *range* = 0 - 29.44).

Self-Report Scales. All self-report scales were summed and averaged to create an index for each of the variables of interest.

The Combined Rescorla-Wagner and Softmax Action Selector Model. To evaluate the frequency at which participants updated their forward model in response to each new trial outcome, an estimated learning rate (α) was calculated for both the stable and volatile conditions. In order to change the learned association between selecting a particular box and it's most probable outcome, an individual must first recognise a disparity between the expected outcome and the observed consequence; that is, the observed consequence must be perceived as surprising by the individual. Subsequently, the individual must then decide that, in order to ensure that their comprehension of the world is accurate, it is necessary to adjust their understanding of the relationship between the action and its outcome in response to this new, surprising outcome (Browning et al., 2015). Hence, an α represents the extent to which a participant's choices were reliant on the outcomes of recent actions relative to the outcomes of actions that occurred further in the past. High αs indicate that the individual's choices were largely guided by the outcomes of their most recent actions compared to those of actions committed on more antecedent trials. That is to say, the discrepancy between an expected and actual outcome on a given trial had a substantial effect on the modification of any subsequent outcome predictions and choices made. Whereas, a lower α demonstrates that unexpected outcomes caused only minimal changes to successive outcome predictions and choice behaviour.

To calculate an α for the stable and volatile conditions, the choices made by the participant in each of the conditions were fitted to a Rescorla-Wagner model, which was in turn, attached to a softmax action selector, in keeping with the data analysis procedure used in research by Browning et al. (2015). The process used to fit participants' choice data to the combined Rescorla-Wagner and softmax action selector model was as follows: first, forty potential α values were generated, in keeping with Browning et al. (2015). These values were spaced equally from log(0.01) to log(1). Using the Rescorla-Wagner model (see equation 6.1), the outcome probability, i.e., the probability of the red box being the correct choice, was calculated for each possible α value. In equation 6.1, the notation, $r_{(i+1)}$, indicates the estimated outcome probability on the next trial, i.e., the probability that selecting the red box will lead to a reward outcome. $r_{(i)}$ denotes the estimated outcome probability on the current trial. α refers to the learning rate, and $\varepsilon_{(i)}$ represents the error in the participant's prediction of the outcome on the current trial.

$$r_{(i\mp1)} = r_{(i)} + \alpha\varepsilon_{(i)} \tag{6.1}$$

Next, the estimated value of each colour choice was calculated for each trial by multiplying the outcome probability values and the actual reward magnitude of each colour option. This was achieved using equation 6.2 and 6.3, modified from Browning et al. (2015). In equations 6.2 and 6.3, $g_{\ red(i+1)}$ and $g_{\ blue(i+1)}$ signify the estimated values of the red and blue stimuli. Whereas, $f_{\ \boldsymbol{red}(\boldsymbol{i}+\mathbf{1})}$ and $f_{\ blue(i+1)}$ represent the actual reward magnitudes for the two stimuli. $r_{(i+1)}$ refers to the probability that the red box was the correct choice and $1 - r_{(i+1)}$ shows the probability that the blue box was the correct choice. As the points values did not differ between the two colour options on each trial, this calculation essentially weighted each points value presented to participants by the underlying outcome probabilities. For instance, in the event that the red box had an 80% chance of leading to a reward outcome, whereas the blue box had a 20% chance of leading to a reward, and both choices were worth 10 points, then the value of selecting the red box would be 8 (0.8 x 10) and the value of selecting the blue box would be 2 (0.2 x 10).

$$g_{\ red(i+1)} = r_{(i+1)} \times f_{\ \boldsymbol{red(i+1)}} \tag{6.2}$$

$$g_{\ blue(i+1)} = 1 - r_{(i+1)} \times f_{\ \boldsymbol{blue(i+1)}} \tag{6.3}$$

The relative advantage of selecting the red box rather than the blue box on each trial was then obtained by subtracting the estimated value of the blue box from the estimated value of the red box. For instance, in the above example, the relative advantage of selecting the red box over the blue box would be 6 (8 - 2). Evidently, positive advantage values indicated that it was better to have chosen red compared to blue box on the current trial. Whereas, negative advantage values signified that the opposite was true.

Following this, the probability that a participant would choose either the red or blue box on each trial was calculated using equation 6.4, as outlined in Browning et al. (2015). Again, $g_{\ (red)}$ and $g_{\ (blue)}$ refer to the estimated value of the red and blue boxes, respectfully. Hence, the notation, $g_{\ (red)} - g_{\ (blue)}$, denotes the relative advantage of choosing the red box over the blue box. $P(choice = red)$ represents the probability that the participant would choose the red box on a given trial. The probability of the blue box being chosen on a trial was given by 1-$P(choice = red)$. Finally, β denotes the inverse decision temperature.

$$P(choice = red) = \frac{1}{1 + exp(-\beta(g_{\ (red)} - g_{\ (blue)})} \tag{6.4}$$

The β was an estimate of the extent to which knowledge of the relative advantage of selecting red over blue, and hence, an understanding of the relative outcome probabilities of each option, had an impact on a participant's choice behaviour. The higher the β value, the greater the extent to which the participant's choices were guided by an understanding of the relative advantage of selecting one box over the other. Comparatively, the lower the β value the greater the extent to which the participant's choices appeared to be made at random, irrespective of the relative advantage offered by either box. Thirty β values were generated, each equally-spaced between log(1) to log(100). The probability that the participant selected each box was then computed for each of the 30 possible β values.

To clarify, the combined Rescorla-Wagner and softmax action selector model contained two free parameters: the learning rate (α) and the inverse decision temperature (β). For each participant, the two parameters were estimated separately for the stable and volatile conditions. A joint posterior probability density function was created for both parameters based on the likelihood that each potential parameter value fit to the participant's actual choice behaviour. Estimates of the two parameters, and the corresponding SD of those estimates (σ_α and σ_β) were then calculated as the anticipated value and SD of a marginal probability density function over each potential parameter value, calculated through direct integration. Given that the two free parameters were used to multiply other values in the equations described above, all statistical analyses were conducted on the logarithms of the parameter estimates, in keeping with Browning et al. (2015).

The difference in learning rate and inverse decision temperature between conditions

As the learning rate and inverse decision temperature parameters were multiplicative, the difference in these parameters between the two conditions was calculated as the difference of their log values, consistent with Browning et al. (2015). More specifically, the difference in learning rate ($\Delta\alpha$) between the two conditions was calculated by subtracting the log value for the stable condition α from the log value for the volatile condition α. This revealed the extent to which participant's choice behaviour was more strongly influenced by recent trial outcomes than more antecedent trial outcomes on the volatile condition relative to the stable condition. In the volatile condition, only the 20 most recent trials provide outcome probability information useful to the current choice decision. Hence, participants will achieve better performance in the volatile condition if they base their choice decisions on the outcomes of more recent trials compared to the outcomes of trials that occurred further in the past. Consequently, if the participant was able to adaptively base their decisions on recent trial outcomes more so in the volatile condition than in the stable condition, then their learning rate would be higher in the volatile than the stable condition, resulting in a positive $\Delta\alpha$ value. Ergo, the more positive a participant's $\Delta\alpha$ value, the greater their ability to regulate the extent to which their forward model was updated in light of new outcome information in accordance with the relative volatility of the current context.

In addition to the $\Delta\alpha$, the difference in the variability of the learning rate ($\Delta\sigma_{\alpha}$) between the two conditions was also calculated by subtracting the log value for the stable condition σ_{α} from the log value for the volatile condition σ_{α}. This demonstrated the impact of the contextual volatility on the consistency with which participants' based their decision on more recent trials relative to more anterior trials. A more positive $\Delta\sigma_{\alpha}$ suggests that a participant's α was more consistent in the volatile condition relative to the stable condition. In contrast, a more negative $\Delta\sigma_{\alpha}$ would indicate that the opposite was true, with greater α consistency in the stable condition over the volatile condition.

The difference in inverse temperature ($\Delta\beta$) and the the difference in the variability of the inverse temperature ($\Delta\sigma_{\beta}$) were also calculated in the same manner as the $\Delta\alpha$ and the $\Delta\sigma_{\alpha}$ by subtracting the stable condition log values from the volatile condition log values. The $\Delta\beta$ shows the extent to which participants' choice behaviour was guided by an understanding of the outcome probabilities to a greater degree in one condition than the other. A more positive $\Delta\beta$ indicates that advantage knowledge had a greater impact on the participant's choices in the volatile condition than in the stable condition. Whereas, a more negative $\Delta\beta$ is reflective of the opposite. Finally, the $\Delta\sigma_{\beta}$ shows the impact of contextual volatility on the consistency with which participants made choices informed by advantage knowledge, as opposed to merely random selections. Once again, a more positive $\Delta\sigma_{\beta}$ suggests that the participant's β was more consistent in the volatile condition relative to the stable condition, whereas, a more negative $\Delta\sigma_{\beta}$ indicates that the reverse was true.

Outlier Detection. To exclude any potentially confounding anomalous data, the Tukey's fences method for outlier detection was applied to the log values for each DV. For the stable condition, 12 σ_{α}, 3 β, and 11 σ_{β} datapoints were found to be more than 1.5 interquartile ranges away from the nearest quartile. Whereas, for the volatile condition, 24 σ_{α}, 1 β, and 17 σ_{β} datapoints were more than 1.5 interquartile ranges away from the nearest quartile. Finally, 16 $\Delta\alpha$, 48 $\Delta\sigma_{\alpha}$, 15 $\Delta\beta$, 9 $\Delta\sigma_{\beta}$ datapoints were more than 1.5 interquartile ranges away from the nearest quartile.

Upon comparison, it was found that the removal of the anomalous data points did not affect the direction or significance of the results regarding the influence of age, impulsivity and sex on the σ_{α} or β for either condition, nor on the $\Delta\sigma_{\alpha}$, $\Delta\beta$, or the $\Delta\sigma_{\beta}$. Hence, the identified data points for these variables were not removed in order to ensure the completeness of the data. However, excluding the extraneous data points for the σ_{β} in both conditions and the $\Delta\alpha$ did have an effect on the findings. Prior to outlier removal, the σ_{β} in both conditions and the $\Delta\alpha$ were not predicted by any predictor variable. Whereas, the σ_{β} was significantly predicted by impulsivity in both the stable (β = .13, t= 2.26, p= .02) and volatile (β = .12, t= 2, p= .046) conditions and no other predictor variable (all p>.05) after outlier removal. Similarly, $\Delta\alpha$ was significantly predicted by age (β = .13, t= 2.29, p= .02) and no other predictor variable (all p>.05) after the outlier data points were excluded. As a result of these changes, the 44 data points identified as outliers for these three variables were removed from the data so as not to statistically bias the results.

Statistical Analyses. All statistical analyses were run in SPSS. As a brief reminder, due the fact that the learning rate and inverse decision parameters were multiplicative, all statistical analyses were conducted on the logarithms of each parameter estimate. For the reasons outlined in chapter 2, impulsivity and sex were included as nuisance variables in the current study. In further support of the inclusion of sex as a nuisance variable, past research has argued that males demonstrate a superior ability to adapt their choice behaviour to sudden changes in learned action-outcome contingencies compared to females (Evans & Hampson, 2015; Overman, 2004; although, see Chowdhury et al., 2019 for an alternative account). It has been suggested that this may be due to sex differences in the rate at which regions of the brain implicated in cognitive control (Vijayakumar et al., 2014) and reward learning (Chahal et al., 2021) develop, particularly during adolescence. Therefore, sex was included as a nuisance variable in the statistical analyses.

Twelve stepwise multiple linear regressions were performed to investigate whether any of the 12 DVs could be predicted by age, impulsivity or sex. In each regression analysis, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block. The purpose of the statistical tests performed on the learning rate variables (α, σ_{α}, $\Delta\alpha$, $\Delta\sigma_{\alpha}$) was to reveal the extent to which age influenced participants' ability to adapt their actions in response to past trial outcomes in each condition when the respective influences of impulsivity and sex are taken into account. Whereas, the purpose of the tests conducted on the inverse decision temperature variables (β, σ_{β}, $\Delta\beta$, $\Delta\sigma_{\beta}$) was to determine the impact of age on the extent to which participants' choice behaviour was guided by knowledge of the relative outcome probabilities in each condition when impulsivity and sex are controlled for. G*Power analysis revealed that 77 participants were required to obtain a medium sized-effect (f^2=.15) in a multiple linear regression with three predictor variables, 80% power, and a 5% alpha level (Faul et al., 2009). As the sample contained 320 participants, the analyses were sufficiently powered.

Results

In the current study, participants completed 180 trials of an outcome learning task. To briefly reiterate, participants were asked to select one of two prize boxes on each trial. The chance that selecting each box would result in points earned varied per condition. In the stable condition, there was a fixed 75% chance that selecting one of the boxes (i.e., the red box) would lead to a reward. The alternative box then had a fixed 25% chance of giving a reward if selected. Whereas, in the volatile condition, the two boxes had an 80% and 20% chance of leading to a reward if chosen. Every 20 trials, the 80% and 20% reward probabilities swapped between the two boxes. The choices made by participants on each trial were then recorded.

Task Performance

On average, participants selected the correct box on 58% of trials across the task (*SD*= 5.96, *range*= 42.33-71.67). This performance level was significantly different from chance, $t(319)$= 26.43, p<.001, d=5.96. When each condition was examined in isolation, participants selected the correct box on 60% of trials (*SD*= 7.53, *range*= 35.29-76.4) on average in the stable condition, and on 57% of occasions (*SD*= 7.61, *range*= 27.78-74.7) on average in the volatile condition. Performance level was significantly different from chance in both the stable ($t(319)$ = 24.43, p<.001, d=7.53) and the volatile ($t(319)$ = 17.19, p<.001, d= 7.61) conditions. This suggests that participants understood the goal of the task.

The Influence of Age on the Learning Rate in the Stable and Volatile Conditions

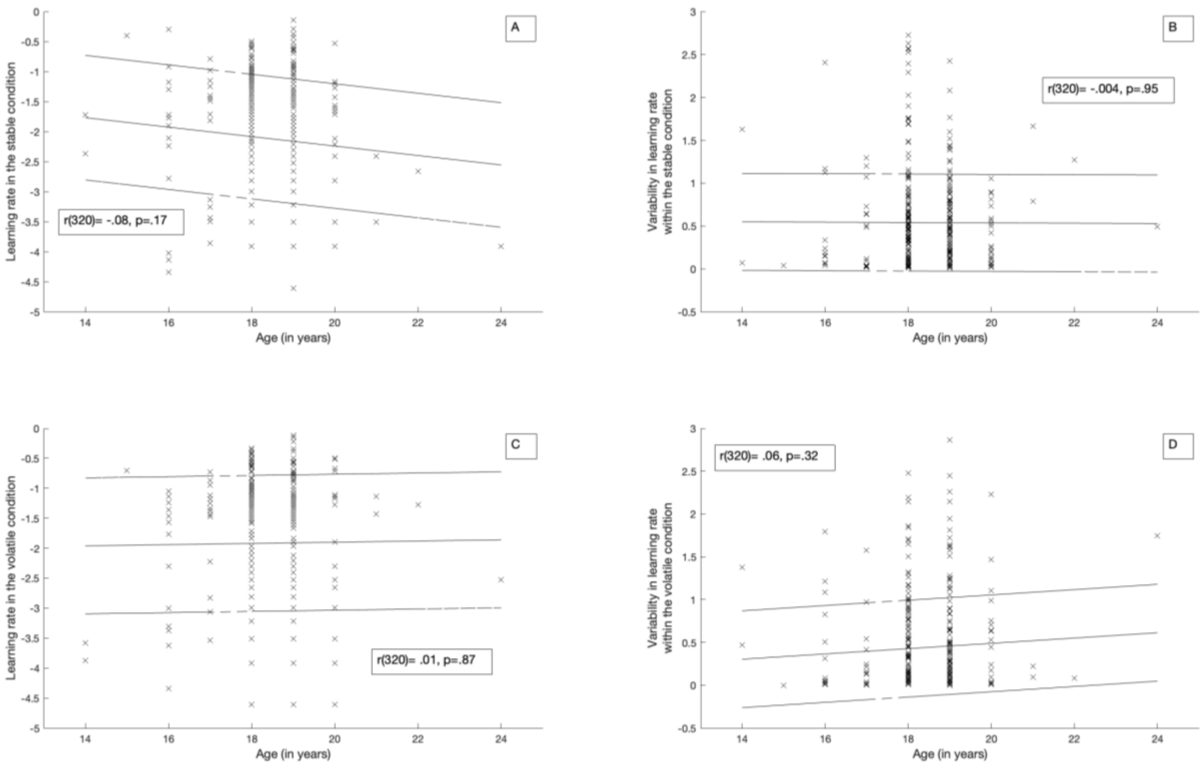
To explore the extent to which participants' choice behaviour was informed by the outcomes of recent actions relative to the outcomes of actions that occurred further in the past, an estimate of their α and σ_α were obtained for each condition. The higher a participant's α, the greater the extent to which their choices were guided by the outcomes of their most recent actions compared to those of actions committed on more antecedent trials. Whereas, the higher a participant's σ_α, the greater the variability in the extent to which participants' choices were directed by more recent trial outcomes compared to more prior outcomes.

Initially, four one-sample t-tests were conducted to ensure that a sufficient level of learning, and variability in learning, occurred on each condition to warrant further analysis. It was revealed that the stable condition α (M= -2.11, SD = 1.03; t(319)= -36.46, p<.001, d= 1.04), the stable condition σ_α (M= .19, SD = .17; t(319)= 17.09, p<.001, d= .57), the volatile condition α (M= -1.92, SD = 1.13; t(319)= -30.22, p<.001, d= 1.13), and the volatile condition σ_α (M= .24, SD = .19; t(319)= 13.94, p<.001, d= .57) all significantly differed from zero. This suggests that sufficient learning occurred to facilitate further analysis regarding the impact of age on participants' rate of learning within both conditions.

To investigate the influence of age on participants' α and σ_α in each condition, four separate stepwise multiple linear regressions were conducted on the stable condition α and σ_α and the volatile condition α and σ_α. In each regression, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block. For both conditions, it was revealed that the α was not significantly predicted by any of the predictor variables (all p>.05). Similarly, it was also revealed that the σ_α was not significantly predicted by any of the predictor variables in either condition (all p>.05). For visualisation purposes, see figure 6.2 for the α and σ_α for each condition plotted against participants' unlogged age in years.

Figure 6.2.

The α and σ_α *as a function of participants' unlogged age in years and condition*



Note. 6.2A. A figure showing learning rate in the stable condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.2B. A figure showing the variability in learning rate within the stable condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.2C. A figure showing learning rate in the volatile condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.2D. A figure showing the variability in learning rate within the volatile condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

The Influence of Age on the Inverse Decision Temperature in the Stable and Volatile Conditions

To explore the extent to which participants' choice behaviour was guided by an understanding of the relative advantage of choosing one box over the alternative option, an estimate of their inverse decision temperature (β) and an estimate of the variability in their inverse decision temperature (σ_β) were obtained for each condition. The higher a participant's β, the greater the extent to which their choices were informed by an understanding of the relative outcome probabilities associated with each colour option. Whereas, the higher a participant's σ_β, the greater the variability in the extent to which participants' choices were guided by this understanding of the underlying outcome probabilities.

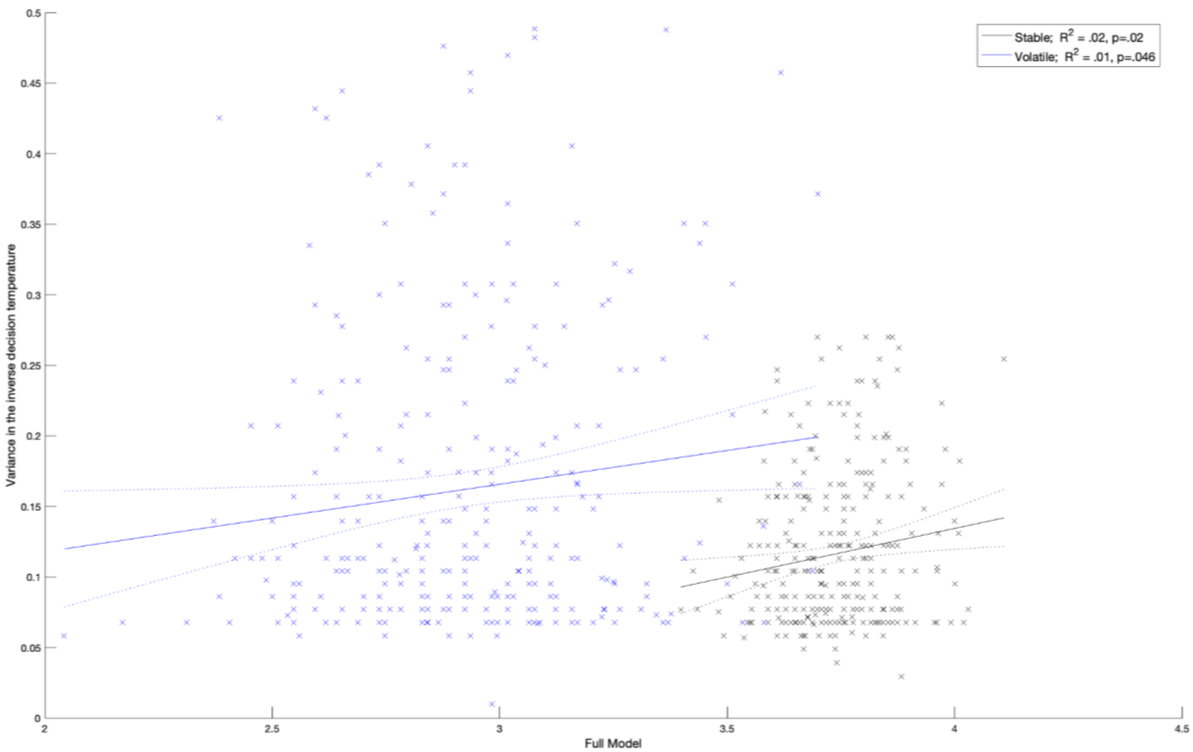
To ensure that the recorded β and σ_β were of a sufficient magnitude to facilitate the planned regression analyses, four one-sample t-tests were conducted on the β and σ_β for each condition. Indeed, it was revealed that the stable condition β (*M*= 1.55, *SD*= .97; *t*(319)= 28.51 *p*<.001, *d*= .97), the stable condition σ_β (*M*= .12, *SD*= .05; *t*(308)= 38.23, *p*<.001, *d*= .05), the volatile condition β (*M*= 1.54, *SD*= .89; *t*(319)= 30.75, *p*<.001, *d*= .89, and the volatile condition σ_β (*M*= .16, *SD*= .11; *t*(302)= 26.39, *p*<.001, *d*= .11) all significantly differed from zero. This suggests that these parameters are sufficient in magnitude to warrant further analysis.

To establish the influence of age on participants' β and σ_β in each condition, four separate stepwise multiple linear regressions were conducted on the stable condition β and σ_β and the volatile condition β and σ_β. In each regression, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block. For both conditions, it was revealed that β was not significantly predicted by any of the predictor variables (all *p*>.05).

In terms of the σ_β, a significant model was revealed for the stable condition, *F*(1,307)= 5.12, *p*= .02. It was revealed that the σ_β was significantly predicted by impulsivity (β = .13, *t*= 2.26, *p*= .02) and was not predicted by age (β = .07, *t*= 1.18, *p*= .24) or sex (β = .03, *t*= 53, *p*= .6). This suggests that as impulsivity increased, the variance in the inverse decision temperature increased within the stable condition. The overall model fit was R^2 = .02 (*SE*= .05). Likewise, a significant model was also revealed for the volatile condition, *F*(1,301)= 4.02, *p*= .046. It was found that the σ_β was significantly predicted by impulsivity (β = .12, *t*= 2, *p*= .046) and was not predicted by age (β = .003, *t*= .05, *p*= .96) or sex (β = -.04, *t*= -.69, *p*= .49). This suggests that, much like the stable condition, as impulsivity increased, the variance in the inverse decision temperature also increased within the volatile condition. The overall model fit was $R^{2\,=}$ .01 (*SE*= .11). Figure 6.3 shows the σ_β plotted against a model of age, impulsivity and sex for each condition. To further visualise the relationships between age and the β and σ_β in each condition, see the scatterplots presented in figure 6.4.

Figure 6.3.

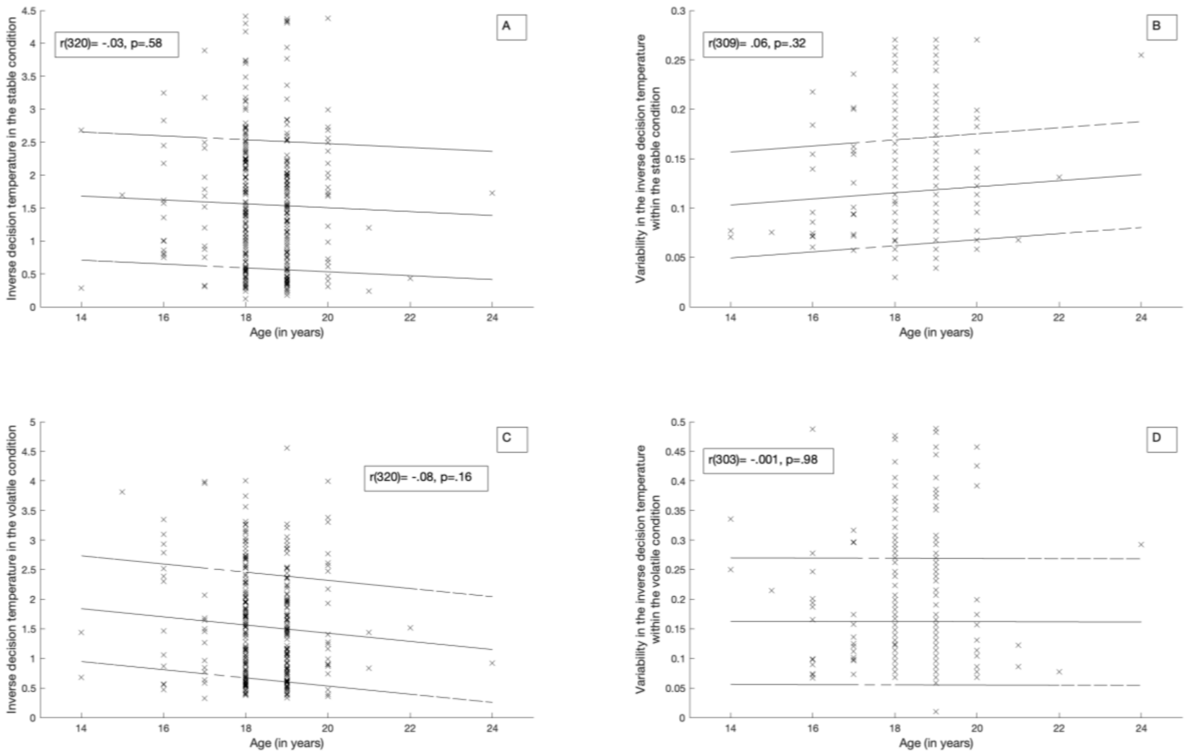
The variance in the inverse decision temperature plotted against a model of age, impulsivity and sex for the stable and volatile conditions.



Note. A figure showing the variance in the inverse decision temperature for both the stable condition and volatile condition plotted against a model of age, impulsivity and sex. The stable condition is represented in black. Whereas, the volatile condition is depicted in blue. Error bars on both models represent +/- 1 confidence interval.

Figure 6.4.

The β and σ_β *as a function of participants' unlogged age* *in years and condition*



Note. 6.4A. A figure showing inverse decision temperature in the stable condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.4B. A figure showing the variability in the inverse decision temperature within the stable condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.4C. A figure showing inverse decision temperature in the volatile condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.4D. A figure showing the variability in the inverse decision temperature within the volatile condition plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

The Influence of Age on the Difference in Learning Rate and Inverse Decision Temperature Between the Two Conditions

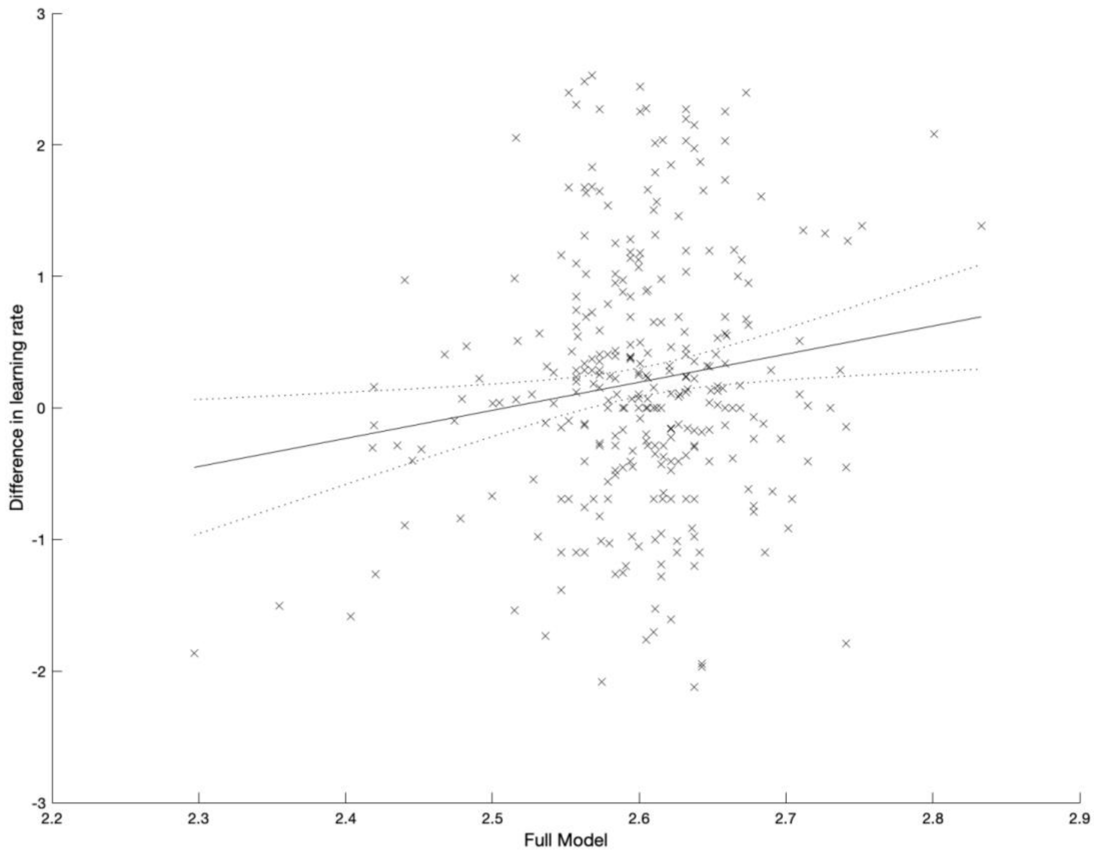
To determine the extent to which participants' α, σ_α, β, or σ_β differed between the stable and volatile conditions, the difference in learning rate ($\Delta\alpha$), the difference in the variability in learning rate ($\Delta\sigma_\alpha$), the difference in inverse decision temperature ($\Delta\beta$), and the difference in the variability in their inverse decision temperature ($\Delta\sigma_\beta$) were calculated. In the case of the $\Delta\alpha$ and the $\Delta\beta$, higher values indicated that a participant showed a higher α or β in the volatile condition than in the stable condition. Whereas, a higher $\Delta\sigma_\alpha$ or $\Delta\sigma_\beta$ demonstrate that the participant showed greater variability in their α or β in the volatile condition relative to the stable condition.

To verify that the difference in α, σ_α, β, and σ_β was of a sufficient magnitude to warrant the planned regression analyses, four one-sample t-tests were conducted on the $\Delta\alpha$ (M= .2, SD= .99), $\Delta\sigma_\alpha$ (M= -.01, SD= .67), $\Delta\beta$ (M= -.01, SD= .8), and $\Delta\sigma_\beta$ (M= .04, SD= .11) values. It was revealed that the $\Delta\alpha$ ($t(303)$= 3.57, p<.001, d= .97), the $\Delta\sigma_\alpha$($t(319)$= -2.67, p= .01, d= .67), and the volatile condition $\Delta\sigma_\beta$ ($t(319)$= 6.44, p<.001, d= .11) all significantly differed from zero. Whereas, the $\Delta\beta$, ($t(319)$= -.33, p= .74, d= .8) did not significantly differ from zero. For this reason, the planned regression for the $\Delta\beta$ will not be performed. In addition, as the $\Delta\alpha$ is positive, it can be interpreted that participants tended to have a higher learning rate in the volatile condition relative to the stable condition. This provides further indication that the task was completed as intended.

To investigate the influence of age on the $\Delta\alpha$, the $\Delta\sigma_\alpha$, and the $\Delta\sigma_\beta$, three separate stepwise multiple linear regressions were conducted. In each regression, impulsivity and sex were entered in an initial block as nuisance variables, and age was entered alone in a second block. It was revealed that none of the predictor variables explained a significant proportion of the variation in the $\Delta\sigma_\alpha$ or the $\Delta\sigma_\beta$ (all p>.05). Whereas, a significant model was found for $\Delta\alpha$, $F(1,302)$= 5.23, p= .02. It was revealed that $\Delta\alpha$ was significantly predicted by age (β= .13 , t= 2.29, p= .02), and was not significantly predicted by impulsivity (β= -.06, t= -.1.08, p= .28). or sex (β= -.01, t= -.23, p= .82). This suggests that as age increased, the difference in learning rate also increased. The overall model fit was R^2 = .02 (SE= .96). Figure 6.5 shows the $\Delta\alpha$ plotted against a model of age, impulsivity and sex. For visualisation purposes, see figure 6.6 for the $\Delta\alpha$, $\Delta\sigma_\alpha$, $\Delta\sigma_\beta$ plotted against participants' unlogged age in years.

Figure 6.5.

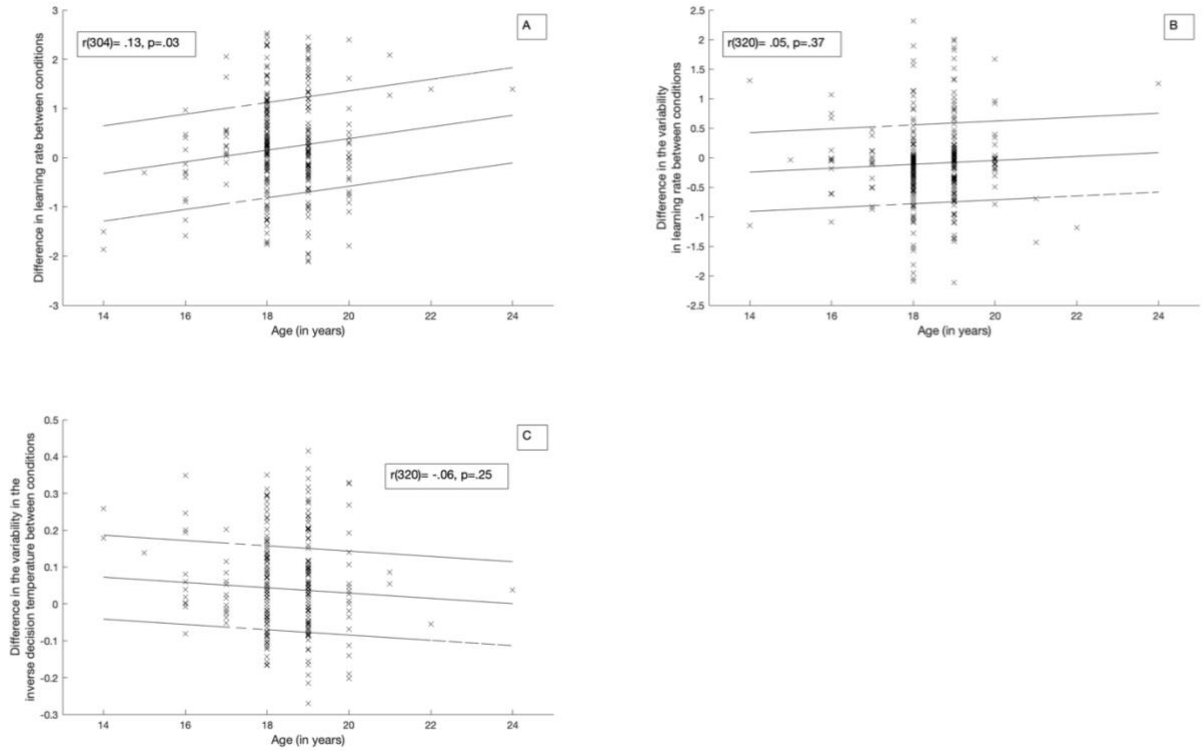
The difference in learning rate between conditions plotted against a model of age, impulsivity and sex.



Note. A figure showing the difference in learning rate between conditions plotted against a model of age, impulsivity and sex. Error bars represent +/- 1 confidence interval.

Figure 6.6.

The $\Delta\alpha$, $\Delta\sigma_{\alpha}$, $\Delta\sigma_{\beta}$ as a function of participants' unlogged age in years



Note. 6.6A. A figure showing difference in learning rate between the two conditions plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.6B. A figure showing the difference in the variability in learning rate between the two conditions plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation. 6.6C. A figure showing the difference in the variability in the inverse decision temperature between the conditions plotted against participants' unlogged age in years. Error bars represent +/- 1 standard deviation.

Discussion

The purpose of the present study was to determine the influence of age on young people's ability to update their forward model to an appropriate degree with respect to the volatility of the current context. To achieve this goal, participants aged 14-24 completed 180 trials of an outcome learning task. To briefly recap, participants selected between two prize boxes on each trial with the goal of obtaining a points-based reward. In a stable condition, the probabilistic action-outcome relationships remained fixed; one box always had a higher chance of delivering a reward. Whereas, in a volatile condition, these action-outcome associations shifted between the two boxes every 20 trials. The extent to which a participant's choices were guided more so by recent trial outcomes as opposed to more antecedent outcomes was indexed separately for each condition via a learning rate. The better the participant's capacity to modify their behaviour according to the relative volatility of the active action-outcome relationships, the greater the distance between their two learning rates should be, with a higher learning rate employed in the volatile context relative to the stable context (Browning et al., 2015).

The present study found that the degree to which individuals updated their forward model in response to an observed action-outcome did not vary with age in either the stable or the volatile context. This finding is inconsistent with both the first and second hypotheses of the current study and the results of past literature (e.g., van den Bos et al., 2012; Eckstein et al., 2022; Jepma et al., 2020), as it contradicts the idea that the tendency to employ a low learning rate within a stable context and a high learning rate in a volatile context increases with age. Likewise, this also contradicts the alternative notion proposed by Hauser et al. (2015) that adolescents are better equipped to respond to unexpected action-outcomes than adults, resulting in a decline in the volatile learning rate with age. Arguably, the current results could imply that the capacity to update one's forward model to the correct degree in response to an unexpected outcome is already present from age 14. However, this notion is largely speculative given that the results were null and only a linear change model was tested (see chapter 7 for a discussion on the limitations of fitting the data to only one type of model).

The precise reason for the disparity between the current findings and those of past research is unclear. This is particularly true, given that the two-forced choice probabilistic learning task employed in the current task was near identical to the tasks used in the majority past studies, with similar stimuli, trial structure and task length (e.g., van den Bos et al., 2012; Hauser et al., 2015; Eckstein et al., 2020). One potential explanation for the lack of significant effects in the current study is that only 10% of the current sample were aged <18. In contrast, past studies have tended to recruit an approximately equivalent number of adolescent and adult participants (e.g., van den Bos et al., 2012; Hauser et al., 2015; Jepma et al., 2020). Therefore, the role of age as a determinant in young people's outcome learning skills may have been obscured by the homogeneity of the current sample. In further support of this suggestion, the only other study to report no difference in learning rate between adolescents and adults on a probabilistic learning task, also attributed their results to a lack of age diversity amongst their participants (Javadi et al., 2014). Ergo, future research with greater variation in participants' ages is needed to verify the legitimacy of the current results.

Whilst the current findings appear to suggest that adolescents and adults do not differ in their ability to modify their learning rate to align with both stable and volatile contexts, the relative magnitude to which they adapted their learning rate differently for each context was found to increase with age. This finding aligns with the third hypothesis of the current study, as it implies that the magnitude to which individuals can flexibly optimise their learning rate to each context refines with age. Therefore, this suggests that, from adolescence to adulthood, subtle refinements occur in individuals' capacity to flexibly modify their forward model according to the current environmental volatility. However, it should be noted that only a low proportion of the variance in the difference in learning rate was explained by the final model (R^2 = .02). Hence, the current results should be interpreted with caution.

In addition to the learning rate, participants' inverse decision temperature was also calculated for each condition in order to quantify their ability to maintain and utilise an accurate conceptualisation of the current action-outcome associations. It was found that this did not vary with age for either condition, nor between conditions. This finding is at least partially consistent with past literature. For instance, Decker et al. (2016) argued that children tended to rely solely on a simplistic stimulus-response strategy to guide their choices on a probabilistic learning task. Whereas, from early adolescence, individuals tended to use knowledge of the underlying action-outcome structure to inform their decisions. This tendency was reported to then improve with age from adolescence to adulthood. In agreement with the results of Decker et al. (2016), the current findings appear to show that adolescents from age 14 onwards do possess the capacity to understand the active action-outcome associations and use this to drive their choices. However, given that these are null results, they should be interpreted with caution. Furthermore, in contrast to the study reported by Decker et al. (2016), no age-related refinement was found. Speculatively, it may be argued that this discrepancy may have occurred due to the noted lack of diversity in participants' ages, which may have obscured these subtle age-related changes.

Moving forward, future research should modify the combined Rescorla-Wagner and softmax action selector model to account for age-related differences in the rate at which the perceived reward probabilities for chosen and unchosen options are updated (Fischer & Ullsperger, 2013). The learning model used in the current study functioned under the assumption that all participants use a counterfactual learning strategy. This means that their understanding of the reward probabilities for both the chosen and unchosen boxes were updated simultaneously after each trial (Boorman et al., 2011). However, past research has shown that, whilst adults do tend to use this counterfactual learning strategy, adolescents' choice behaviour is better explained by a more basic model where only the reward probability of the chosen option is updated (Palminteri et al., 2016). This would suggest that adults are better able to update all relevant action-outcome associations in a given context than adolescents, and thereby, will possess a more accurate forward model. Therefore, future research should test multiple potential learning models, in order to best capture the true influence of age on individuals' capacity to adapt their understanding of the causal structure of their environment in light of new information.

Similarly, the combined Rescorla-Wagner and softmax action selector model should also be modified to account for the idea that age has a differential effect on the rate at which individuals learn from positive outcomes and negative outcomes (Ferdinand et al., 2016). Notably, the evidence regarding the precise manner in which age and outcome valence interact remains mixed. For instance, some learning studies have reported a shift towards greater emphasis on positive over negative feedback to inform choice behaviour from adolescence to adulthood (e.g., van der Schaaf et al., 2011; Hartley & Somerville, 2015), whilst others have argued the reverse to be true (e.g., van Duijvenvoorde et al., 2008). This suggests that the precise manner in which age interacts with outcome valence to influence the rate at which individuals can learn from past outcomes is unclear. Unfortunately, learning model used in the current research did not differentiate between rewarded and unrewarded outcomes. Ergo, future research is needed to elucidate the impact of outcome valence on the ability to learn from past action experience at different ages.

To conclude, the purpose of the present study was to determine how young people's ability to update their forward model to an appropriate degree with respect to the volatility of the current context changes with age. Although the ability to appropriately modify one's learning rate to align with both stable and volatile contexts was found to be age-invariant, the relative magnitude to which one can adapt their learning rate differently for each context was found to improve with age. This suggests that the capacity to update the forward model appropriately is present from early adolescence, and refines with age. Notably, the lack of heterogeneity in participants' ages and the lack of differentiation between learning rates for chosen and unchosen, as well as reward and unrewarded, outcomes should be acknowledged as potential confounds. Consequently, future research should initially be concerned with evaluating whether the current findings can be replicated once the current design has been amended to remove the influence of these factors.

Chapter 7: General Discussion

Chapter Summary

Chapter 7 provides a general discussion of the current research, beginning with a brief reiteration of previous knowledge and the aims of the thesis. The major findings from the current studies will then be outlined, in addition to their implications for the forward model development literature. Next, the significance of the current findings to previous knowledge on the developmental trajectory of SoA will be discussed. Following this, the general strengths and weaknesses of the thesis will be explored. Avenues for future research will then be examined, before ending with the final conclusion of the thesis.

Aims of the Thesis and Brief Reiteration of Previous Knowledge

The purpose of this thesis was to amend the noted absence of adolescents from prior SoA development literature, and thus, determine the full trajectory at which the capacity to experience a veridical SoA develops from childhood to adulthood. To recap, a SoA refers to an individual's awareness of their control over their voluntary actions and the sensory consequences of those actions (Haggard & Chambon, 2012). It has been argued that a consensus regarding the manner in which SoA matures from childhood to adulthood remains absent from prior literature (Choudhury et al., 2007). Whilst some prior research has concluded that children demonstrate a reduced SoA compared to adults (e.g., Cavazzana et al., 2014, 2017), these studies have tended to ignore adolescents' experience of agency. Indeed, only two prior studies have included adolescents within their investigation of how SoA reaches an adult-like level of precision (Aytemür et al., 2021; Aytemür & Levita, 2021). When taken together, these two studies reported that adolescents have a less precise SoA compared to both children (Aytemür & Levita, 2021) and adults (Aytemür et al., 2021; Aytemür & Levita, 2021). However, the fact that both studies employed a version of the intentional binding effect to measure SoA (Aytemür et al., 2021; Aytemür & Levita, 2021) raises concern over the legitimacy of their findings. This may be said as the legitimacy of this effect as a pure SoA measure has been criticised within past literature (Suzuki et al., 2019). Hence, the precise manner through which SoA matures from childhood to adulthood warranted further investigation.

A SoA is believed to be produced via an internal computation through which the predicted and observed consequences of self-authored actions are compared within a forward model (Haggard & Chambon, 2012). On that basis, the proficiency of the forward model system was examined as a proxy measure of agency in the current research. More specifically, the functionality of an individual's forward model system was interrogated via their ability to i) accurately predict the outcome of their action, and ii) update learned action-outcome knowledge in response to post-action feedback; two skills indicative of a precise forward model. Similar to previous SoA literature (e.g., Cavazzana et al., 2014, 2017), the majority of past studies have focused on understanding how children and adults differ in their ability to predict action consequences (e.g., Franchak, 2019; Perchet & Garcia-Larrea, 2005) and adapt behaviour in response to sensory feedback (e.g., Tahej et al., 2012; Scheerer et al., 2016). Conversely, few studies have directly examined the manner in which the forward model system develops across adolescence (Quatman-Yates et al., 2012; Barlaam et al. 2012; Dahl et al. 2018). Hence, the primary aims of the current thesis were twofold:

1. Based on past literature (e.g., Van Gerven et al., 2016), the first goal of this thesis was to test the idea that the ability to form accurate action-outcome predictions improves with age from childhood to adulthood.

2. In addition, the second goal of this thesis was to evaluate the suggestion that the ability to appropriately update learned action-outcome associations in light of post-action feedback improves with age from childhood to adulthood, as suggested by past research (e.g., Master et al., 2020).

By accomplishing these two initial goals, it was reasoned that the ultimate thesis aim could be achieved:

3. To assess the assertion that SoA matures at a linear rate from childhood to adulthood, as suggested by past child studies (e.g., Cavazzana et al., 2014; 2017) and neuroimaging research (e.g., Blakemore et al., 2012), in light of the contradictory evidence presented by Aytemür and Levita (2021).

Summary of Major Findings and Their Implications for the Forward Model Development Literature

Attention will now turn to discussing how the primary aims of the thesis were investigated within each empirical chapter, the key findings that were obtained, and how these findings extend our knowledge on how the functionality of the forward model system changes with age from childhood to adulthood.

Summary of Chapter 3

Beginning with chapter 3, this chapter aimed to determine how sensorimotor continuation changes with age across childhood, adolescence, and young adulthood. Sensorimotor continuation refers to an individual's ability to maintain a specified inter-response-interval when producing a series of isochronous motor responses (McPherson et al., 2018). Sensorimotor continuation was measured using a synchronisation-continuation task. Participants aged 4-25 years first made keypresses in time with a series of isochronous tones played at either a high, medium or low frequency. They then continued making keypresses at the same pace after the tones were removed. To accurately and consistently replicate the set response pace, participants had to use their forward model to predict the precise time at which to make their next response and correct any disparities between their produced inter-response-interval and the target inter-response-interval (Maes, 2016). Therefore, the greater the accuracy and consistency with which participants could maintain the target response pace, the better their ability to both form veridical forward model predictions and make appropriate updates to their forward model in light of tactile feedback from their keypress. Hence, the results of this study answer both of the two primary aims of this thesis.

It has been argued that evidence on how sensorimotor continuation changes with age from childhood to adulthood is limited within past literature. Only two prior studies have investigated how this ability reaches adult-like maturity, both of which failed to reach a consensus. One study argued that sensorimotor continuation improves with age from childhood to adulthood (McAuley et al., 2006), whilst the other concluded that adolescents and adults do not differ in their sensorimotor continuation skills (Witt & Stevens, 2013). Hence, the purpose of chapter 3 was to resolve this discrepancy, and thereby, advance our current understanding of how sensorimotor continuation develops with age from childhood to adulthood. In agreement with the conclusion drawn by McAuley et al. (2006), the current research found that the accuracy and consistency of participants' sensorimotor continuation improved with age from childhood to young adulthood. Furthermore, this suggests that the accuracy and consistency of participants' forward model predictions and their ability to appropriately amend their action-outcome knowledge both refine with age. However, it should be noted that the findings may have been confounded by interactions between age and participants' working memory capacity (Gomes et al., 1999), time perception (Droit-Volet et al., 2007), and experience of music training (Thompson et al., 2015). Hence, future research is needed to verify whether the current findings can be replicated after the impact of these potential confounds upon the results has been controlled.

Summary of Chapter 4

Building on the results of chapter 3, the aim of chapter 4 was to examine, in isolation, how the tendency to form forward model predictions changes with age from childhood to adulthood. This was achieved by interrogating the influence of age on predictive motor timing in children, adolescents and young adults. Predictive motor timing refers to the ability to manipulate the timing of an intended action such that its occurrence aligns with the predicted onset of an imminent stimulus (Tanaka et al., 2021). This ability was measured using a cued reaction time task. Participants aged 4-25-years were first presented with an amber cue stimulus, followed by a target green stimulus after a variable interval. Participants' objective was to respond as soon as the target stimulus became visible.

Crucially, participants could achieve the task objective by either making an anticipatory response or a reactive response (Braver, 2012). Anticipatory responses required participants to use their forward model to predict when to respond in order for their keypress to temporally align with the onset time of the target stimulus. Whereas, reactive responses were triggered by the onset of the target stimulus, and thus, required no internal action preparation via the forward model in advance of the target stimulus' arrival (Burnett Heyes et al., 2012). Notably, an anticipatory response would achieve a faster reaction time relative to a reactive response, meaning that anticipatory responses were more advantageous for the task objective. Hence, the higher the ratio of anticipatory to reactive responses, and the faster and more consistent the rate of rise in participants' anticipatory decision process, the greater their tendency to form forward model predictions.

Previous literature has tended to focus on comparing the performance of children and young adults on cued RT tasks (Iselin & DeCoster, 2009), often reporting that the latter demonstrate greater anticipatory response behaviour than the former (Brown, 2019; Perchet & Garcia-Larrea, 2005). Whilst these findings appear to suggest that predictive motor timing skills improve from childhood to adulthood, it has been argued that further empirical investigation is needed to verify this assertion (Debrabant et al., 2012). To that end, the current research found that the tendency to make anticipatory responses over reactive responses increased with age from childhood to adulthood. Greater age was also associated with a faster and more consistent rate of rise in the anticipatory decision process towards the threshold required for action execution. These results extend the findings of previous research (Brown, 2019; Perchet & Garcia-Larrea, 2005), as they suggest that predictive motor timing refines with age from childhood to adulthood. These findings are consistent with the results of both behavioural (Van Gerven et al., 2016) and neuroimaging studies (Killikelly & Szűcs, 2013), which have also reported that the ability to prepare responses in advance of anticipated stimuli improves with age from childhood to adulthood. Subsequently, in congruence with the results of chapter 3, the current findings suggest that the ability to conceptualise the outcome of a planned action using a forward model refines with age.

Although, it should be noted that, less than a third of the responses made were anticipatory across participants, which is lower than documented in previous research (e.g., Brown, 2019). This jeopardises the legitimacy of the results, as it suggests that an aspect of the current task artificially deterred some participants from preparing anticipatory responses. The precise factor that caused this low average rate of anticipatory response behaviour to occur is difficult to identify retrospectively. Hence, future research is required to determine why this transpired and verify whether the results of the current study can be replicated in scenarios where a higher average rate of anticipatory responding is observed across participants.

Summary of Chapter 5

To further extend the findings of chapter 4, the goal of chapter 5 was to establish how the ability to use appropriate prior knowledge to guide forward model predictions changes with age from childhood to adulthood. There are numerous instances in daily life where individuals must shift flexibly between different tasks, each with their own relevant action-outcome associations. To facilitate this, relevant prior knowledge must be combined with available context cues so that the most appropriate action can be selected for the current task (Berniker & Körding, 2011). Switch costs are often recorded as an indicator of an individual's ability to switch between action-outcome pairings as the contextual information changes between trials. Whereas, mixing costs provide an index of an individual's ability to maintain, and select between, different action-outcome associations (Kray et al., 2008). In both instances, lower costs indicate a lower frequency of errors made when switching or selecting between different action-outcome associations, and hence, greater accuracy in one's forward model predictions.

Thus far, past research has failed to achieve a consensus on how the ability to maintain, and switch between, different action-outcome pairings changes from childhood to adulthood. Previous studies have predominantly found that children demonstrate higher switch costs and mixing costs than young adults on task-switching paradigms (e.g., Davidson et al., 2006; Kray et al., 2004; Kray et al., 2008). However, only two previous studies have included adolescents within their sample, both of which failed to find conclusive evidence to support the existence of an age-related decline in either switch costs or mixing costs (Manzi et al., 2011; Reimers & Maylor, 2005). To rectify this issue, the current research measured the switch costs and mixing costs accrued by individuals aged 5-21 on a goal-switching task. During the task, participants switched between two objectives: when green stimuli appeared, participants made pro-saccade responses; whereas, red stimuli warranted anti-saccades.

In agreement with some previous studies (e.g., Davidson et al., 2006), it was found that older age was associated with lower switch costs. Therefore, this suggests that the ability to flexibly switch between relevant action-outcome associations develops with age from childhood to adulthood. This result is also consistent with the idea that the ability to inhibit previously relevant action-outcome associations improves with age from childhood through to adulthood

(Diamond, 2013), in line with the maturation of the frontoparietal network (Marek & Dosenbach, 2022). Unfortunately, the impact of age on the ability to maintain, and select between, different action-outcome mappings could not be examined due to the fact that low mixing costs were obtained across participants. Additionally, confidence in the reliability of the current findings was undermined by concerns regarding the spatial acuity of the eye-tracking software used to measure saccades and the lack of age-diversity in the recruited sample. Therefore, future research is needed to verify whether the similar findings can be obtained once these issues have been resolved.

Summary of Chapter 6

Complementary to chapters 4 and 5, the aim of chapter 6 was to investigate, in isolation, how the ability to make appropriate updates to learned action-outcome knowledge in light of post-action feedback changes with age from adolescence to adulthood. In order to maintain an up-to-date understanding of the probabilistic associations between actions and their effects, the feedback observed after an action must be incorporated into the individual's prior estimate (Berniker & Körding, 2011). Crucially, the rate at which these amendments are made to learned action-outcome associations must be modulated according to the volatility of the current context (Gershman, 2015). To recap, the extent to which the prior is modified in response to an observed outcome can be expressed as a learning rate (Hohwy, 2017). The higher the learning rate, the greater the influence of recent outcomes on the prior, relative to the wider history of observed feedback (Eckstein et al., 2022). In a relatively stable context, where probabilistic action-outcome relationships remain fixed over time, it is optimal to possess a low learning rate (Behrens et al., 2007). Whereas, in a more volatile context, where action-outcome associations are subject to frequent change, a high learning rate is favourable (Browning et al., 2015). By examining an individual's ability to optimally modify their learning rate according to the volatility of current context, it is possible to measure their capacity to make appropriate updates to their forward model.

Few past studies have investigated how the ability to adapt one's learning rate to the volatility of the current context develops from adolescence to adulthood (DePasque & Galván 2017). Furthermore, whilst some studies have reported that adults tend to successfully employ a lower learning rate than adolescents within stable contexts (van den Bos et al., 2012; Jepma et al., 2020), the evidence in regard to the learning rates employed within volatile contexts appear to be more mixed. Some studies have argued that adolescents use a higher learning rate than adults within volatile contexts (Hauser et al., 2015). This suggests that observed outcomes trigger a larger update in adolescents' prior compared to that of adults, regardless of the current contextual volatility. Opposingly, other studies have argued the converse to be true; adults were shown to utilise a higher learning rate than adolescents (Eckstein et al., 2020). This suggests that the ability to correctly deflate one's learning rate in a stable context, and inflate one's learning rate in a volatile context improves with age across adolescence.

In order to resolve the noted discrepancies between the findings of past studies, current research measured adolescents' and young adults' ability to update action-outcome knowledge via an outcome learning task. In both a stable context and a volatile context, participants aged 14-24-years selected between two boxes with the goal of finding a reward. In conflict with the results of past literature (e.g., van den Bos et al., 2012; Eckstein et al., 2020; Hauser et al., 2015; Jepma et al., 2020), it was found that learning rate did not vary with age in either the stable or the volatile context. Whereas, the degree to which individuals adjusted their learning rate differently for the stable context relative to the volatile context increased with age. On the one hand, when taken together with the findings of chapter 3, the results appear to suggest that the capacity to flexibly modify action-outcome knowledge largely develops from childhood to adolescence, with minor refinements in this ability occurring from adolescence to adulthood. On the other hand, given that only 10% of the current sample were aged less than 18-years-old, the null results obtained from this study could indicate that the true impact of age on the variation in participants' learning rate in each context was obscured. Hence, future research is required to verify whether these results can be replicated with a more age-diverse sample.

Conclusions Drawn Across the Empirical Chapters

Drawing across the four empirical chapters, the results from chapters 3, 4 and 5 demonstrate that the ability to generate accurate action-outcome predictions improves with age from childhood to young adulthood. More specifically, the findings show that both the rate at which individuals engage in action-outcome prediction to guide their actions (*chapter 4*) and the quality of those predictions (*chapters 3 and 5*) improves with age across this developmental period. Ergo, the results obtained from all three chapters successfully address the first aim of the current thesis.

Furthermore, these findings are consistent with past studies which have suggested that, as individuals move from childhood to adulthood, they shift from responding reactively to perceived stimuli to preparing responses proactively in anticipation of expected stimuli (Braver, 2012; Van Gerven et al., 2016) as a result of maturation in the brain (Smith et al., 2011; Vijayakumar et al., 2014). Moreover, the current findings support the idea that the efficiency at which individuals can predict the consequences of planned actions via a forward model develops with age from childhood through to adulthood.

Whilst the results from chapters 3, 4 and 5 were effective in addressing the first aim of this thesis, the findings from chapters 3 and 6 were successful in responding to the second thesis goal. This may be said as, collectively, the results demonstrate that the ability to appropriately update learned action-outcome associations in light of post-action feedback improves with age from childhood to young adulthood. More precisely, the findings indicate that the accuracy (*chapter 3*) and magnitude (*chapter 6*) to which individuals can update action-outcome knowledge in response to sensory feedback refines with age from childhood to adulthood. These findings are congruous with previous learning studies which have shown that the proficiency with which individuals can incorporate past outcome evidence into their prior estimate and use this knowledge to guide current action improves from childhood to adulthood (Barash et al., 2019; Chambers et al., 2018; Master et al., 2020). Overall, the current findings reinforce the notion that the ability to alter one's constructed forward model in light of new outcome evidence continues to develop from childhood through to young adulthood.

Imperatively, as evident across the empirical chapters, the majority of past studies which have contributed to our understanding of forward model development have tended to compare the capabilities of children and adults (e.g., Perchet & Garcia-Larrea, 2005; Davidson et al., 2006). Whereas, few prior studies have explicitly sought to establish a consensus regarding the trajectory at which the functional efficiency of the forward model system progresses across childhood, adolescence and young adulthood (Quatman-Yates et al., 2012; Barlaam et al. 2012; Dahl et al. 2018). In addition, prior studies which have included adolescents within their sample have tended to report contradictory conclusions on how the ability to maintain and/or operationalise action-outcome knowledge changes with age (e.g., Hauser et al., 2015 and Eckstein et al., 2020, as discussed in chapter 6). To that end, the findings presented within this thesis provide a crucial contribution to the forward model development literature, as they advance our understanding of the trajectory at which the forward model develops as individuals progress from childhood to adolescence to young adulthood.

Implications of the Current Findings for SoA Development Literature

To briefly recap, a SoA is believed to emerge as the result of an internal computation occurring within a forward model where the level of concordance between expected and observed action-outcomes is evaluated (Haggard & Chambon, 2012). Subsequently, the accuracy with which an individual can experience a SoA over an observed sensory event is believed to be dependent on the precision of their forward model predictions (Asai, 2017). In turn, the veridicality of those predictions is thought to be reliant on the individual's ability to maintain an up-to-date conceptualisation of the action-outcome contingencies relevant to the current context (Berniker & Körding, 2011). Therefore, the quality of one's SoA experience is underpinned by the functional efficiency of their forward model. For this reason, the implications of the present results for current knowledge of how a SoA develops from childhood to adulthood will now be discussed.

Overall, the findings presented throughout this thesis suggest that the precision of the forward model improves with age as individuals move through childhood, adolescence, and young adulthood. This implies that the accuracy with which an individual can experience a SoA over an observed sensory event also improves with age across this period, thus fulfilling the ultimate goal of the current thesis. These findings extend our current knowledge on how SoA matures from childhood to adulthood. As noted previously, whilst there has been past evidence to suggest that children experience a reduced SoA compared to adults (e.g., Cavazzana et al., 2014, 2017), these studies failed to consider how the precision of one's SoA might alter across adolescence. Only two prior studies included adolescents within their investigation of how the quality of one's agency experience matures with age (Aytemür et al., 2021; Aytemür & Levita, 2021). Thus, the studies presented in this thesis are successful in extending our past SoA development knowledge.

When compared directly with prior literature, it can be said that the current findings are consistent with the results of Cavazzana et al. (2014, 2017), as they support the idea that adults demonstrate a more precise SoA over the consequences of their actions compared to children. The current results also partially align with the findings reported by Aytemür et al. (2021). In research conducted by Aytemür et al. (2021), it was concluded that mid-adolescents (13-14) had a less precise SoA compared to adults (25-28), as they demonstrated a larger outcome binding effect, and hence, were more likely to subconsciously attribute the occurrence of a tone delivered 450ms after their keypress to themselves than adults. To recap, the outcome binding effect refers to a phenomenon in which an individual judges a sensory event to have shifted temporally towards their action. This effect is believed to only occur when the individual believes that their action caused the event to transpire (Render & Jansen, 2021). Hence, the larger the outcome binding effect, the greater the individual's SoA over the observed event (Borhani et al., 2017). Therefore, the result reported by Aytemür et al. (2021) is congruent with the overall conclusion of the current research.

The current findings are also partly consistent with the results reported by Aytemür and Levita (2021). Consistent with the current results, Aytemür and Levita (2021) concluded that adults demonstrate a greater SoA over self-produced sensory events than late-adolescents. This suggests that SoA improves from late adolescence to adulthood. However, Aytemür and Levita (2021) also reported that the magnitude of agency experienced declines from childhood to late-adolescence, with children actually experiencing a SoA comparable to that of adults. Taken together, both findings contradict the conclusion of the current thesis, as they suggest that, instead of maturing in a linear fashion from childhood to adulthood, SoA follows a U-shaped developmental trajectory, with a marked decrement in adolescence.

The noted discrepancy between the results acquired by Aytemür and Levita (2021) and those obtained from the current research can be explained by differences in the way in which participants' capacity to experience a SoA was indexed. Aytemür and Levita (2021) proposed that children in their study failed to achieve an initial feeling of control over the tone due to a lack of precision in their ability to predict the outcome of their action. Consequently, they argued that children formed a retrospective JoA regarding the causal association between their keypress and the tone. As a brief reminder, a JoA refers to a higher-order, introspective belief regarding the most likely cause of an observed outcome that is produced through a process of conscious reasoning

(Desantis et al., 2011; Weiss et al., 2014). This JoA was then believed to have inflated the outcome binding effect demonstrated by children relative to late adolescents (Aytemür & Levita, 2021). As prediction skills improve with age, Aytemür and Levita (2021) argued that child participants' reliance on retrospective JoA cues when attributing causality over the tone diminished. This is reported to have caused the decline in the outcome binding effect from childhood to late-adolescence. Adults' heightened outcome binding effect relative to adolescents' levels of outcome binding was then said to be indicative of adults' superior accuracy in predicting the outcome of their action.

In support of the explanation posited by Aytemür and Levita (2021) to account for their results, it has previously been shown that, when an action is performed involuntarily, and thus, no prediction of the most probable action outcome is available, the magnitude of the outcome binding effect can be modulated by participants' retrospective beliefs regarding their control over the tone (Dogge et al., 2012). This suggests that the outcome binding effect can be driven by a JoA made post-action when pre-action prediction cues are either unreliable or unavailable. In addition, it has been argued that children tend to attribute observed events to their own actions when the true cause of the event is ambiguous or probabilistic (Kushnir et al., 2009). Therefore, it is plausible that the finding of an increased outcome binding effect in children relative to late adolescents resulted from a greater tendency to rely on a retrospective JoA within childhood, in absence of the prediction skills required to compute a reliable FoA.

The outcome binding measure used in research by Aytemür and Levita (2021) assessed the extent to which participants attributed the occurrence of the tone to their own keypress. In contrast, as the current research used the functional efficiency of the forward model system as a proxy measure of agency, the degree to which participants felt control over their action-outcomes was not directly interrogated. This suggests that, unlike the results obtained by Aytemür and Levita (2021), the current findings could not be directly affected by participants' retrospective agency beliefs, as they were unrelated to the goal of each task. This explains the disparity between the conclusions of this thesis and those reported in research by Aytemür and Levita (2021). Furthermore, both the results of research by Aytemür and Levita (2021) and the current findings suggest that an individual's ability to predict the consequences of their actions improves from childhood to adulthood. Moreover, the current results make vital contribution to the SoA literature, as they advance our understanding of how the precision at which a SoA can be experienced develops with age across childhood, adolescence, and young adulthood.

General Strengths and Limitations of the Current Research

The current research possessed a number of strengths and limitations, each of which will now be discussed.

General Strengths of the Current Research

To begin with the strengths, it may be argued that the tasks employed within the current research demonstrate an effective means of assessing the capacity to compute a reliable FoA, in absence of any contamination from one's JoA. Prior to the current research, a FoA had predominantly been indexed using the intentional binding effect in past studies (e.g., Haggard & Clark, 2003). To recap, the intentional binding effect refers to a temporal compression which occurs between the perceived timing of an action and its sensory effect when the effect is thought to be self-authored (Haggard, 2017). As noted previously within this chapter, it has been argued that the intentional binding effect can also reflect an individual's JoA over a sensory event in situations where outcome predictions are unavailable (Aytemür & Levita, 2021; Dogge et al., 2012; Synofzik et al., 2008). Therefore, this suggests that the intentional binding effect does not necessarily provide a reliable means through which to record the FoA exclusively. Conversely, as the current tasks examine the functionality of the cognitive model underlying the FoA, it may be argued that they offer a more direct route to assess the quality with which an individual can experience a FoA. Future research can then use any of the current tasks in tandem with a JoA measure to track how the FoA and JoA are utilised differently throughout development to achieve one's action goals.

Aside from establishing an effective method of assessing an individual's capacity to experience a veridical FoA, the current research was also successful in contributing to current knowledge regarding the feasibility of online behavioural experiments. For instance, as discussed in chapter 5 and consistent with prior literature (e.g., Papoutsaki et al., 2018; Semmelmann & Weigelt, 2018; Slim & Hartsuiker, 2022), the current research noted that WebGazer software lacked the spatial and temporal acuity required to reliably record peak velocity and saccade duration online. This implies that the breadth of eye-tracking data that can be collected online is narrower compared to lab-based studies where it is possible to measure these variables with higher precision (Semmelmann & Weigelt, 2018). Similarly, it was also found that younger participants were worse at maintaining their position in view of the webcam and holding their gaze at the centre of the screen at the start of each trial than older participants. This suggests that the online eye-tracking tasks do not provide a suitable means through which to collect saccade data from children. Ergo, the current results provide a valuable insight into the viability of utilising online studies to collect behavioural data.

General Limitations of the Current Research

Attention will now turn to exploring the limitations of the current research. Admittedly, although the tasks described within this thesis were selected with the intention of exclusively measuring either individuals' ability to predict action-outcomes or their capacity to update action-outcome knowledge, these processes do not operate independently of one another (Wolpert & Flanagan, 2001). Forward model predictions are believed to be produced by combining prior knowledge with current contextual information via Bayes' theorem (Faisal et al., 2008). Consequently, an individual's ability to generate accurate predictions is dependent on their capacity to maintain an up-to-date conceptualisation of relevant action-outcome contingencies (Wolpert & Ghahramani, 2000). For instance, one could argue that, in order to form effective predictions regarding the most likely onset of the green light within the cued RT task (*chapter 4*), participants needed to incorporate the amber durations observed on past trials into their prior estimate (Burke & Roodenrys, 2000; although see appendix C for evidence contrary to this idea). This would then maximise the informativeness of the prior estimate, and thus, lead to the construction of an accurate prediction. Therefore, this suggests that the success with which participants could execute an anticipatory response was directly determined by their ability to learn from past action experience. Ergo, it can be argued that the current findings cannot, definitively, be attributed to the development of either the prediction or the updating process in isolation.

In the real world, action decisions are seldom made without reference to the relative valence of the outcomes associated with each option (Zheng et al., 2015). Indeed, the win-stay and lose-shift choice strategy has been well-documented in past literature (e.g., Worthy & Maddox, 2014). Previous research has reported mixed evidence on how the ability to learn from positive and negative outcomes changes with age. For instance, some studies have reported a shift towards a greater influence of positive over negative feedback on subsequent choice behaviour from childhood to adulthood (e.g., van der Schaaf et al., 2011; Hartley & Somerville, 2015), whilst others have argued the reverse to be true (van Duijvenvoorde et al., 2008). This suggests that the degree to which an individual can learn from an observed outcome is modulated by the interaction between their age and the valence of the observed outcome. Unfortunately, the majority of the present studies either failed to investigate the relative influence of past win or loss outcomes on participants' subsequent predictions (*chapters 4-6*) or failed to offer any feedback on participants' prediction accuracy (*chapter 3*). Therefore, future research should evaluate the comparative effects of positive and negative action-outcomes on participants' forward model predictions and how they might vary with age in order to achieve results with higher ecological validity.

The lack of variability in participants' age across all four empirical studies must also be acknowledged as a limitation of the current research. For instance, only 28% of the final sample reported in chapter 3 were aged less than 18-years-old, with similar numbers seen for the other three chapters (chapter 4: 40%; chapter 5: 18%; chapter 6: 10%). This consistent lack of younger participants can, at least partly, be attributed to the government-imposed restrictions on face-to-face teaching introduced in response to the Covid-19 pandemic. As all face-to-face teaching was suspended for many months during 2020 (Brown & Kirk-Wade, 2021), this decreased the number of opportunities available for recruiting child and adolescent participants via schools. Arguably, this lack of variance in age suggests that the current results are unlikely to be representative of the full developmental timecourse from early childhood to young adulthood. Thus, future research is needed in order to identify whether these findings can be replicated using samples with greater age variance.

Throughout this thesis, participants' data was fit exclusively to linear regression models. However, it is important to acknowledge that a linear pattern is only one of the possible trajectories that forward model development might follow. By taking a solely linear approach to fitting the data, the current research ignored the possibility that an alternative growth pattern, such as a quadratic model as suggested by Aytemür and Levita (2021), might have been a more suitable fit for the data. Furthermore, to conclusively assess whether the forward model does develop at a linear rate, future research should compare how well participants' data fit to a linear model relative to alternative types of development models, such as quadratic (follow a U-shaped trajectory), phase-change (progress through a series of categorical stages), and sigmoid (undergo a period of rapid growth prior to reaching a plateau). From this, it will then be possible to draw more concrete conclusions on the true trajectory at which the forward model develops throughout childhood, adolescence and young adulthood.

Avenues for Future Research

In addition to the calls for further research that have made throughout this chapter, there are further avenues through which future studies could build upon the findings presented within this thesis, some examples of which will now be outlined.

Evidently, the findings of this thesis revealed the causal impact of an individual's age on the precision of their forward model. Arguably, this age-related change can, in part, be attributed to the maturational alterations that occur within the brain from childhood to young adulthood (Blakemore et al., 2012; Zito et al., 2017). However, the role of motor experience on the precision of the forward model should also be acknowledged. For example, as discussed in chapter 3, past research has shown that years of music training was positively associated with participants' ability to synchronise their finger-taps with a series of tones (Thompson et al., 2015). Experience of learning to play a musical instrument is believed to offer individuals with the opportunity to practice predicting when beats will occur (Slater et al., 2018) and adjusting the timing of movements according to auditory cues (Krause et al., 2010). This suggests that greater experience in performing a specific action, or activity, can result in a more precise forward model that is specific to that context.

Indeed, the relevance of motor experience to the accuracy of the forward model is also evident outside of music training. Research by Kretch and Adolph (2012) showed that infants who were experienced walkers were more likely to avoid crossing over a 90-degree cliff edge compared to infants who had only recently begun to walk. This reinforces the idea that, the greater the extent to which participants' possess prior knowledge relevant to the situation (e.g.., walking), the more reliable their prior estimate, and hence, the more precise their outcome predictions (Hohwy, 2017). Intuitively, older individuals will have had a greater timescale in which to practice performing specific actions and activities compared to younger individuals. Hence, it seems plausible that the observed age-related improvement in forward model precision was, in part, due to greater motor experience amongst older participants. Therefore, future studies should aim to untangle the relative influences of age and motor experience on forward model development. Doing so would provide a more comprehensive account of how the forward model, and thereby one's capacity for agency, becomes more sophisticated from childhood to adulthood.

Alongside exploring the unique influences of age and motor experience on forward model development, future research can also expand upon the current results by examining how the forward model differs between neurotypical individuals and those with neurological conditions that have been shown to struggle with agency. For instance, past studies have shown that adults with Tourette Syndrome demonstrate a diminished SoA compared to neurotypical controls (Zapparoli et al., 2020). Therefore, it can be reasoned that individuals with this condition also have a less accurate forward model. In support of this idea, Kim et al. (2019b) reported that, when making a series of reach-and-return movements, adults with Tourette Syndrome were less accurate in adjusting the direction of their return movement to account for discrepancies between the expected and actual endpoints of their initial reaching action in comparison to neurotypical controls. This suggests that adults with Tourette Syndrome lack the ability to maintain an accurate forward model. The current results then provide a crucial comparison point which future studies can use to identify how the developmental trajectory of the forward model differs in individuals with Tourette Syndrome relative to neurotypical individuals.

Final Conclusion

To conclude, the purpose of the current thesis was to rectify the noted absence of adolescents within prior SoA development literature (e.g., Cavazzana et al., 2017), and thereby, determine the full trajectory at which the capacity to experience a veridical SoA develops from childhood to young adulthood. Given the relevance of the forward model system to the construction of a SoA (Haggard & Chambon, 2012), this aim was achieved by evaluating the impact of age on the functional efficiency of the forward model using four online behavioural tasks. Taken together, the current findings suggest that the precision of the forward model improves with age from childhood to young adulthood. More specifically, the ability to form accurate action-outcome predictions and the capacity to update learned action-outcome associations appropriately were both found to mature with age. This implies that the extent to which an individual can experience a SoA over a self-authored event improves with age across childhood, adolescence, and young adulthood. In order to solidify these conclusions, future research is needed to identify whether these findings can be replicated using samples with greater age variance.

Overall, the findings presented within this thesis make four key contributions to current knowledge. First, the current results extend our understanding of the trajectory at which SoA matures from childhood to adulthood, which has been understudied in past literature (Choudhury et al., 2007). Second, the current findings improve our knowledge of how the precision of the forward model changes with across this period, which has also been predominantly neglected in past research (Dahl et al. 2018). Thirdly, the tasks presented in this thesis each provide a novel alternative to intentional binding studies (e.g. Haggard & Clark, 2003) as a means through which to assess an individual's capacity to compute a reliable FoA. Finally, the current research provides a valuable insight into the viability of utilising online studies to collect behavioural data. Future research can build upon the findings presented within this thesis by exploring the relative influences of age and motor experience to forward model development.

References

Aagaard, J. (2019). Multitasking as distraction: A conceptual analysis of media multitasking research. *Theory & Psychology*, *29*(1), 87-99. https://doi.org/10.1177/0959354318815766

Adam, R., Bays, P. M., & Husain, M. (2012). Rapid decision-making under risk. *Cognitive Neuroscience*, *3*(1), 52-61. https://doi.org/10.1080/17588928.2011.613988

Adjerid, I., & Kelley, K. (2018). Big data in psychology: A framework for research advancement. *American Psychologist*, *73*(7), 899. https://doi.org/10.1037/amp0000190

Adolph, K. E., & Franchak, J. M. (2017). The development of motor behavior. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1-2), e1430. https://doi.org/10.1002/wcs.1430

Alahyane, N., Brien, D. C., Coe, B. C., Stroman, P. W., & Munoz, D. P. (2014). Developmental improvements in voluntary control of behavior: Effect of preparation in the fronto-parietal network? *NeuroImage*, *98*, 103-117. https://doi.org/10.1016/j.neuroimage.2014.03.008

Alibali, M. W., & Nathan, M. J. (2010). Conducting research in schools: A practical guide. *Journal of Cognition and Development*, *11*(4), 397-407. https://doi.org/10.1080/15248372.2010.516417

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*, 388-407. https://doi.org/10.3758/s13428-019-01237-x

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, *53*, 1407-1425. https://doi.org/10.3758/s13428-020-01501-5

Armitage, J., & Eerola, T. (2020). Reaction time data in music cognition: Comparison of pilot data from lab, crowdsourced, and convenience web samples. *Frontiers in Psychology*, *10*, 2883. https://doi.org/10.3389/fpsyg.2019.02883

Armstrong, R. A., & Hilton, A. C. (2010). Stepwise multiple regression. *Statistical Analysis in Microbiology: Statnotes*, 135-138. https://doi.org/10.1002/9780470905173.ch26

Asai, T. (2017). Know thy agency in predictive coding: Meta-monitoring over forward modeling. *Consciousness and Cognition*, *51*, 82-99. https://doi.org/10.1016/j.concog.2017.03.001

Assaiante, C. (2012). Action and representation of action during childhood and adolescence: A functional approach. *Neurophysiologie Clinique/Clinical Neurophysiology*, *42*(1-2), 43-51. https://doi.org/10.1016/j.neucli.2011.09.002

Aytemür, A., & Levita, L. (2021). A reduction in the implicit sense of agency during adolescence compared to childhood and adulthood. *Consciousness and Cognition*, *87*, 103060. https://doi.org/10.1016/j.concog.2020.103060

Aytemür, A., Lee, K. H., & Levita, L. (2021). Neural correlates of implicit agency during the transition from adolescence to adulthood: An ERP study. *Neuropsychologia*, *158*, 107908. https://doi.org/10.1016/j.neuropsychologia.2021.107908

Bakhshani, N. M. (2014). Impulsivity: A predisposition toward risky behaviors. *International Journal of High Risk Behaviors & Addiction*, *3*(2). https://doi.org/10.5812/ijhrba.20428

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*(9), 1175. https://doi.org/10.1037/0003-066X.44.9.1175

Barash, J., Brocas, I., Carrillo, J. D., & Kodaverdian, N. (2019). Heuristic to Bayesian: The evolution of reasoning from childhood to adulthood. *Journal of Economic Behavior & Organization*, *159*, 305-322. https://doi.org/10.1016/j.jebo.2018.05.008

Barcelo, F., Escera, C., Corral, M. J., & Periánez, J. A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *Journal of Cognitive Neuroscience*, *18*(10), 1734-1748. https://doi.org/10.1162/jocn.2006.18.10.1734

Barchard, K. A., & Williams, J. (2008). Practical advice for conducting ethical online experiments and questionnaires for United States psychologists. *Behavior Research Methods*, *40*, 1111-1128. https://doi.org/10.3758/BRM.40.4.1111

Barlaam, F., Fortin, C., Vaugoyeau, M., Schmitz, C., & Assaiante, C. (2012). Development of action representation during adolescence as assessed from anticipatory control in a bimanual load-lifting task. *Neuroscience*, *221*, 56-68. https://doi.org/10.1016/j.neuroscience.2012.06.062

Barnhoorn, J. S., Haasnoot, E., Bocanegra, B. R., & van Steenbergen, H. (2015). QRTEngine: An easy solution for running online reaction time experiments using Qualtrics. *Behavior Research Methods*, *47*, 918-929. https://doi.org/10.3758/s13428-014-0530-7

Barratt, E. S., Patton, J., Greger Olsson, N., & Zuker, G. (1981). Impulsivity and paced tapping. *Journal of Motor Behavior*, *13*(4), 286-300. https://doi.org/10.1080/00222895.1981.10735254

Bartlett, R., Wright, T., Olarinde, T., Holmes, T., Beamon, E. R., & Wallace, D. (2017). Schools as sites for recruiting participants and implementing research. *Journal of Community Health Nursing*, *34*(2), 80-88. https://doi.org/10.1080/07370016.2017.1304146

Baruch, C., Panissal-Vieu, N., & Drake, C. (2004). Preferred perceptual tempo for sound sequences: comparison of adults, children, and infants. *Perceptual and Motor Skills*, *98*(1), 325-339. https://doi.org/10.2466/pms.98.1.325-339

Beames, J. R., Christensen, H., & Werner-Seidler, A. (2021). School teachers: The forgotten frontline workers of COVID-19. *Australasian Psychiatry*, *29*(4), 420-422. https://doi.org/10.1177/10398562211006145

Bechara, A., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, *6*(2), 215-225. https://doi.org/10.1093/cercor/6.2.215

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214-1221. https://doi.org/10.1038/nn1954

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature*, *456*(7219), 245-249. https://doi.org/10.1038/nature07538

Bejjanki, V. R., Randrup, E. R., & Aslin, R. N. (2020). Young children combine sensory cues with learned information in a statistically efficient manner: But task complexity matters. *Developmental Science*, *23*(3), e12912. https://doi.org/10.1111/desc.12912

Bender, S., Weisbrod, M., Bornfleth, H., Resch, F., & Oelkers-Ax, R. (2005). How do children prepare to react? Imaging maturation of motor preparation and stimulus anticipation by late contingent negative variation. *NeuroImage*, *27*(4), 737-752. https://doi.org/10.1016/j.neuroimage.2005.05.020

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*(3), 351-368. https://doi.org/10.1093/pan/mpr057

Berniker, M., & Körding, K. (2011). Bayesian approaches to sensory integration for motor control. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(4), 419-428. https://doi.org/10.1002/wcs.125

Bethlehem, R. A., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., ... & Schaare, H. L. (2022). Brain charts for the human lifespan. *Nature*, *604*(7906), 525-533. https://doi.org/10.1038/s41586-022-04554-y

Bianco, V., Berchicci, M., Quinzi, F., Perri, R. L., Spinelli, D., & Di Russo, F. (2020). Females are more proactive, males are more reactive: neural basis of the gender-related speed/accuracy trade-off in visuo-motor tasks. *Brain Structure and Function*, *225*, 187-201. https://doi.org/10.1007/s00429-019-01998-3

Birnbaum, M. H. (2004). Methodological and Ethical Issues in Conducting Social Psychology Research via the Internet. In C. Sansone, C. C. Morf, & A. T. Panter (Eds.), *The Sage Handbook of Methods in Social Psychology* (pp. 359–382). Sage Publications, Inc.

Bishop, D. V., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, *14*(2), 402-416. https://doi.org/10.1111/j.1467-7687.2010.00990.x

Blakemore, S. J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*, *61*(2), 397-406. https://doi.org/10.1016/j.neuroimage.2011.11.080

Bolger, L. E., Bolger, L. A., O’Neill, C., Coughlan, E., O’Brien, W., Lacey, S., ... & Bardid, F. (2021). Global levels of fundamental motor skills in children: A systematic review. *Journal of Sports Sciences*, *39*(7), 717-753. https://doi.org/10.1080/02640414.2020.1841405

Boorman, E. D., Behrens, T. E., & Rushworth, M. F. (2011). Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biology*, *9*(6), e1001093. https://doi.org/10.1371/journal.pbio.1001093

Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, doing, and controlling: Implicit sense of agency over somatosensory events. *Psychological Science*, *28*(7), 882-893. https://doi.org/10.1177/0956797617697693

Braun Janzen, T., Thompson, W. F., & Ranvaud, R. (2014). A developmental study of the effect of music training on timed movements. *Frontiers in Human Neuroscience*, *8*, 801. https://doi.org/10.3389/fnhum.2014.00801

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106-113. https://doi.org/10.1016/j.tics.2011.12.010

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*, e9414. https://doi.org/10.7717/peerj.9414

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology*, *26*(2), 571-593. https://doi.org/10.1207/s15326942dn2602_3

Brown, B. J. (2019). *The neural and social correlates of automatic behaviours.* [Doctoral thesis, University of Nottingham]. Nottingham eTheses. https://eprints.nottingham.ac.uk/id/eprint/55830

Brown, J., & Kirk-Wade, E. (2021). Coronavirus: A history of ‘lockdown laws’ in England. *House of Commons Library.* https://researchbriefings.files.parliament.uk/documents/CBP-9068/CBP-9068.pdf

Browning, M., Behrens, T. E., Jocham, G., O'reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590-596. https://doi.org/10.1038/nn.3961

Burke, D., & Roodenrys, S. (2000). Implicit learning in a simple cued reaction-time task. *Learning and Motivation*, *31*(4), 364-380. https://doi.org/10.1006/lmot.2000.1062

Burnett Heyes, S., Adam, R. J., Urner, M., van der Leer, L., Bahrami, B., Bays, P. M., & Husain, M. (2012). Impulsivity and rapid decision-making for reward. *Frontiers in Psychology*, *3*, 153. https://doi.org/10.3389/fpsyg.2012.00153

Cáceres, P., & San Martín, R. (2017). Low cognitive impulsivity is associated with better gain and loss learning in a probabilistic decision-making task. *Frontiers in Psychology*, *8*, 204. https://doi.org/10.3389/fpsyg.2017.00204

Carpenter, R. H. S., Reddi, B. A. J., & Anderson, A. J. (2009). A simple two-stage model predicts response time distributions. *The Journal of Physiology*, *587*(16), 4051-4062. https://doi.org/10.1113/jphysiol.2009.173955

Carruthers, G. (2012). The case for the comparator model as an explanation of the sense of agency and its breakdowns. *Consciousness and Cognition*, *21*(1), 30-45. https://doi.org/10.1016/j.concog.2010.08.005

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156-2160. https://doi.org/10.1016/j.chb.2013.05.009

Cavanna, A. E., Servo, S., Monaco, F., & Robertson, M. M. (2009). The behavioral spectrum of Gilles de la Tourette syndrome. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *21*(1), 13-23.

Cavazzana, A., Begliomini, C., & Bisiacchi, P. S. (2014). Intentional binding effect in children: Insights from a new paradigm. *Frontiers in Human Neuroscience*, *8*, 651. https://doi.org/10.3389/fnhum.2014.00651

Cavazzana, A., Begliomini, C., & Bisiacchi, P. S. (2017). Intentional binding as a marker of agency across the lifespan. *Consciousness and Cognition*, *52*, 104-114. https://doi.org/10.1016/j.concog.2017.04.016

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology*, *37*(5), 715. https://doi.org/10.1037/0012-1649.37.5.715

Chahal, R., Delevich, K., Kirshenbaum, J. S., Borchers, L. R., Ho, T. C., & Gotlib, I. H. (2021). Sex differences in pubertal associations with fronto-accumbal white matter morphometry: Implications for understanding sensitivity to reward and punishment. *NeuroImage*, *226*, 117598. https://doi.org/10.1016/j.neuroimage.2020.117598

Chambers, C., Sokhey, T., Gaebler-Spira, D., & Körding, K. P. (2018). The development of Bayesian integration in sensorimotor estimation. *Journal of Vision*, *18*(12), 8-8.

Chambon, V., Sidarus, N., & Haggard, P. (2014). From action intentions to action effects: How does the sense of agency come about? *Frontiers in Human Neuroscience*, *8*, 320. https://doi.org/10.3389/fnhum.2014.00320

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences*, *106*(14), 5529-5533. https://doi.org/10.1073/pnas.0810002106

Choudhury, S., Charman, T., Bird, V., & Blakemore, S. J. (2007). Development of action representation during adolescence. *Neuropsychologia*, *45*(2), 255-262. https://doi.org/10.1016/j.neuropsychologia.2006.07.010

Chowdhury, T. G., Wallin-Miller, K. G., Rear, A. A., Park, J., Diaz, V., Simon, N. W., & Moghaddam, B. (2019). Sex differences in reward-and punishment-guided actions. *Cognitive, Affective, & Behavioral Neuroscience*, *19*, 1404-1417. https://doi.org/10.3758/s13415-019-00736-w

Christakou, A., Halari, R., Smith, A. B., Ifkovits, E., Brammer, M., & Rubia, K. (2009). Sex-dependent age modulation of frontostriatal and temporo-parietal activation during cognitive control. *NeuroImage*, *48*(1), 223-236. https://doi.org/10.1016/j.neuroimage.2009.06.070

Claassen, D. O., Jones, C. R., Yu, M., Dirnberger, G., Malone, T., Parkinson, M., ... & Jahanshahi, M. (2013). Deciphering the impact of cerebellar and basal ganglia dysfunction in accuracy and variability of motor timing. *Neuropsychologia*, *51*(2), 267-274. https://doi.org/10.1016/j.neuropsychologia.2012.09.018

Commodari, E., & La Rosa, V. L. (2021). Adolescents and distance learning during the first wave of the COVID-19 pandemic in Italy: What impact on students' well-being and learning processes and what future prospects? *European Journal of Investigation in Health, Psychology and Education*, *11*(3), 726-735. https://doi.org/10.3390/ejihpe11030052

Cooper, P. S., Wong, A. S., Fulham, W. R., Thienel, R., Mansfield, E., Michie, P. T., & Karayanidis, F. (2015). Theta frontoparietal connectivity associated with proactive and reactive cognitive control processes. *NeuroImage*, *108*, 354-363. https://doi.org/10.1016/j.neuroimage.2014.12.028

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PloS One*, *8*(3), e57410. https://doi.org/10.1371/journal.pone.0057410

Cyders, M. A., & Smith, G. T. (2008). Emotion-based dispositions to rash action: Positive and negative urgency. *Psychological Bulletin*, *134*(6), 807. https://doi.org/10.1037/a0013341

Cyders, M. A., Littlefield, A. K., Coffey, S., & Karyadi, K. A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. *Addictive Behaviors*, *39*(9), 1372-1376. https://doi.org/10.1016/j.addbeh.2014.02.013

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, *554*(7693), 441-450. https://doi.org/10.1038/nature25770

David, N., Stenzel, A., Schneider, T. R., & Engel, A. K. (2011). The feeling of agency: Empirical indicators for a pre-reflective level of action awareness. *Frontiers in Psychology*, *2*, 149. https://doi.org/10.3389/fpsyg.2011.00149

Davidson, M. C., Amso, D., Anderson, L. C., & 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*(11), 2037-2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Bellis, M. D., Keshavan, M. S., Beers, S. R., Hall, J., Frustaci, K., Masalehdan, A., ... & Boring, A. M. (2001). Sex differences in brain maturation during childhood and adolescence. *Cerebral Cortex*, *11*(6), 552-557. https://doi.org/10.1093/cercor/11.6.552

De Guio, F., Jacobson, S. W., Molteno, C. D., Jacobson, J. L., & Meintjes, E. M. (2012). Functional magnetic resonance imaging study comparing rhythmic finger tapping in children and adults. *Pediatric Neurology*, *46*(2), 94-100. https://doi.org/10.1016/j.pediatrneurol.2011.11.019

De Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, *48*, 1-12. https://doi.org/10.3758/s13428-015-0567-2

Debrabant, J., Gheysen, F., Vingerhoets, G., & Van Waelvelde, H. (2012). Age-related differences in predictive response timing in children: Evidence from regularly relative to irregularly paced reaction time performance. *Human Movement Science*, *31*(4), 801-810. https://doi.org/10.1016/j.humov.2011.09.006

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (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

DePasque, S., & Galván, A. (2017). Frontostriatal development and probabilistic reinforcement learning during adolescence. *Neurobiology of Learning and Memory*, *143*, 1-7. https://doi.org/10.1016/j.nlm.2017.04.009

Desantis, A., Roussel, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness and Cognition*, *20*(4), 1211-1220. https://doi.org/10.1016/j.concog.2011.02.012

Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, *4*(11), 423-431. https://doi.org/10.1016/S1364-6613(00)01537-0

Di Luca, M., & Rhodes, D. (2016). Optimal perceived timing: Integrating sensory information with dynamically updated expectations. *Scientific Reports*, *6*(1), 28563. https://doi.org/10.1038/srep28563

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Dogge, M., Schaap, M., Custers, R., Wegner, D. M., & Aarts, H. (2012). When moving without volition: Implied self-causation enhances binding strength between involuntary actions and effects. *Consciousness and Cognition*, *21*(1), 501-506. https://doi.org/10.1016/j.concog.2011.10.014

Donati, M. A., Beccari, C., Bacherini, A., Capitanucci, D., & Primi, C. (2021). Psychometric properties of the short UPPS-P scale in adolescents: gender, age invariance, and validity among italian youth. *Addictive Behaviors*, *120*, 106987. https://doi.org/10.1016/j.addbeh.2021.106987

Drewing, K., Aschersleben, G., & Li, S. C. (2006). Sensorimotor synchronization across the life span. *International Journal of Behavioral Development*, *30*(3), 280-287. https://doi.org/10.1177/0165025406066764

Drody, A. C., Pereira, E. J., & Smilek, D. (2023). A desire for distraction: uncovering the rates of media multitasking during online research studies. *Scientific Reports*, *13*(1), 781. https://doi.org/10.1038/s41598-023-27606-3

Droit-Volet, S., Delgado, M., & Rattat, A. C. (2006). The development of the ability to judge time in children. In Marrrow, J. R. (Ed.), *Focus on Child Psychology Research*, (pp. 81-104). Nova Science Publishers, Inc.

Droit-Volet, S., Meck, W. H., & Penney, T. B. (2007). Sensory modality and time perception in children and adults. *Behavioural Processes*, *74*(2), 244-250. https://doi.org/10.1016/j.beproc.2006.09.012

Dugré, J. R., Giguére, C. É., Percie du Sert, O., Potvin, S., Dumais, A., & Consortium Signature. (2019). The psychometric properties of a short UPPS-P impulsive behavior scale among psychiatric patients evaluated in an emergency setting. *Frontiers in Psychiatry*, *10*, 139. https://doi.org/10.3389/fpsyt.2019.00139

Dworkin, J., Hessel, H., Gliske, K., & Rudi, J. H. (2016). A comparison of three online recruitment strategies for engaging parents. *Family Relations*, *65*(4), 550-561. https://doi.org/10.1111/fare.12206

Dykiert, D., Der, G., Starr, J. M., & Deary, I. J. (2012). Age differences in intra-individual variability in simple and choice reaction time: Systematic review and meta-analysis. https://doi.org/10.1371/journal.pone.0045759

Eckstein, M. K., Master, S. L., Dahl, R., Wilbrecht13, L., & Collins, A. G. (2019). Modeling the development of decision making in volatile environments using strategies, reinforcement learning, and Bayesian inference. In *Conference on Cognitive Computational Neuroscience (CCN 2019)* (pp. 48-51). https://ccneuro.org/2019/proceedings/0000048.pdf

Eckstein, M. K., Master, S. L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. (2022). Reinforcement learning and Bayesian inference provide complementary models for the unique advantage of adolescents in stochastic reversal. *Developmental Cognitive Neuroscience*, *55*, 101106. https://doi.org/10.1016/j.dcn.2022.101106

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429-433. https://doi.org/10.1038/415429a

Evans, K. L., & Hampson, E. (2015). Sex differences on prefrontally-dependent cognitive tasks. *Brain and Cognition*, *93*, 42-53. https://doi.org/10.1016/j.bandc.2014.11.006

Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, *9*(4), 292-303. https://doi.org/10.1038/nrn2258

Faria, I., Diniz, A., & Barreiros, J. (2017). Manual asymmetries in bimanual isochronous tapping tasks in children. *Acta Psychologica*, *172*, 41-48. https://doi.org/10.1016/j.actpsy.2016.11.005

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Ferdinand, N. K., Becker, A. M., Kray, J., & Gehring, W. J. (2016). Feedback processing in children and adolescents: Is there a sensitivity for processing rewarding feedback? *Neuropsychologia*, *82*, 31-38. https://doi.org/10.1016/j.neuropsychologia.2016.01.007

Figueroa-Varela, M., Rodríguez-Ruiz, S., Muñoz, M. A., Santaella, F., Vila, J., & Anllo-Vento, L. (2010). Subjective and behavioral dimensions of impulsivity and their relation to a self-report adaptation of the SWAN in a sample of Spanish adolescents. *European Child & Adolescent Psychiatry*, *19,* pS21-S22.

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243-1255. https://doi.org/10.1016/j.neuron.2013.07.006

Forrest, W., Hay, C., Widdowson, A. O., & Rocque, M. (2019). Development of impulsivity and risk-seeking: Implications for the dimensionality and stability of self-control. *Criminology*, *57*(3), 512-543. https://doi.org/10.1111/1745-9125.12214

Forsberg, A., Guitard, D., Adams, E. J., Pattanakul, D., & Cowan, N. (2022). Children's long-term retention is directly constrained by their working memory capacity limitations. *Developmental Science*, *25*(2), e13164. https://doi.org/10.1111/desc.13164

Fradkin, I., Adams, R. A., Parr, T., Roiser, J. P., & Huppert, J. D. (2020). Searching for an anchor in an unpredictable world: A computational model of obsessive compulsive disorder. *Psychological Review*, *127*(5), 672. https://doi.org/10.1037/rev0000188

Franchak, J. M. (2019). Development of affordance perception and recalibration in children and adults. *Journal of Experimental Child Psychology*, *183*, 100-114. https://doi.org/10.1016/j.jecp.2019.01.016

Franken, I. H., van Strien, J. W., Nijs, I., & Muris, P. (2008). Impulsivity is associated with behavioral decision-making deficits. *Psychiatry Research*, *158*(2), 155-163. https://doi.org/10.1016/j.psychres.2007.06.002

Franklin, D. W., & Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, *72*(3), 425-442. https://doi.org/10.1016/j.neuron.2011.10.006

Friedman, M. S., Chiu, C. J., Croft, C., Guadamuz, T. E., Stall, R., & Marshal, M. P. (2016). Ethics of online assent: Comparing strategies to ensure informed assent among youth. *Journal of Empirical Research on Human Research Ethics*, *11*(1), 15-20. https://doi.org/10.1177/1556264615624

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

Gabitov, E., Lungu, O., Albouy, G., & Doyon, J. (2020). Movement errors during skilled motor performance engage distinct prediction error mechanisms. *Communications Biology*, *3*(1), 763. https://doi.org/10.1038/s42003-020-01465-4

Galesic, M. (2006). Dropouts on the web: Effects of interest and burden experienced during an online survey. *Journal of Official Statistics*, *22*(2), 313.

Garaizar, P., Cubillas, C. P., & Matute, H. (2016). A HTML5 open source tool to conduct studies based on Libet's clock paradigm. *Scientific Reports*, *6*(1), 32689. https://doi.org/10.1038/srep32689

Gentsch, A., Schütz-Bosbach, S., Endrass, T., & Kathmann, N. (2012). Dysfunctional forward model mechanisms and aberrant sense of agency in obsessive-compulsive disorder. *Biological Psychiatry*, *71*(7), 652-659. https://doi.org/10.1016/j.biopsych.2011.12.022

George, W. K. (2016). Could time be logarithmic? *Journal of Cosmology*, *26*(6), 14118-14134. https://thejournalofcosmology.com/WKGfinalRev.pdf

Gershman, S. J. (2015). A unifying probabilistic view of associative learning. *PLoS Computational Biology*, *11*(11), e1004567. https://doi.org/10.1371/journal.pcbi.1004567

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., ... & Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, *101*(21), 8174-8179. https://doi.org/10.1073/pnas.0402680101

Gomes, H., Sussman, E., Ritter, W., Kurtzberg, D., Cowan, N., & Vaughan Jr, H. G. (1999). Electrophysiological evidence of developmental changes in the duration of auditory sensory memory. *Developmental Psychology*, *35*(1), 294. https://doi.org/10.1037/0012-1649.35.1.294

Gonzalez-Gadea, M. L., Chennu, S., Bekinschtein, T. A., Rattazzi, A., Beraudi, A., Tripicchio, P., ... & Ibanez, A. (2015). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00543.2015

Gooch, C. M., Wiener, M., Wencil, E. B., & Coslett, H. B. (2010). Interval timing disruptions in subjects with cerebellar lesions. *Neuropsychologia*, *48*(4), 1022-1031. https://doi.org/10.1016/j.neuropsychologia.2009.11.028

Gopnik, A., & Bonawitz, E. (2015). Bayesian models of child development. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 75-86. https://doi.org/10.1002/wcs.1330

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., ... & Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences*, *114*(30), 7892-7899. https://doi.org/10.1073/pnas.1700811114

Gould, S. J., Cox, A. L., Brumby, D. P., & Wiseman, S. (2015). Home is where the lab is: A comparison of online and lab data from a time-sensitive study of interruption. *Human Computation*, *2*(1), 45-67.

Graziola, F., Pellorca, C., Di Criscio, L., Vigevano, F., Curatolo, P., & Capuano, A. (2020). Impaired motor timing in Tourette Syndrome: Results from a case–control study in children. *Frontiers in Neurology*, 1331. https://doi.org/10.3389/fneur.2020.552701

Gredebäck, G., Lindskog, M., Juvrud, J. C., Green, D., & Marciszko, C. (2018). Action prediction allows hypothesis testing via internal forward models at 6 months of age. *Frontiers in Psychology*, *9*, 290. https://doi.org/10.3389/fpsyg.2018.00290

Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2011). Bayes and blickets: Effects of knowledge on causal induction in children and adults. *Cognitive Science*, *35*(8), 1407-1455. https://doi.org/10.1111/j.1551-6709.2011.01203.x

Grootswagers, T. (2020). A primer on running human behavioural experiments online. *Behavior Research Methods*, *52*, 2283-2286. https://doi.org/10.3758/s13428-020-01395-3

Gu, L. L., Skierkowski, D., Florin, P., Friend, K., & Ye, Y. (2016). Facebook, Twitter, & QR codes: An exploratory trial examining the feasibility of social media mechanisms for sample recruitment. *Computers in Human Behavior*, *60*, 86-96. https://doi.org/10.1016/j.chb.2016.02.006

Gvirts Probolovski, H. Z., & Dahan, A. (2021). The potential role of dopamine in mediating motor function and interpersonal synchrony. *Biomedicines*, *9*(4), 382. https://doi.org/10.3390/biomedicines9040382

Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology*, *22*(10), R390-R392.

Haggard, P., & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition*, *12*(4), 695-707. https://doi.org/10.1016/S1053-8100(03)00052-7

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382-385. https://doi.org/10.1038/nn827

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196-207. https://doi.org/10.1038/nrn.2017.14

Hallez, Q., Damsma, A., Rhodes, D., Van Rijn, H., & Droit-Volet, S. (2019). The dynamic effect of context on interval timing in children and adults. *Acta Psychologica*, *192*, 87-93. https://doi.org/10.1016/j.actpsy.2018.10.004

Hammerschmidt, D., Frieler, K., & Wöllner, C. (2021). Spontaneous motor tempo: Investigating psychological, chronobiological, and demographic factors in a large-scale online tapping experiment. *Frontiers in Psychology*, *12*, 677201. https://doi.org/10.3389/fpsyg.2021.677201

Harden, K. P., & Tucker-Drob, E. M. (2011). Individual differences in the development of sensation seeking and impulsivity during adolescence: Further evidence for a dual systems model. *Developmental Psychology*, *47*(3), 739. https://doi.org/10.1037/a0023279

Hartley, C. A., & Somerville, L. H. (2015). The neuroscience of adolescent decision-making. *Current Opinion in Behavioral Sciences*, *5*, 108-115. https://doi.org/10.1016/j.cobeha.2015.09.004

Haswell, C. C., Izawa, J., Dowell, L. R., Mostofsky, S. H., & Shadmehr, R. (2009). Representation of internal models of action in the autistic brain. *Nature Neuroscience*, *12*(8), 970-972. https://doi.org/10.1038/nn.2356

Hauger, D., Paramythis, A., & Weibelzahl, S. (2011). Using browser interaction data to determine page reading behavior. In Konstan, J.A., Conejo, R., Marzo, J.L., Oliver, N. (Eds.), *User Modeling, Adaption and Personalization: 19th International Conference 2011* (pp. 147-158). Springer. https://doi.org/10.1007/978-3-642-22362-4_13

Hauser, T. U., Iannaccone, R., Walitza, S., Brandeis, D., & Brem, S. (2015). Cognitive flexibility in adolescence: Neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *NeuroImage*, *104*, 347-354. https://doi.org/10.1016/j.neuroimage.2014.09.018

Hermann, R. P., Novak, C. B., & Mackinnon, S. E. (1996). Establishing normal values of moving two-point discrimination in children and adolescents. *Developmental Medicine & Child Neurology*, *38*(3), 255-261. https://doi.org/10.1111/j.1469-8749.1996.tb15087.x

Hilbig, B. E. (2016). Reaction time effects in lab-versus Web-based research: Experimental evidence. *Behavior Research Methods*, *48*, 1718-1724. https://doi.org/10.3758/s13428-015-0678-9

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science*, *15*(5), 688-696. https://doi.org/10.1111/j.1467-7687.2012.01171.x

Hoerger, M. (2010). Participant dropout as a function of survey length in Internet-mediated university studies: Implications for study design and voluntary participation in psychological research. *Cyberpsychology, Behavior, and Social Networking*, *13*(6), 697-700. https://doi.org/10.1089/cyber.2009.0445

Hogarth, L., Chase, H. W., & Baess, K. (2012). Impaired goal-directed behavioural control in human impulsivity. *Quarterly Journal of Experimental Psychology*, *65*(2), 305-316. http://dx.doi.org/10.1080/17470218.2010.518242

Hohwy, J. (2017). Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. *Consciousness and Cognition*, *47*, 75-85. https://doi.org/10.1016/j.concog.2016.09.004

Hove, M. J., Gravel, N., Spencer, R. M., & Valera, E. M. (2017). Finger tapping and pre-attentive sensorimotor timing in adults with ADHD. *Experimental Brain Research*, *235*, 3663-3672. https://doi.org/10.1007/s00221-017-5089-y

Howlett, M. (2022). Looking at the ‘field’ through a Zoom lens: Methodological reflections on conducting online research during a global pandemic. *Qualitative Research*, *22*(3), 387-402. https://doi.org/10.1177/146879412098569

Iselin, A. M. R., & DeCoster, J. (2009). Reactive and proactive control in incarcerated and community adolescents and young adults. *Cognitive Development*, *24*(2), 192-206. https://doi.org/10.1016/j.cogdev.2008.07.001

Jacobs, R. A., & Kruschke, J. K. (2011). Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(1), 8-21. https://doi.org/10.1002/wcs.80

Jaime, M., Longard, J., & Moore, C. (2014). Developmental changes in the visual–proprioceptive integration threshold of children. *Journal of Experimental Child Psychology*, *125*, 1-12. https://doi.org/10.1016/j.jecp.2013.11.004

Jamadar, S., Hughes, M., Fulham, W. R., Michie, P. T., & Karayanidis, F. (2010). The spatial and temporal dynamics of anticipatory preparation and response inhibition in task-switching. *NeuroImage*, *51*(1), 432-449. https://doi.org/10.1016/j.neuroimage.2010.01.090

Javadi, A. H., Schmidt, D. H., & Smolka, M. N. (2014). Adolescents adapt more slowly than adults to varying reward contingencies. *Journal of Cognitive Neuroscience*, *26*(12), 2670-2681. https://doi.org/10.1162/jocn_a_00677

Jaworska, N., & MacQueen, G. (2015). Adolescence as a unique developmental period. *Journal of Psychiatry & Neuroscience: JPN*, *40*(5), 291. https://doi.org/10.1503/jpn.150268

Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science*, *4*(2), 104-107. https://doi.org/10.1111/j.1467-9280.1993.tb00469.x

Jepma, M., Schaaf, J. V., Visser, I., & Huizenga, H. M. (2020). Uncertainty-driven regulation of learning and exploration in adolescents: A computational account. *PLoS Computational Biology*, *16*(9), e1008276. https://doi.org/10.1371/journal.pcbi.1008276

Jia, R., Guo, H., Wang, Y., & Zhang, J. (2018). Analysis and test of sound delay on web audio under different situations. In *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)* (pp. 1515-1519). IEEE https://doi.org/10.1109/ICIEA.2018.8397949

Jung, J., Jackson, S. R., Nam, K., Hollis, C., & Jackson, G. M. (2015). Enhanced saccadic control in young people with Tourette syndrome despite slowed pro-saccades. *Journal of Neuropsychology*, *9*(2), 172-183. https://doi.org/10.1111/jnp.12044

Karaminis, T., Cicchini, G. M., Neil, L., Cappagli, G., Aagten-Murphy, D., Burr, D., & Pellicano, E. (2016). Central tendency effects in time interval reproduction in autism. *Scientific Reports*, *6*(1), 1-13. https://doi.org/10.1038/srep28570

Karayanidis, F., & McKewen, M. (2021). More than "just a test"—Task-switching paradigms offer an early warning system for cognitive decline. In Federmeier, K. D. (Ed.), *Psychology of learning and motivation.* (pp. 141-193). Academic Press. https://doi.org/10.1016/bs.plm.2021.02.006

Kawato, M., & Wolpert, D. (2007). Internal models for motor control. In Bock, G. R., & Goode, J. A., *Novartis Foundation Symposium 218-Sensory Guidance of Movement: Sensory Guidance of Movement: Novartis Foundation Symposium 218* (pp. 291-307). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470515563.ch16

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*, *136*(5), 849. https://doi.org/10.1037/a0019842

Kim, L. E., Oxley, L., & Asbury, K. (2022). "My brain feels like a browser with 100 tabs open": A longitudinal study of teachers' mental health and well-being during the COVID-19 pandemic. *British Journal of Educational Psychology*, *92*(1), 299-318. https://doi.org/10.1111/bjep.12450

Kim, J., Gabriel, U., & Gygax, P. (2019a). Testing the effectiveness of the Internet-based instrument PsyToolkit: A comparison between web-based (PsyToolkit) and lab-based (E-Prime 3.0) measurements of response choice and response time in a complex psycholinguistic task. *PloS One*, *14*(9), e0221802. https://doi.org/10.1371/journal.pone.0221802

Kim, S., Jackson, G. M., Dyke, K., & Jackson, S. R. (2019b). Impaired forward model updating in young adults with Tourette Syndrome. *Brain*, *142*(1), 209-219. https://doi.org/10.1093/brain/awy306

Kirsch, W., Pfister, R., & Kunde, W. (2016). Spatial action-effect binding. *Attention, Perception, & Psychophysics*, *78*, 133-142. https://doi.org/10.3758/s13414-015-0997-z

Kirsch, W., Kunde, W., & Herbort, O. (2019). Intentional binding is unrelated to action intention. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(3), 378. https://doi.org/10.1037/xhp0000612

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?* Pion Ltd. https://pure.mpg.de/rest/items/item_1790332/component/file_3136265/content

Klevberg, G. L., & Anderson, D. I. (2002). Visual and haptic perception of postural affordances in children and adults. *Human Movement Science*, *21*(2), 169-186. https://doi.org/10.1016/S0167-9457(02)00100-8

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712-719. https://doi.org/10.1016/j.tins.2004.10.007

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review*, *17*(1), 1-14. https://doi.org/10.3758/PBR.17.1.1

Kochari, A. R. (2019). Conducting web-based experiments for numerical cognition research. *Journal of Cognition*, *2*(1). https://doi.org/10.5334/joc.85

Kononowicz, T. W., & Penney, T. B. (2016). The contingent negative variation (CNV): Timing isn't everything. *Current Opinion in Behavioral Sciences*, *8*, 231-237. https://doi.org/10.1016/j.cobeha.2016.02.022

Koolschijn, P. C. M., & Crone, E. A. (2013). Sex differences and structural brain maturation from childhood to early adulthood. *Developmental Cognitive Neuroscience*, *5*, 106-118. https://doi.org/10.1016/j.dcn.2013.02.003

Körding, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, *10*(7), 319-326. https://doi.org/10.1016/j.tics.2006.05.003

Krause, V., Pollok, B., & Schnitzler, A. (2010). Perception in action: The impact of sensory information on sensorimotor synchronization in musicians and non-musicians. *Acta Psychologica*, *133*(1), 28-37. https://doi.org/10.1016/j.actpsy.2009.08.003

Kray, J., Eber, J., & Lindenberger, U. (2004). Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. *Acta Psychologica*, *115*(2-3), 143-165. https://doi.org/10.1016/j.actpsy.2003.12.001

Kray, J., Eber, J., & Karbach, J. (2008). Verbal self-instructions in task switching: A compensatory tool for action-control deficits in childhood and old age?. *Developmental Science*, *11*(2), 223-236. https://doi.org/10.1111/j.1467-7687.2008.00673.x

Kretch, K. S., & Adolph, K. E. (2013). Cliff or step? Posture-specific learning at the edge of a drop-off. *Child Development*, *84*(1), 226-240. https://doi.org/10.1111/j.1467-8624.2012.01842.x

Kühn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. *Cortex*, *49*(7), 1935-1942. https://doi.org/10.1016/j.cortex.2012.09.002

Kushnir, T., Wellman, H. M., & Gelman, S. A. (2009). A self-agency bias in preschoolers' causal inferences. *Developmental Psychology*, *45*(2), 597. https://doi.org/10.1037/a0014727

Lacey, N. (2016). Responsibility without consciousness. *Oxford Journal of Legal Studies*, *36*(2), 219-241. https://doi.org/10.1093/ojls/gqv032

Ladouceur, C. D., Dahl, R. E., & Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Developmental Science, 10*(6), 874-891. https://doi.org/10.1111/j.1467-7687.2007.00639.x

Lakes, K. D., Swanson, J. M., & Riggs, M. (2012). The reliability and validity of the English and Spanish strengths and weaknesses of ADHD and normal behavior rating scales in a preschool sample: Continuum measures of hyperactivity and inattention. *Journal of Attention Disorders*, *16*(6), 510-516. https://doi.org/10.1177/1087054711413550

Larsen, B., & Luna, B. (2018). Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neuroscience & Biobehavioral Reviews*, *94*, 179-195. https://doi.org/10.1016/j.neubiorev.2018.09.005

Leech, N. L., Gliner, J. A., Morgan, G. A., & Harmon, R. J. (2003). Use and interpretation of multiple regression. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(6), 738-740. https://doi.org/10.1097/01.CHI.0000046845.56865.22

Legaspi, R., & Toyoizumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nature Communications*, *10*(1), 4250. https://doi.org/10.1038/s41467-019-12170-0

Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience & Biobehavioral Reviews, 30*(6), 718-729. https://doi.org/10.1016/j.neubiorev.2006.06.001

Leshem, R., & Yefet, M. (2019). Does impulsivity converge distinctively with inhibitory control? Disentangling the cold and hot aspects of inhibitory control. *Personality and Individual Differences*, *145*, 44-51. https://doi.org/10.1016/j.paid.2019.03.003

Lewis, P. A., Wing, A. M., Pope, P. A., Praamstra, P., & Miall, R. C. (2004). Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping. *Neuropsychologia*, *42*(10), 1301-1312. https://doi.org/10.1016/j.neuropsychologia.2004.03.001

Liljeholm, M. (2021). Agency and goal-directed choice. *Current Opinion in Behavioral Sciences*, *41*, 78-84. https://doi.org/10.1016/j.cobeha.2021.04.004

Lim, M. S., Jocham, G., Hunt, L. T., Behrens, T. E., & Rogers, R. D. (2015). Impulsivity and predictive control are associated with suboptimal action-selection and action-value learning in regular gamblers. *International Gambling Studies*, *15*(3), 489-505. https://doi.org/10.1080/14459795.2015.1078835

Looyestyn, J., Kernot, J., Boshoff, K., Ryan, J., Edney, S., & Maher, C. (2017). Does gamification increase engagement with online programs? A systematic review. *PloS One*, *12*(3), e0173403. https://doi.org/10.1371/journal.pone.0173403

Lorimer, S., McCormack, T., Blakey, E., Lagnado, D. A., Hoerl, C., Tecwyn, E. C., & Buehner, M. J. (2020). The developmental profile of temporal binding: From childhood to adulthood. *Quarterly Journal of Experimental Psychology*, *73*(10), 1575-1586. https://doi.org/10.1177/1747021820925075

Lourenco, F., & Casey, B. J. (2013). Adjusting behavior to changing environmental demands with development. *Neuroscience & Biobehavioral Reviews*, *37*(9), 2233-2242. https://doi.org/10.1016/j.neubiorev.2013.03.003

Lozano, Ó. M., Díaz-Batanero, C., Rojas, A. J., Pilatti, A., & Fernández-Calderón, F. (2018). Concordance between the original and short version of the Impulsive Behaviour Scale UPPS-P using an IRT model. *PLoS One*, *13*(3), e0194390. https://doi.org/10.1371/journal.pone.0194390

Lucenet, J., & Blaye, A. (2014). Age-related changes in the temporal dynamics of executive control: A study in 5-and 6-year-old children. *Frontiers in Psychology*, *5*, 831. https://doi.org/10.3389/fpsyg.2014.00831

Mackenzie, E., Berger, N., Holmes, K., & Walker, M. (2021). Online educational research with middle adolescent populations: Ethical considerations and recommendations. *Research Ethics*, *17*(2), 217-227. https://doi.org/10.1177/1747016120963160

Maes, P. J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology*, *4*, 1008. https://doi.org/10.3389/fpsyg.2013.01008

Maes, P. J. (2016). Sensorimotor grounding of musical embodiment and the role of prediction: A review. *Frontiers in Psychology*, *7*, 308. https://doi.org/10.3389/fpsyg.2016.00308

Manzi, A., Nessler, D., Czernochowski, D., & Friedman, D. (2011). The development of anticipatory cognitive control processes in task-switching: An ERP study in children, adolescents, and young adults. *Psychophysiology*, *48*(9), 1258-1275. https://doi.org/10.1111/j.1469-8986.2011.01192.x

Marek, S., & Dosenbach, N. U. (2022). The frontoparietal network: Function, electrophysiology, and importance of individual precision mapping. *Dialogues in Clinical Neuroscience, 20*(2). https://doi.org/10.31887/DCNS.2018.20.2/smarek

Marzinzik, F., Wahl, M., Krüger, D., Gentschow, L., Colla, M., & Klostermann, F. (2012). Abnormal distracter processing in adults with attention-deficit-hyperactivity disorder. *PLoS One*, *7*(3), e33691. https://doi.org/10.1371/journal.pone.0033691

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1-23. https://doi.org/10.3758/s13428-011-0124-6

Master, S. L., Eckstein, M. K., Gotlieb, N., Dahl, R., Wilbrecht, L., & Collins, A. G. (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

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*, 372-380. https://doi.org/10.3758/BF03206780

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348. https://doi.org/10.1037/0096-3445.135.3.348

McCabe, S. E. (2004). Comparison of web and mail surveys in collecting illicit drug use data: A randomized experiment. *Journal of Drug Education*, *34*(1), 61-72. https://doi.org/10.2190/4hey-vwxl-dvr3-hakv

McPherson, T., Berger, D., Alagapan, S., & Fröhlich, F. (2018). Intrinsic rhythmicity predicts synchronization-continuation entrainment performance. *Scientific Reports*, *8*(1), 11782. https://doi.org/10.1038/s41598-018-29267-z

Metcalfe, J., Eich, T. S., & Castel, A. D. (2010). Metacognition of agency across the lifespan. *Cognition*, *116*(2), 267-282. https://doi.org/10.1016/j.cognition.2010.05.009

Miller, R., Schmidt, K., Kirschbaum, C., & Enge, S. (2018). Comparability, stability, and reliability of internet-based mental chronometry in domestic and laboratory settings. *Behavior Research Methods*, *50*, 1345-1358. https://doi.org/10.3758/s13428-018-1036-5

Mills, P. F., van der Steen, M. M., Schultz, B. G., & Keller, P. E. (2015). Individual differences in temporal anticipation and adaptation during sensorimotor synchronization. *Timing & Time Perception*, *3*(1-2), 13-31. https://doi.org/10.1163/22134468-03002040

Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, *65*, 276-291. https://doi.org/10.1016/j.neubiorev.2016.03.020

Monier, F., & Droit-Volet, S. (2019). Development of sensorimotor synchronization abilities: Motor and cognitive components. *Child Neuropsychology*, *25*(8), 1043-1062. https://doi.org/10.1080/09297049.2019.1569607

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, *21*(1), 59-68. https://doi.org/10.1016/j.concog.2011.08.010

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*(1), 546-561. https://doi.org/10.1016/j.concog.2011.12.002

Mu, Y., Huang, Y., Ji, C., Gu, L., & Wu, X. (2018). Auditory over visual advantage of sensorimotor synchronization in 6-to 7-year-old children but not in 12-to 15-year-old children and adults. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(5), 818. https://doi.org/10.1037/xhp0000500

Narain, D., Remington, E. D., Zeeuw, C. I. D., & Jazayeri, M. (2018). A cerebellar mechanism for learning prior distributions of time intervals. *Nature Communications*, *9*(1), 469. https://doi.org/10.1038/s41467-017-02516-x

Neri, P. (2010). How inherently noisy is human sensory processing? *Psychonomic Bulletin & Review*, *17*(6), 802-808. https://doi.org/10.3758/PBR.17.6.802

Nobusako, S., Sakai, A., Tsujimoto, T., Shuto, T., Nishi, Y., Asano, D., ... & Nakai, A. (2018). Manual dexterity is a strong predictor of visuo-motor temporal integration in children. *Frontiers in Psychology, 9,* 948. https://doi.org/10.3389/fpsyg.2018.00948

Nobusako, S., Tsujimoto, T., Sakai, A., Shuto, T., Hashimoto, Y., Furukawa, E., ... & Morioka, S. (2020). The time window for sense of agency in school-age children is different from that in young adults. *Cognitive Development*, *54*, 100891. https://doi.org/10.1016/j.cogdev.2020.100891

Nobusako, S., Wen, W., Nagakura, Y., Tatsumi, M., Kataoka, S., Tsujimoto, T., ... & Morioka, S. (2022). Developmental changes in action-outcome regularity perceptual sensitivity and its relationship to hand motor function in 5–16-year-old children. *Scientific Reports, 12*(1), 17606. https://doi.org/10.1038/s41598-022-21827-8

Noreika, V., Falter, C. M., & Rubia, K. (2013). Timing deficits in attention-deficit/hyperactivity disorder (ADHD): Evidence from neurocognitive and neuroimaging studies. *Neuropsychologia*, *51*(2), 235-266. https://doi.org/10.1016/j.neuropsychologia.2012.09.036

Overman, W. H. (2004). Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain and Cognition*, *55*(1), 134-147. https://doi.org/10.1016/S0278-2626(03)00279-3

Padilla, M. L., Pfefferbaum, A., Sullivan, E. V., Baker, F. C., & Colrain, I. M. (2014). Dissociation of preparatory attention and response monitoring maturation during adolescence. *Clinical Neurophysiology*, *125*(5), 962-970. https://doi.org/10.1016/j.clinph.2013.10.012

Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S. J. (2016). The computational development of reinforcement learning during adolescence. *PLoS Computational Biology*, *12*(6), e1004953. https://doi.org/10.1371/journal.pcbi.1004953

Papoutsaki, A., Laskey, J., & Huang, J. (2017). Searchgazer: Webcam eye tracking for remote studies of web search. In *Proceedings of the 2017 Conference on Human Information Interaction and Retrieval* (pp. 17-26). Association for Computing Machinery. https://doi.org/10.1145/3020165.3020170

Papoutsaki, A., Gokaslan, A., Tompkin, J., He, Y., & Huang, J. (2018). The eye of the typer: A benchmark and analysis of gaze behavior during typing. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (pp. 1-9). Association for Computing Machinery. https://doi.org/10.1145/3204493.3204552

Paulus, M., Hunnius, S., Van Elk, M., & Bekkering, H. (2012). How learning to shake a rattle affects 8-month-old infants’ perception of the rattle's sound: Electrophysiological evidence for action-effect binding in infancy. *Developmental Cognitive Neuroscience*, *2*(1), 90-96. https://doi.org/10.1016/j.dcn.2011.05.006

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 1-20. https://doi.org/10.3758/s13428-021-01694-3

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195-203. https://doi.org/10.3758/s13428-018-01193-y

Peirce, J., Hirst, R., & MacAskill, M. (2022). *Building experiments in PsychoPy*. Sage.

Perchet, C., & Garcia-Larrea, L. (2005). Learning to react: Anticipatory mechanisms in children and adults during a visuospatial attention task. *Clinical Neurophysiology*, *116*(8), 1906-1917. https://doi.org/10.1016/j.clinph.2005.03.022

Plumert, J. M. (1995). Relations between children's overestimation of their physical abilities and accident proneness. *Developmental Psychology*, *31*(5), 866. https://doi.org/10.1037/0012-1649.31.5.866

Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 2½-and 4-year-old children. *International Journal of Behavioral Development*, *27*(3), 220-231. https://doi.org/10.1080/01650250244000290

Pulkkinen, M. L., & Aaltonen, J. (2003). Sense of agency in narrative processes of repeatedly convicted drunk drivers. *Counselling Psychology Quarterly*, *16*(2), 145-159. https://doi.org/10.1080/0951507031000151525

Quatman-Yates, C. C., Quatman, C. E., Meszaros, A. J., Paterno, M. V., & Hewett, T. E. (2012). A systematic review of sensorimotor function during adolescence: a developmental stage of increased motor awkwardness?. *British Journal of Sports Medicine*, *46*(9), 649-655. http://dx.doi.org/10.1136/bjsm.2010.079616

Rait, M. A., Prochaska, J. J., & Rubinstein, M. L. (2015). Recruitment of adolescents for a smoking study: use of traditional strategies and social media. *Translational Behavioral Medicine*, *5*(3), 254-259. https://doi.org/10.1007/s13142-015-0312-5

Reimers, S., & Maylor, E. A. (2005). Task switching across the life span: Effects of age on general and specific switch costs. *Developmental Psychology*, *41*(4), 661. https://doi.org/10.1037/0012-1649.41.4.661

Render, A., & Jansen, P. (2021). Influence of arousal on intentional binding: Impaired action binding, intact outcome binding. *Attention, Perception, & Psychophysics*, *83*, 103-113. https://doi.org/10.3758/s13414-020-02105-z

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*, 403-452. https://doi.org/10.3758/s13423-012-0371-2

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969-992. https://doi.org/10.3758/BF03206433

Rhodes, M., Rizzo, M. T., Foster-Hanson, E., Moty, K., Leshin, R. A., Wang, M., ... & Ocampo, J. D. (2020). Advancing developmental science via unmoderated remote research with children. *Journal of Cognition and Development*, *21*(4), 477-493. https://doi.org/10.1080/15248372.2020.1797751

Ruess, M., Thomaschke, R., & Kiesel, A. (2018). Intentional binding of visual effects. *Attention, Perception, & Psychophysics*, *80*, 713-722. https://doi.org/10.3758/s13414-017-1479-2

Saarikallio, S. H., Randall, W. M., & Baltazar, M. (2020). Music listening for supporting adolescents' sense of agency in daily life. *Frontiers in Psychology*, *10*, 2911. https://doi.org/10.3389/fpsyg.2019.02911

Saito, N., Takahata, K., Murai, T., & Takahashi, H. (2015). Discrepancy between explicit judgement of agency and implicit feeling of agency: Implications for sense of agency and its disorders. *Consciousness and Cognition*, *37*, 1-7. https://doi.org/10.1016/j.concog.2015.07.011

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, *94*(3), 241-255. https://doi.org/10.1016/j.cognition.2004.04.003

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences*, *10*(4), 251. https://doi.org/10.3390/brainsci10040251

Scheerer, N. E., Jacobson, D. S., & Jones, J. A. (2016). Sensorimotor learning in children and adults: Exposure to frequency-altered auditory feedback during speech production. *Neuroscience*, *314*, 106-115. https://doi.org/10.1016/j.neuroscience.2015.11.037

Schultchen, D., Zaudig, M., Krauseneck, T., Berberich, G., & Pollatos, O. (2019). Interoceptive deficits in patients with Obsessive-Compulsive Disorder in the time course of cognitive-behavioral therapy. *PLoS One*, *14*(5), e0217237. https://doi.org/10.1371/journal.pone.0217237

Schwartze, M., Keller, P. E., Patel, A. D., & Kotz, S. A. (2011). The impact of basal ganglia lesions on sensorimotor synchronization, spontaneous motor tempo, and the detection of tempo changes. *Behavioural Brain Research*, *216*(2), 685-691. https://doi.org/10.1016/j.bbr.2010.09.015

Schwertman, N. C., Owens, M. A., & Adnan, R. (2004). A simple more general boxplot method for identifying outliers. *Computational Statistics & Data Analysis*, *47*(1), 165-174. https://doi.org/10.1016/j.csda.2003.10.012

Scott, K., & Schulz, L. (2017). Lookit (part 1): A new online platform for developmental research. *Open Mind*, *1*(1), 4-14. https://doi.org/10.1162/OPMI_a_00002

Seghezzi, S., & Zapparoli, L. (2020). Predicting the sensory consequences of self-generated actions: Pre-supplementary motor area as supra-modal hub in the sense of agency experience. *Brain Sciences*, *10*(11), 825. https://doi.org/10.3390/brainsci10110825

Seghezzi, S., Zirone, E., Paulesu, E., & Zapparoli, L. (2019). The brain in (willed) action: A meta-analytical comparison of imaging studies on motor intentionality and sense of agency. *Frontiers in Psychology*, *10*, 804. https://doi.org/10.3389/fpsyg.2019.00804

Semmelmann, K., Hönekopp, A., & Weigelt, S. (2017). Looking tasks online: Utilizing webcams to collect video data from home. *Frontiers in Psychology*, *8*, 1582. https://doi.org/10.3389/fpsyg.2017.01582

Semmelmann, K., & Weigelt, S. (2018). Online webcam-based eye tracking in cognitive science: A first look. *Behavior Research Methods*, *50*, 451-465. https://doi.org/10.3758/s13428-017-0913-7

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., ... & Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience*, *28*(14), 3586-3594. https://doi.org/10.1523/JNEUROSCI.5309-07.2008

Shi, Z., Church, R. M., & Meck, W. H. (2013). Bayesian optimization of time perception. *Trends in Cognitive Sciences*, *17*(11), 556-564. https://doi.org/10.1016/j.tics.2013.09.009

Sidarus, N., Travers, E., Haggard, P., & Beyer, F. (2020). How social contexts affect cognition: Mentalizing interferes with sense of agency during voluntary action. *Journal of Experimental Social Psychology*, *89*, 103994. https://doi.org/10.1016/j.jesp.2020.103994

Siegel, L. S. (1993). Amazing new discovery: Piaget was wrong! *Canadian Psychology/ Psychologie Canadienne*, *34*(3), 239–245. https://doi.org/10.1037/h0078835

Sisk, L. M., & Gee, D. G. (2022). Stress and adolescence: Vulnerability and opportunity during a sensitive window of development. *Current Opinion in Psychology*, *44*, 286-292. https://doi.org/10.1016/j.copsyc.2021.10.005

Slater, J., Ashley, R., Tierney, A., & Kraus, N. (2018). Got rhythm? Better inhibitory control is linked with more consistent drumming and enhanced neural tracking of the musical beat in adult percussionists and nonpercussionists. *Journal of Cognitive Neuroscience*, *30*(1), 14-24. https://doi.org/10.1162/jocn_a_01189

Slim, M. S., & Hartsuiker, R. J. (2022). Moving visual world experiments online? A web-based replication of Dijkgraaf, Hartsuiker, and Duyck (2017) using PCIbex and WebGazer.js. *Behavior Research Methods*, 1-19. https://doi.org/10.3758/s13428-022-01989-z

Smith, A. B., Giampietro, V., Brammer, M., Halari, R., Simmons, A., & Rubia, K. (2011). Functional development of fronto-striato-parietal networks associated with time perception. *Frontiers in Human Neuroscience*, *5,* 136. https://doi.org/10.3389/fnhum.2011.00136

Smittenaar, P., Rutledge, R. B., Zeidman, P., Adams, R. A., Brown, H., Lewis, G., & Dolan, R. J. (2015). Proactive and reactive response inhibition across the lifespan. *PLoS One*, *10*(10), e0140383. https://doi.org/10.1371/journal.pone.0140383

Snyder, A. C., Byron, M. Y., & Smith, M. A. (2021). A stable population code for attention in prefrontal cortex leads a dynamic attention code in visual cortex. *Journal of Neuroscience*, *41*(44), 9163-9176. https://doi.org/10.1523/JNEUROSCI.0608-21.2021

Sobel, D. M., & Munro, S. (2006. When Mr. Blicket wants it, children are Bayesian. In *Proceedings of the Annual Meeting of the Cognitive Science Society*. (pp. 810-816). https://escholarship.org/uc/item/7fk9h80q

Somerville, L. H., Sasse, S. F., Garrad, M. C., Drysdale, A. T., Abi Akar, N., Insel, C., & Wilson, R. C. (2017). Charting the expansion of strategic exploratory behavior during adolescence. *Journal of Experimental Psychology: General*, *146*(2), 155. https://doi.org/10.1037/xge0000250

Sowell, E. R., Thompson, P. M., Holmes, C. J., Batth, R., Jernigan, T. L., & Toga, A. W. (1999). Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *NeuroImage*, *9*(6), 587-597. https://doi.org/10.1006/nimg.1999.0436

Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self-and external-agency attribution: A brief review and meta-analysis. *Brain Structure and Function*, *216*, 151-157. https://doi.org/10.1007/s00429-010-0298-1

Subramaniam, K. (2021). The role of the medial prefontal cortex in self-agency in Schizophrenia. *Journal of Psychiatry and Brain Science*, *6*. https://doi.org/10.20900/jpbs.20210017

Sumich, A. L., Sarkar, S., Hermens, D. F., Ibrahimovic, A., Kelesidi, K., Wilson, D., & Rubia, K. (2012). Sex differences in brain maturation as measured using event-related potentials. *Developmental Neuropsychology*, *37*(5), 415-433. https://doi.org/10.1080/87565641.2011.653461

Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional binding without intentional action. *Psychological Science*, *30*(6), 842-853. https://doi.org/10.1177/0956797619842191

Swanson, J., Deutsch, C., Cantwell, D., Posner, M., Kennedy, J. L., Barr, C. L., ... & Wasdell, M. (2001). Genes and Attention-Deficit Hyperactivity Disorder. *Clinical Neuroscience Research*, *1*(3), 207-216. https://doi.org/10.1016/S1566-2772(01)00007-X

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219-239. https://doi.org/10.1016/j.concog.2007.03.010

Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: An interplay between prediction and postdiction. *Frontiers in Psychology*, *4*, 127. https://doi.org/10.3389/fpsyg.2013.00127

Tahej, P. K., Ferrel-Chapus, C., Olivier, I., Ginhac, D., & Rolland, J. P. (2012). Multiple representations and mechanisms for visuomotor adaptation in young children. *Human Movement Science*, *31*(6), 1425-1435. https://doi.org/10.1016/j.humov.2012.02.016

Tanaka, M., Kunimatsu, J., Suzuki, T. W., Kameda, M., Ohmae, S., Uematsu, A., & Takeya, R. (2021). Roles of the cerebellum in motor preparation and prediction of timing. *Neuroscience*, *462*, 220-234. https://doi.org/10.1016/j.neuroscience.2020.04.039

Tanaka, H., Ishikawa, T., Lee, J., & Kakei, S. (2020). The cerebro-cerebellum as a locus of forward model: A review. *Frontiers in Systems Neuroscience*, *14*, 19. https://doi.org/10.3389/fnsys.2020.00019

Thillay, A., Roux, S., Gissot, V., Carteau-Martin, I., Knight, R. T., Bonnet-Brilhault, F., & Bidet-Caulet, A. (2015). Sustained attention and prediction: Distinct brain maturation trajectories during adolescence. *Frontiers in Human Neuroscience*, *9*, 519. https://doi.org/10.3389/fnhum.2015.00519

Thompson, E. C., White-Schwoch, T., Tierney, A., & Kraus, N. (2015). Beat synchronization across the lifespan: Intersection of development and musical experience. *PLoS One*, *10*(6), e0128839. https://doi.org/10.1371/journal.pone.0128839

Tukey, J. W. (1977). *Exploratory data analysis*. Addison Wesley Publishing Company.

Valera, E. M., Spencer, R. M., Zeffiro, T. A., Makris, N., Spencer, T. J., Faraone, S. V., ... & Seidman, L. J. (2010). Neural substrates of impaired sensorimotor timing in adult Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *68*(4), 359-367. https://doi.org/10.1016/j.biopsych.2010.05.012

van Beers, R. J., Wolpert, D. M., & Haggard, P. (2002). When feeling is more important than seeing in sensorimotor adaptation. *Current Biology*, *12*(10), 834-837. https://doi.org/10.1016/S0960-9822(02)00836-9

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., De-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, *121*(4), 649. https://doi.org/10.1037/a0037665

Van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum–medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral Cortex*, *22*(6), 1247-1255. https://doi.org/10.1093/cercor/bhr198

van der Schaaf, M. E., Warmerdam, E., Crone, E. A., & Cools, R. (2011). Distinct linear and non-linear trajectories of reward and punishment reversal learning during development: Relevance for dopamine's role in adolescent decision making. *Developmental Cognitive Neuroscience*, *1*(4), 578-590. https://doi.org/10.1016/j.dcn.2011.06.007

Van Duijvenvoorde, A. C., Jansen, B. R., Griffioen, E. S., Van der Molen, M. W., & Huizenga, H. M. (2013). Decomposing developmental differences in probabilistic feedback learning: A combined performance and heart-rate analysis. *Biological Psychology*, *93*(1), 175-183. https://doi.org/10.1016/j.biopsycho.2013.01.006

van Duijvenvoorde, A. C., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *Journal of Neuroscience*, *28*(38), 9495-9503. https://doi.org/10.1523/JNEUROSCI.1485-08.2008

van Elk, M., Rutjens, B. T., & van der Pligt, J. (2015). The development of the illusion of control and sense of agency in 7-to-12-year old children and adults. *Cognition*, *145*, 1-12. https://doi.org/10.1016/j.cognition.2015.08.004

Van Gerven, P. W., Hurks, P. P., Bovend'Eerdt, T. J., & Adam, J. J. (2016). Switch hands! Mapping proactive and reactive cognitive control across the life span. *Developmental Psychology*, *52*(6), 960. https://doi.org/10.1037/dev0000116 van Kemenade, B. M., Arikan, B. E., Podranski, K., Steinsträter, O., Kircher, T., & Straube, B. (2019). Distinct roles for the cerebellum, angular gyrus, and middle temporal gyrus in action–feedback monitoring. *Cerebral Cortex*, *29*(4), 1520-1531. https://doi.org/10.1093/cercor/bhy048

van Laarhoven, T., Stekelenburg, J. J., Eussen, M. L., & Vroomen, J. (2019). Electrophysiological alterations in motor-auditory predictive coding in Autism Spectrum Disorder. *Autism Research*, *12*(4), 589-599. https://doi.org/10.1002/aur.2087

Veale, J. F. (2014). Edinburgh handedness inventory–short form: A revised version based on confirmatory factor analysis. *Laterality: Asymmetries of Body, Brain and Cognition*, *19*(2), 164-177. https://doi.org/10.1080/1357650X.2013.783045

Vijayakumar, N., Whittle, S., Yücel, M., Dennison, M., Simmons, J., & Allen, N. B. (2014). Prefrontal structural correlates of cognitive control during adolescent development: A 4-year longitudinal study. *Journal of Cognitive Neuroscience*, *26*(5), 1118-1130. https://doi.org/10.1162/jocn_a_00549

Vilares, I., & Körding, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences*, *1224*(1), 22-39. https://doi.org/10.1111/j.1749-6632.2011.05965.x

Vilaza, G. N., Haselager, W. F. F., Campos, A. M., & Vuurpijl, L. (2014). Using games to investigate sense of agency and attribution of responsibility. *Proceedings of the 2014 SBGames.* (pp. 393 - 399). SBGames. http://www.sbgames.org/sbgames2014/papers/culture/full/Cult_Full_Using%20games%20to%20investigate.pdf

Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64*, 105-123. https://doi.org/10.1016/j.neuropsychologia.2014.08.005

Watanabe, H., & Taga, G. (2006). General to specific development of movement patterns and memory for contingency between actions and events in young infants. *Infant Behavior and Development*, *29*(3), 402-422. https://doi.org/10.1016/j.infbeh.2006.02.001

Weijs, M. L., Macartney, E., Daum, M. M., & Lenggenhager, B. (2021). Development of the bodily self: Effects of visuomotor synchrony and visual appearance on virtual embodiment in children and adults. *Journal of Experimental Child Psychology*, *210*, 105200. https://doi.org/10.1016/j.jecp.2021.105200

Weiss, C., Tsakiris, M., Haggard, P., & Schütz-Bosbach, S. (2014). Agency in the sensorimotor system and its relation to explicit action awareness. *Neuropsychologia*, *52*, 82-92. https://doi.org/10.1016/j.neuropsychologia.2013.09.034

Welniarz, Q., Worbe, Y., & Gallea, C. (2021). The forward model: A unifying theory for the role of the cerebellum in motor control and sense of agency. *Frontiers in Systems Neuroscience*, *15*, 644059. https://doi.org/10.3389/fnsys.2021.644059

Wen, W. (2019). Does delay in feedback diminish sense of agency? A review. *Consciousness and Cognition*, *73*, 102759. https://doi.org/10.1016/j.concog.2019.05.007

Whitford, T. J., Rennie, C. J., Grieve, S. M., Clark, C. R., Gordon, E., & Williams, L. M. (2007). Brain maturation in adolescence: Concurrent changes in neuroanatomy and neurophysiology. *Human Brain Mapping*, *28*(3), 228-237. https://doi.org/10.1002/hbm.20273

Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., & Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage*, *96*, 67-72. https://doi.org/10.1016/j.neuroimage.2014.03.072

Wilson, P. H., & Hyde, C. (2013). The development of rapid online control in children aged 6–12 years: Reaching performance. *Human Movement Science, 32*(5), 1138-1150. https://doi.org/10.1016/j.humov.2013.02.008

Witt, S. T., & Stevens, M. C. (2013). The role of top-down control in different phases of a sensorimotor timing task: A DCM study of adults and adolescents. *Brain Imaging and Behavior*, *7*, 260-273. https://doi.org/10.1007/s11682-013-9224-5

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology, 11*(18), R729-R732.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience, 3*(11), 1212-1217. https://doi.org/10.1038/81497

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science, 269*(5232), 1880-1882. https://doi.org/10.1126/science.7569931

Worthy, D. A., & Maddox, W. T. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W.K. Estes. *Journal of Mathematical Psychology*, *59*, 41-49. https://doi.org/10.1016/j.jmp.2013.10.001

Xia, L., Master, S. L., Eckstein, M. K., Baribault, B., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2021). Modeling changes in probabilistic reinforcement learning during adolescence. *PLoS Computational Biology*, *17*(7), e1008524. https://doi.org/10.1371/journal.pcbi.1008524

Xue, Z. X., Hu, Y. J., Wang, J., Huang, L. J., Liu, W., & Sun, F. D. (2017). Reliability and validity of the short version of UPPS-P Impulsive Behavior Scale in college students. *Chinese Journal of Clinical Psychology*, 25(4), 662–666.

Yin, S., Bi, T., Chen, A., & Egner, T. (2021). Ventromedial prefrontal cortex drives the prioritization of self-associated stimuli in working memory. *Journal of Neuroscience, 41*(9), 2012-2023. https://doi.org/10.1523/JNEUROSCI.1783-20.2020

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology*, *31*(17), R1026-R1032. https://doi.org/10.1016/j.cub.2021.07.044

Yon, D., Heyes, C., & Press, C. (2020). Beliefs and desires in the predictive brain. *Nature Communications*, *11*(1), 4404. https://doi.org/10.1038/s41467-020-18332-9

Yu, A. J., & Dayan, P. (2004). Inference, attention, and decision in a Bayesian neural architecture. In Saul, L., Weiss, Y., & Bottou, L. (Eds.), *Advances in Neural Information Processing Systems 17*. (pp. 1577–1584). MIT Press. https://proceedings.neurips.cc/paper/2004/file/0e4a2c65bdaddd66a53422d93daebe68-Paper.pdf

Yücel, M., Fornito, A., Youssef, G., Dwyer, D., Whittle, S., Wood, S. J., ... & Allen, N. B. (2012). Inhibitory control in young adolescents: The role of sex, intelligence, and temperament. *Neuropsychology*, *26*(3), 347. https://doi.org/10.1037/a0027693

Zapparoli, L., Seghezzi, S., Devoto, F., Mariano, M., Banfi, G., Porta, M., & Paulesu, E. (2020). Altered sense of agency in Gilles de la Tourette syndrome: Behavioural, clinical and functional magnetic resonance imaging findings. *Brain Communications*, *2*(2), fcaa204. https://doi.org/10.1093/braincomms/fcaa204

Zélanti, P. S., & Droit-Volet, S. (2012). Auditory and visual differences in time perception? An investigation from a developmental perspective with neuropsychological tests. *Journal of Experimental Child Psychology*, *112*(3), 296-311. https://doi.org/10.1016/j.jecp.2012.01.003

Zheng, Y., Li, Q., Wang, K., Wu, H., & Liu, X. (2015). Contextual valence modulates the neural dynamics of risk processing. *Psychophysiology*, *52*(7), 895-904. https://doi.org/10.1111/psyp.12415

Zito, G. A., Wiest, R., & Aybek, S. (2020). Neural correlates of sense of agency in motor control: A neuroimaging meta-analysis. *PLoS One*, *15*(6), e0234321. https://doi.org/10.1371/journal.pone.0234321

Appendices

Appendix A.

Edinburgh Handedness Inventory – Short Form (EHI-SF)

The Edinburgh Handedness Inventory – Short Form (EHI-SF) is a 4-item self-report scale used to measure participants' hand preference (Veale, 2014). Participants respond by indicating on a 5-point Likert scale which hand they most often use when performing various daily activities. Example items include, "using a spoon" and "throwing a ball". Participants recruited through a high school, RPS, or social media completed the EHI-SF when providing their demographic information via a survey hosted on Qualtrics. Whereas, participants recruited through SSM completed the EHI-SF during the left-hand vs right-hand task (see appendix B). Hand preference was recorded to provide additional detail about the demographics of the current sample.

Appendix B.

The Left-Hand vs Right-Hand Task

The left-hand vs right-hand task was used to measure hand preference in participants aged 4-12 recruited via SSM. The task was designed using PsychoPy (Peirce, 2019) and ran online via Pavlovia (Peirce, 2022). Stimuli consisted of five 340x340 pixel white circles. Each circle contained one Likert scale response option from the EHI-SF, such as “usually left” and “usually right”. At the beginning of the left-hand vs right-hand task, participants first saw a black instructions screen with details about how to complete the task. On each trial, participants saw a black screen with an item from the EHI-SF presented at the top of the screen. In the centre of the screen, five circles were shown. Each circle contained one Likert scale response option from the EHI-SF, such as “usually left” and “usually right”. To respond, participants clicked on the circle that best reflected their experience. There was no time limit for responses. The left-hand vs right-hand task took approximately 1-minute to complete.

Appendix C.

Chapter 4b: Determining How the Ability to Learn From Past Action Experience Changes With Age From Childhood to Adulthood

Chapter Summary

The aim of the present chapter was to determine how the ability to acquire goal-related information from past action experience and use this to inform future actions changes with age from childhood to adulthood. To achieve this aim, a Bayesian learning model was fitted to the data obtained via the experiment outlined in chapter 4. This revealed the weight that participants' tended to attribute to an average estimate of all past amber light durations (i.e., the prior) relative to the amber light duration shown on the most recent trial (i.e., the likelihood) when predicting the most probable time at which the green light would onset, and therefore, when best to respond (i.e., the posterior). In accordance with typical learning tasks (Jacobs & Kruschke, 2011; Yu & Dayan, 2004), it was expected that the weight on prior would increase over time as a greater number of amber durations were observed. Unfortunately, it was found that all participants failed to construct a reliable representation of the prior. This suggests that either the cued RT task was not suitable for this type of analysis, or the Bayesian learning model used was incorrect. Hence, the influence of age on the weight attributed to the prior and the likelihood was not examined. Suggestions for why this may have occurred are discussed.

Introduction

Chapter 4 demonstrated that the extent to which individuals' ability to form predictions about the most likely outcome of their intended action develops from childhood to adulthood. The aim of the current chapter was to extend the findings of chapter 4 by examining the extent to which sensory information gained from past iterations of a target action are used to inform subsequent forward model predictions. As a brief reminder of the cued RT task described in chapter 4, participants were instructed to click the screen in response to the onset of a green target stimulus. The target stimulus was preceded by an amber cue stimulus, which remained onscreen for a variable interval. Participants' objective during the task was to minimise the temporal discrepancy between the target stimulus' onset and their response time.

Evidently, in order to determine when best to respond on a given trial, the individual had to estimate the most probable duration of the current amber cue. From a Bayesian perspective, this prediction is formed by weighting an average estimate of all past cue durations (i.e., the prior) against the cue duration observed on the most recent trial (i.e., the likelihood; Vilares & Körding, 2011). Given that the cue duration was drawn from a Gaussian distribution, the prior estimate will be more informative of the most probable cue duration on the next trial than the likelihood. Subsequently, this suggests that on average, it is more advantageous to attribute greater weight to the prior than to the likelihood on the cued RT task. Ergo, by estimating the average weight to which the participant attributed to the prior and to the likelihood, it is possible to determine the extent to which they updated their prior knowledge in response to past action experience.

As noted in chapter 1, past research has argued that the acuity with which children can perceive the duration of sensory stimuli is inferior compared to that of adults (Zélanti & Droit-Volet, 2012). Consequently, in order to resolve the uncertainty introduced by these noisy sensory estimates, it has been suggested that children place greater weight on the prior relative to the likelihood in comparison to adults (Hallez et al., 2019). Interval reproduction tasks are often cited in support of this idea. For example, Karaminis et al. (2016) presented children aged 6-14 and adults with a series of visual stimuli, each presented for a varying duration of time. In response to each stimulus, participants were instructed to reproduce the duration of the stimulus by holding down a computer key. The task was administered to both autistic and neurotypical participants. However, when focusing purely on the results obtained for the neurotypical participants, it was found that the extent to which produced durations regressed towards the average presented duration decreased with age. This suggests that the tendency to overweight prior observations relative to the most recent sensory information when estimating the temporal duration of a target stimulus declines from childhood to adulthood.

In opposition to the idea that the weight attributed to the prior decreases with age, it may be argued that this finding is dependent on the objective of the task. Crucially, in an interval reproduction task, the participant's goal is to produce a response duration equivalent to the duration of the target stimulus (e.g., Gooch et al., 2010). Arguably, unless the target durations presented on each trial are drawn from a Gaussian distribution, the average stimuli duration observed across past trials will not be informative of the target duration. Therefore, biasing one's duration estimate more towards the prior than the likelihood will result in worse performance. Notably, the stimuli durations presented by Karaminis et al. (2016) were not drawn from a Gaussian distribution. Hence, it may be argued that the results obtained by Karaminis et al. (2016) do not indicate that children bias their posterior towards the prior to a greater extent than adults. Instead, the results merely demonstrate that adults are better than children at reducing the weight on the prior in situations for which it is advantageous to do so. In contrast, to perform well on a learning task where past observations *are* informative of the target stimulus, such as the cued RT task described in chapter 4, participants need to bias their posterior estimates more towards the prior than the likelihood (Jacobs & Kruschke, 2011; Yu & Dayan, 2004).

Given that children have been suggested to lack precision in their sensory perception relative to adults (Droit-Volet et al., 2007; Jensen & Neff, 1993; Zélanti & Droit-Volet, 2012), it may be argued that the quality of the outcome information that they are able to incorporate into their prior after executing an action will be poorer than that of adults. This suggests that the reliability of the prior, and thereby, the weight attributed to the prior over the likelihood, should increase with age from childhood to adulthood. Indeed, past literature has shown that children rely more heavily on the most recently observed trial outcomes than adults on learning tasks (Barash et al., 2019). For instance, Barash et al. (2019) instructed children, adolescents and adults to complete a probabilistic learning task with two alternative choice options. It was found that the extent to which participants' made optimal Bayesian decisions, where the full history of past outcomes were used to guide choice behaviour increased from childhood to adulthood. This supports the idea of an age-related increase in the weight attributed to the prior relative to the likelihood from childhood to adulthood.

In criticism of the results obtained by Barash et al. (2019), it should be noted that age was indexed by separating child and adolescent participants into four categorical bins based on their current year of education. Notably, progression through the school system does not necessarily coincide with motor and cognitive development. Therefore, this may have obscured the true manner at which the ability to learn from past actions changes with age. Hence, further investigation is needed in order to evaluate the conclusion drawn by Barash et al. (2019).

The Current Study

The purpose of the current study was to determine how the ability to integrate past action feedback into the prior changes with age from childhood to adulthood. This would indicate how the capacity to update the forward model after action execution matures across this period. To achieve this aim, a Bayesian learning model was fitted to the data obtained via the experiment outlined in chapter 4. For each trial, the fitted model parameters revealed the weight that participants attributed to an average estimate of all past cue durations (i.e., the prior) relative to the cue duration on the most recent trial (i.e., the likelihood) when predicting the most probable time at which the target stimulus would onset (i.e., the posterior). From this, the average weight on the prior and average weight on the likelihood across trials were calculated. To test the findings of Barash et al. (2019) against the contradictory results obtained by Karaminis et al. (2016), two hypotheses were formed:

1. It was hypothesised that the average weight on the prior would be predicted by age, with older age associated with greater average weight on the prior.

2. It was also hypothesised that the average weight on the likelihood would be predicted by age, with older age associated with lower average weight on the likelihood.

Method

Design

An independent measures design was used in the current study. The independent variables were age, sex and impulsivity. The dependent variables were the average weights attributed to the prior and the likelihood across trials. Both dependent variables provided an indication of participants' ability to learn from visual cues presented across previous trials and use this knowledge to enhance the timing of their next action.

Participants, Materials and Procedure

In the current study, data presented in chapter 4 was re-analysed from a Bayesian learning perspective. Hence, details about the sample, materials, and procedure used to collect this data can be viewed in chapter 4.

Data Analysis

To determine the estimated weight that a participant placed on the prior evidence in comparison to the likelihood evidence on each trial of the cued RT task, their RT_A data was fitted to a Bayesian learning model, as outlined by Shi et al. (2013; see equation C1). The notation, $\bar{a}_{1:x-2}$ in equation C1 refers to the prior evidence available to the participant on a given trial of the task. To clarify, *x* indicates the current trial number. Hence, the prior evidence consists of the mean amber duration across trials ranging from the first trial to the trial that occurred two trials before the current trial. Furthermore, *wp* denotes the weight placed upon the prior evidence by the participant. Whereas, *wl* refers to the weight that the participant placed upon on the likelihood evidence. When combined, *wp* and *wl* must equal 1, as this signifies the participant's full capacity for weighting evidence when forming a decision on how best to respond. Therefore, the value of *wl* may be given through 1-*wp*. The notation, a_{x-1}, represents the likelihood evidence, which was the amber duration on the previous trial. Finally, *y* represents the RT_A value that is produced by the model. Note, all amber duration values inputted into the model were first converted into logarithmic form, in keeping with Shi et al. (2013).

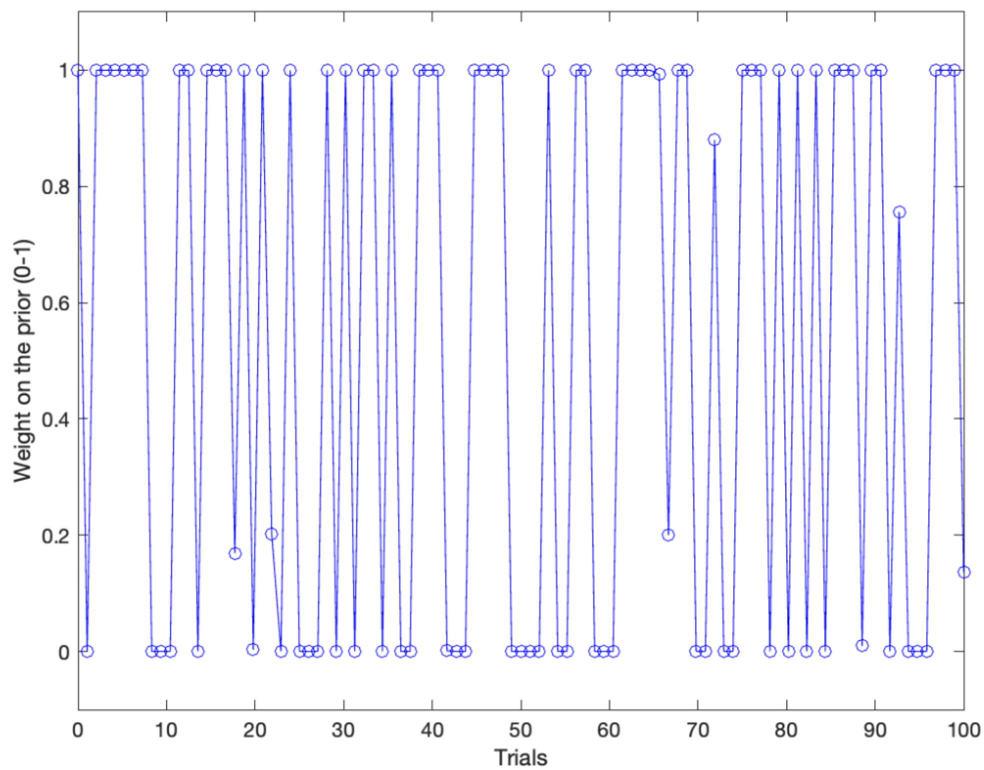
$$y = wp * \bar{a}_{1:x-2} + wl * a_{x-1} \tag{C1}$$

The model outlined in equation C1 computed RT_A values as the weighted combination of the prior evidence and likelihood evidence available to the participant on the current trial. The likelihood evidence referred to the amber duration observed on the previous trial. Past studies which have applied similar Bayesian learning models have often used information available to the participant on the current trial as the likelihood evidence (e.g., Chambers et al., 2018). This information is usually available to the participant before they are prompted to respond. However, on the cued RT task, the amber light is likely to still be present whilst the participant is actively forming their decision of when to respond. Therefore, knowledge on the current trial's amber light duration is not necessarily available to the participant when a decision to respond is made. For this reason, the amber duration observed on the previous trial was used as the likelihood evidence in this task, as this was the most recent amber light duration information available to the participant. In contrast, prior evidence consisted of the mean amber duration across trials ranging from the first trial to the trial that occurred two trials before the current trial. For instance, on trial 10, the likelihood evidence would be the amber light duration observed on trial 9. Whereas, the prior evidence would be an average of the amber light durations observed on trials 1 to 8. Both the prior evidence and likelihood evidence are then multiplied by a weight, and summed to produce a simulated RT_A value.

Initially, the weight on the prior evidence was given an arbitrary value of 0.2 and the weight on the likelihood evidence was set at 0.8. A simulated RT_A value was then produced for each trial via the model using these initial weight values. Beginning at the third trial, the participant's observed RT_A values were then entered into the model one trial at a time. It was necessary to begin at the third trial as there was insufficient prior evidence available to the participant before this point in the task. For each trial, maximum likelihood estimation was used to adjust the weights on the prior and the likelihood such that the sum of squared error between the observed and simulated RT_A values was minimised. Through this process, it was possible to obtain an estimate of the weight that the participant placed on the prior evidence and the likelihood evidence on each trial of the task (see figure C1 for an example). Following this, the mean weight on the prior and the mean weight on the likelihood across trials were calculated in order to capture whether participants tended to base their decisions more heavily on knowledge accumulated throughout the task or on recently observed information.

Figure C1.

An example of the weight on the prior plotted across trials



Note. A figure showing the estimated weight on the prior plotted across all 100 trials for participant 148. An estimated weight closer to 1 indicates that a greater proportion of the decision weight was attributed to the prior than the likelihood. Whereas, an estimated weight closer to 0 demonstrates that the opposite is true; greater weight was awarded to the likelihood information over the prior evidence.

In a learning task, one would expect the weight on the prior to increase at a variable rate as the trials progress and a greater volume of prior evidence is accumulated (Yu & Dayan, 2004). Contrary to this idea, the weight on the prior tended to oscillate between 0 (no weight on the prior) and 1 (maximum weight on the prior) for all participants in the current study, as illustrated in figure C1. This suggests that participants failed to construct a reliable representation of the prior whilst completing the cued RT task. Hence, participants did not make consistent use of their prior across trials. This suggests that the Bayesian learning model outlined in equation C1 was not suitable for the data obtained from the cued RT task. For this reason, the influence of age on the weight attributed to the prior and the likelihood cannot be examined.

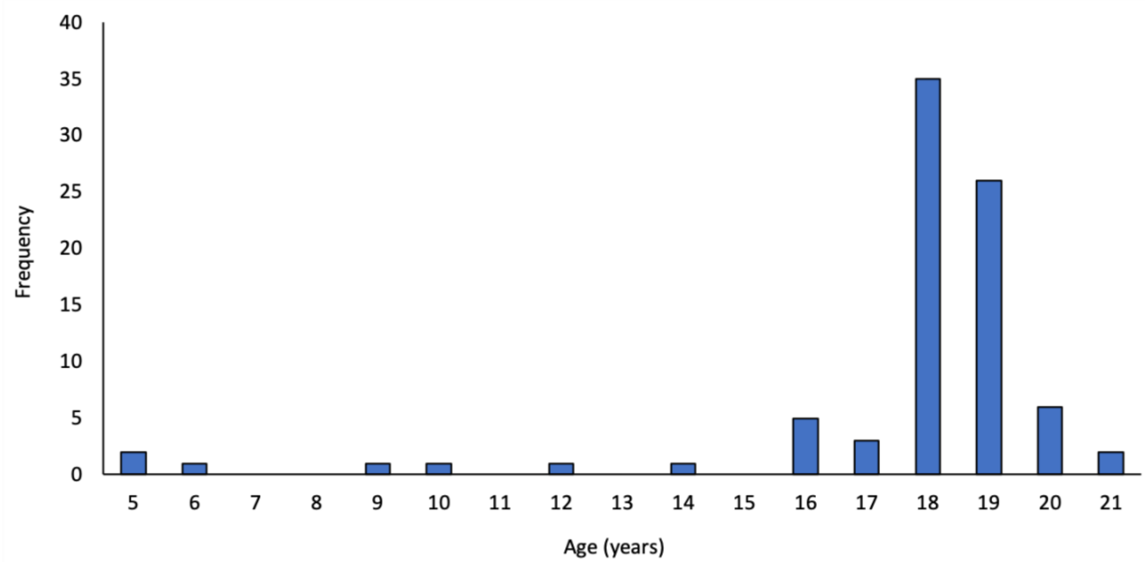
On the one hand, it can be argued that participants' failure to construct a strong prior indicates that learning about the average amber duration was simply not necessary in order to perform well on the cued RT task; accurate predictions could be made using the likelihood evidence alone. On the other hand, it may be the case that it was incorrect to define the likelihood evidence as the amber duration presented on the most recent trial. Bayesian learning studies have tended to conceptualise the most recent sensory information as part of the distribution of prior evidence (Körding & Wolpert, 2006). This suggests that the model used in the current study failed to capture the true prior and likelihood estimates used by participants. Moreover, both suggested reasons for the current results are speculative and warrant further research.

Additionally, it may be queried as to whether participants' assumptions regarding the speed at which stimuli can be presented online contributed to the lack of reliability in their prior. Garaizar et al. (2016) argued that individuals tend to expect larger delays in the presentation of expected stimuli when a behavioural task is presented online as opposed to offline. Crucially, the task instructions used in the current study did not explicitly state that the amber light duration was variable. This suggests that participants may have discounted some of the longer amber light durations as merely unintended delays in the time taken for their internet browser to present the green light. This would have then depleted the reliability of their constructed prior as a genuine estimate of the most probable amber light duration on the next trial. Future research is needed to evaluate the proposed influence of participants' beliefs about the precision of stimuli timing online on their performance in online relative to offline learning tasks.

To briefly conclude, the aim of the current study was to establish how the ability to update prior action knowledge in light of the sensory evidence gained from past experience changes with age from childhood to adulthood. Unfortunately, it was not possible to achieve this aim as all participants, regardless of age, failed to make consistent use of the information learned from past experience. Therefore, future research is required to determine how the ability to update prior action-outcome knowledge develops with age using an alternative learning task.

Appendix D

Participant Frequency as a Function of Age (Chapter 5)



Note. A figure showing the frequency of participants as a function of age in years.