Follow my Lead: Behavioural and Neural Mechanisms of Gaze Leading in Joint Attention

Lisa Joanne Stephenson
6202039
University of East Anglia
School of Psychology
August, 2018

A thesis submitted in partial fulfilment of the requirements of the University of East Anglia for the degree of Doctor of Philosophy.

Abstract

Monitoring others' actions, and our control over those actions, is essential to human social reciprocity. One such everyday social interaction is joint attention when one person follows another's direction of gaze to a referent object. When initiating joint attention (also known as "gaze leading"), reciprocal gaze responses must be processed rapidly. Therefore, we need to detect and sense agency over these social outcomes. If we cause an outcome, a compression of perception of time occurs between our action and its outcome. This phenomenon is termed temporal binding (also called intentional binding), believed to evidence an implicit sense of agency. Using a temporal binding paradigm, Experiments 1-5 evidence an implicit sense of agency for gaze shift responses to gaze leading. Using an old/new recognition paradigm, Experiments 6-7 evidence equal, high performance for recognition of unfamiliar faces for both previously encountered congruent and incongruent gaze responses to gaze leading. Experiment 8 employed electroencephalography to explore whether the neural system differentiates congruency of gaze shift elicited by gaze leading, finding, for the first time, N170-like evidence of this. Combining previous literature and the new findings in this thesis, a new neuro-cognitive model of joint and shared attention is proposed. This encapsulates the processes at work for both the gaze leader and gaze follower, the associated neural mechanisms and the subsequent social cognition processes which can ensue.

Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to express huge gratitude to my supervisor, Andrew Bayliss, whose support, wisdom, patience and kindness have known no bounds throughout my work with him. He invested in me when I was an undergraduate and has inspired and encouraged me every step of the way since. There is simply no better supervisor than him, anywhere. Secondly, I thank Gareth Edwards who has provided brilliant academic and moral support in equal measure, and has been, and continues to be, an ideal mentor.

I thank also my other supervisors, Louis Renoult and Martin Doherty, who have both given me much valued support and advice. I thank Bayliss Lab members, past and present, who all influenced my research in some way and provided excellent examples to follow, particularly Emma Howard.

I owe an enormous debt to all my fellow PhD students, past and present, at UEA School of Psychology in whose warm embrace I have shared good and bad times and always found great support, comfort and friendship and, above all, much laughter. Special mention here goes to Marianna Stella and Laura Forder. I am indebted to all the Psychology lecturers at UEA who have taught me so much in the past seven years since I started my undergraduate degree as a middle-aged student, full of trepidation, during what turned out to be the best mid-life crisis ever. I also thank all the support staff, past and present, in the School of Psychology who work so hard to keep everything running smoothly. Particular thanks to the brilliant Jackie Orford and Scott Steward. I thank, too, all the people who have participated in the research in this thesis.

Last but never least, thanks to my family, to my mother, Lyn, and step-father, Jez, and sister, Claire, who always inspire me and support me way beyond what I deserve. Thanks to my perfect son, Ben, who has grown into a wonderful young man as I have ventured through this PhD and who makes me happy every day. Finally, the greatest debt I owe is to my partner, Andrew Spencer, with whom I have been so lucky to share 25 years of unwedded bliss. He is, and always will be, my hero.

Author's declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Some of the research presented in Chapters 2 and 3 has been communicated to the scientific community by publication in *Cognition* (experiments 1, 2 & 4). This paper can be found in Appendix A.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the School of Psychology Ethics Committee at the University of East Anglia.

SECTION 1 : Introduction and Aims

Chapter 1: Introduction

"The countenance is the portrait of the soul, and the eyes mark its intentions."

Marcus Tullius Cicero, 106-43 B.C.

Aims and Objectives

The aim of this thesis is to increase existing knowledge about the neural and cognitive processes at work when we lead someone's gaze towards an object to engage in joint attention. Specifically, the effects on the joint attention initiator when a gaze bid is reciprocated or rejected will be explored. This chapter will begin by outlining in more detail the purpose and aims of the thesis in a thesis summary. Then, why these processes are an important part of our everyday social encounters will be explained, followed by a review of the relevant literature.

Thesis Summary

The primary aim of this thesis is to add to what is known about the gaze leader during a joint attention interaction, to seek to offer greater balance in the literature which has begun to investigate the gaze leader, yet still more is known about the gaze follower. I seek to do this by investigating the gaze leader's behavioural and neural mechanisms which are deployed during joint attention. One consequence of this additional data about the gaze leader will be to help inform another primary aim of the thesis which is to put forward a novel, comprehensive neuro-cognitive model of joint and shared attention. A secondary aim of the thesis is to examine any individual differences for any effects revealed in the studies, specifically, related to autism-like traits, because the reviewed literature suggests some sub-optimal joint attention behaviours. There are two strands to the behavioural consequences of gaze leading. The first strand examines an action understanding question of whether we feel a sense of agency over the responses we elicit in others' gaze behaviours. The second strand examines any effects upon facial recognition. The third strand is neural, seeking to find electrophysiological evidence that a distinction is made between eliciting congruent and incongruent responses to our joint attention bids. Figure 1 depicts these three strands (two behavioural, one neural).

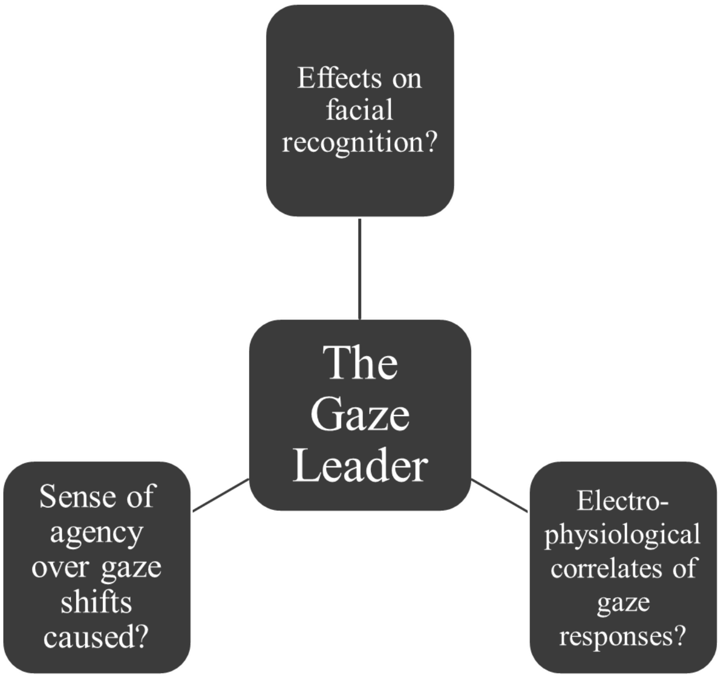

Figure 1. The three empirical strands of this thesis, investigating the gaze leader in a joint attention interaction.

The Importance of Eye Gaze Processing

The morphology of the human eye with distinctive white sclera ideally facilitates detection of gaze signals (Kobayashi & Kohshima, 2001) and so gaze can be prioritised in the visual system by an extensive neural network, identified as involved in gaze processing (see Nummenmaa & Calder, 2009; for a review). The cooperative eye hypothesis is that human eyes have evolved to serve the need for social interactions, and are, therefore, highly visible (Tomasello, Hare, Lehmann, & Call, 2007). This is important because gaze information processing can help us access other's theory of mind (see Bayliss, Frischen, Fenske, & Tipper, 2007; Emery, 2000; Langton, Watt, & Bruce, 2000, for reviews). Furthermore, gaze following is one of the key cognitive processes which enables us to learn through observation (Frith & Frith, 2007).

Direct gaze, which is often a precursor to initiating joint attention, is a crucial stimulus which we prioritise during our social cognition processing and modulates our behaviour, depending upon context (see Hamilton, 2016, for a review). There is even a clear preference for the amount of time we feel comfortable with gaze being directed towards us (Binetti, Harrison, Coutrot, Mareschal, & Johnston, 2015). Therefore, converging evidence shows that gaze processing is a key component of many aspects of social cognition.

Literature Review

This review will summarise findings about joint and 'shared attention,' the latter being a state where at least two individuals are attending the same object, and are aware of one another's states. A great deal is known about joint attention, but the specific contribution of this review will be to synthesise new insights from neuroscience and behavioural work on *initiating joint attention* (also known as "gaze leading"). This is critical for the higher-level state of shared attention.

Because the role of the initiator has been wholly neglected in the literature until the past eight years, a complete picture of both agents during shared attention has been missing but is now starting to emerge. Furthermore, those diagnosed with autism have been found to have more deficits in initiating than responding to joint attention (Mundy & Newell, 2017; Nation & Penny, 2008) and so this review will include the most recent findings about joint attention in autism. Some key terms involved in joint attention and used throughout the thesis are defined in the glossary.

Glossary

Gaze cueing: when a gaze shift towards a location causes another person to reorient their gaze towards the same location.

Gaze following: the act of following the direction of another's gaze in response to gaze cueing.

Gaze leading: the act of the joint attention initiator in attempting to cause the responder to follow their gaze.

Joint attention: a triadic interaction during which one person orients their gaze in the direction of another's gaze towards a referent object.

Shared attention: the same as joint attention except that both parties are aware of their joint attentional state.

Joint and Shared Attention

Joint attention occurs when an individual (the initiator) gazes at an object, causing another individual (the responder) to orient their gaze to the same object. *Shared* attention can be definitionally distinct from *joint* attention in that both agents are aware of their shared attentional state (Emery, 2000). However, this distinction is not always made clear. The two terms are often used interchangeably and some researchers use the term joint attention to include shared knowledge of attentional focus, whilst others do not (see Carpenter & Call, 2013, for more detailed discussion of this). I argue that it is preferable to use two different terms to make the distinction clear. Whilst they are tightly related processes, acknowledging the distinction between them allows a more nuanced examination of their underpinning cognitive mechanisms. Therefore, the definitions of joint attention and shared attention offered by Emery (2000) are adopted here; shared attention requires both parties to know they are mutually attending to the same referent object, whilst joint attention does not.

Initiating shared attention seems to set us apart as a species. Whilst some non-humans show the ability to follow gaze (e.g. macaques; Rosati, Arre, Platt, & Santos, 2016), only humans appear to initiate a share attention interaction for the sole purpose of sharing attention (Carpenter & Call, 2013). Despite this, the research has only recently begun to focus on the *initiator* of the interaction. Over the past 15-20 years, greater understanding of social cognition has resulted from extensive research into gaze following (from the *responder's* perspective) and the accompanying affective, behavioural and neural mechanisms. Much of what we already know about the role of joint attention has come from developmental work on the trajectory of infant-mother social gaze behaviours, and so it is to this work that I turn first.

Developmental Trajectory of Joint Attention

From birth, human infants show orientation towards eye contact (Farroni, Csibra, Simion, & Johnson, 2002). There is some evidence for neonates having an ability to follow eye gaze, at least if they have seen the preceding eye movement (Farroni, Massaccesi, Pividori, & Johnson, 2004), and there is evidence for gaze following ability in three month-olds (Hood, Willen, & Driver, 1998). However, despite these studies, it remains debated precisely when infants do meaningfully follow gaze cues partly because what constitutes gaze following can vary between studies. One longitudinal study found gaze following developed between two and eight months and stabilises by between six and eight months (Gredebäck, Fikke, & Melinder, 2010). Index-finger pointing then emerges at eight to twelve months and has been thought to be the first indication that the desire to share attention is developing, although there is evidence that holding out an object and giving it to an adult is a precursor to this behaviour (Cameron-Faulkner, Theakston, Lieven, & Tomasello, 2015).

Once gaze following has developed, joint attention, a pivotal part of developing social cognition can emerge. “Inter-subjectivity” is the sharing of experiences between people (Bard, 2009; Trevarthen & Aitken, 2001). In an infant’s first year the child and the primary caregiver share attention in their dyad and this is known as the “primary intersubjective” stage (Bruner & Sherwood, 1976; Terrace, 2013). At around 12 months infants ‘check back’ towards the person whose gaze was followed after following their gaze towards the object (Scaife & Bruner, 1975). This coincides with the primary intersubjective phase moving on to the “secondary intersubjective” phase in the infant’s second year. This is when child and the caregiver can start to share attention not only between themselves but including a referent object and so is when joint attention develops (Terrace, 2013).

Mother-child joint attention is positively correlated with efficiency in word learning (Tomasello & Farrar, 1986) and a longitudinal study found infants who gaze followed more at 10.5 months could produce more words associated with mental state at 2.5 years, which, in turn, also correlated with theory of mind ability at 4.5 years (Brooks & Meltzoff, 2015). Recently, jointly attending to a film alongside an experimenter was found to increase the chances of three to four year olds passing a verbal false-belief task presented in the film (Psouni et al., 2018). In addition, the frequency of engaging in joint attention predicts language acquisition (Morales et al., 2000; Mundy & Newell, 2007).

The critical age for joint attention development appears to be during the latter part of the first year of life and during the second year, with initiating joint attention developing later than responding to joint attention (Mundy et al., 2007). In sum, understanding that gaze is referential to objects and people develops by the end of the first year of life (Hoehl, Wiese, & Striano, 2008), whilst joint attention initiation develops later, by 18 months for a typically developing child (see review of joint attention development by Happé & Frith, 2014). The early emergence of joint attention typically within the first two of years of life exemplifies its key role not only in the development of language, but in social cognition processes generally.

Autism and Joint Attention

One key, diagnostic element of Autistic Spectrum Conditions (ASC) is a deficit in nonverbal communication, including eye contact abnormalities (American Psychiatric Association, 2013, who use the term "Autistic Spectrum Disorder"). Although those with autism have relatively spared gaze following behaviour, they are considered unlikely to initiate joint attention or, at least, to have atypical gaze leading behaviour (Billeci et al., 2016; Mundy & Newell, 2007; Nation & Penny, 2008), although see Gillespie-Lynch (2013), for an alternative view. Billeci et al. (2016) found that toddlers with an ASC diagnosis displayed the same eye movements as controls when responding to joint attention, but their patterns of fixations were different when initiating joint attention, for example, fixating for longer on the face than the typical controls and making more transitions from the object to the face.

Another recent study found that recognition memory for pictures was better when children had gaze led to the pictures than when they had been gaze cued to them. Critically, this was found for typically developing children but not for those with an autism diagnosis (Mundy, Kim, Mcintyre, & Lerro, 2016). Most recently, a large study of 338 toddlers made the revealing finding that, when free viewing video scenes, monozygotic twins showed remarkably similar patterns of gaze fixations on the eye regions of faces, $r = 0.91$, compared with $r = 0.35$ for dizygotic twins and no correlation for non-siblings (Constantino et al., 2017). Eye-looking at the mouth region followed a similar pattern of results to the eye region. Moreover, these apparently more highly heritable characteristics of gaze behaviour were reduced for children with autism. Children with autism looked less at eyes and mouth regions of faces than typically developing children. If children with autism look less at eye regions than typical developing children, this fits with the other reviewed findings of reduced eye contact and gaze leading in this population. A recent, revealing study showed that typically developing adults and children preferred a set of stuffed animal toys with visible white sclera over those without, whilst those with a diagnosis of autism did not (Segal, Goetz, & Maldonado, 2016). This is suggestive of the importance of eye gaze to the typical development of social cognition, and supports the cooperative eye hypothesis too.

This is important because those with autism may *appear* to lack motivation for social interaction (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012) and it is joint attention initiation which can signal social motivation to interact with others (Mundy & Newell, 2007). Chevallier et al. (2012) argue that sub-optimal social cognition in autism arises from motivational deficits rather than vice-a-versa. However, the social motivation theory of autism has increasingly been challenged. For example, a recent systematic review of empirical studies into the social motivation hypothesis identified that only 57% of reviewed studies supported the idea (Bottini, 2018). Another paper challenges the theory strongly, including pointing out that those with autism do not report lack of motivation for social interaction (Jaswal & Akhtar, 2018). It is also remarkable, for example, that implicit social biases may be relatively intact in those with an autism diagnosis (Birmingham, Stanley, Nair, & Adolphs, 2015). I do not accept that it is a lack of social motivation which drives the eye contact and gaze leading differences found in autistic individuals, but, nevertheless, these differences in gaze behaviours will inevitably impact social interaction.

There may be individual differences in the broader phenotype too. Edwards, Stephenson, Dalmaso, and Bayliss (2015), across three experiments, found a negative correlation between a “gaze leading” effect (attentional orienting towards faces which had just followed gaze) and level of autism-like traits. The greater the autism-like traits, the less attentional capture from faces who followed gaze. This indicates there may be individual differences in joint attention initiation behaviours across the typically developing population, specifically linked to the levels of typical personality traits found in those with an ASC.

There are associations between social skills and joint attention skills. For example, better joint attention skills in three year old children with an ASC have been associated with better friendships at age eight (Freeman, Gulsrud, & Kasari, 2015). Lawton & Kasari's (2012) intervention to improve joint attention initiation in preschool children with an Aby increased social interaction duration. Other interventions to improve joint attention interaction in children diagnosed with an ASC have resulted in improved language development, play skills and social development (see Goods, Ishijima, Chang, & Kasari, 2013; Murza, Schwartz, Hahs-Vaughn, & Nye, 2016; Reichow & Volkmar, 2010, for reviews). However, improvements from joint attention interventions have often proved short-lived (e.g. Whalen & Schreibman, 2003) or have not been assessed to ascertain whether the improvements are maintained (see Stavropoulos & Carver, 2013, for a review).

Stavropoulos and Carver (2013), in their systematic review, offer a potential explanation for any intervention benefits proving short-lived. The proposed explanation is the lack of social motivation in those with an ASC and so it is suggested that oxytocin may play a key role in the lack of social motivation. Stavropoulos and Carver (2013) suggest administering oxytocin, together with a joint attention intervention, may be more successful. Oxytocin certainly does seem to be involved in social motivation together with other critical neurochemicals such as dopamine (see review, Guastella & Hickie, 2016) and so this proposal is attractive. However, as Guastella & Hickie's (2016) review points out, there are many complexities involved in administering oxytocin to those with such a heterogeneous conditions as autism, and much further research and clinical trialling is needed before more conclusions can be drawn. A recent small clinical trial using intranasal oxytocin treatment took further steps towards this and found that those children with autism who had the lowest pre-existing oxytocin levels showed the most improvement in social abilities (Parker et al., 2017).

All the evidence for deficits in joint attention for those with autism have not only led to a wealth of studies on the efficacy of joint attention skills interventions (see Murza et al., 2016, for a meta-analysis), but also have fuelled debates about what we can learn about autism more generally (see Chevallier et al., 2012). In the field of autism interventions there has been a growing interest in how technology-based interventions, including virtual reality, can be utilised (Grynszpan, Weiss, Perez-Diaz, & Gal, 2014, for a meta-analysis) and the use of assistive robotics specifically is another emerging area (see Boucenna et al., 2014, for a review). Therefore, further research into autism and shared attention can serve the dual purpose of driving forward interventions to improve social skills for those with a diagnosis of autism, but also help explain the elusive question of understanding autism itself. It is encouraging to see the recent steps towards focussing upon joint attention initiation as a key deficit which can be targeted for improvement, in addition to gaze processing more generally since this was first highlighted by Mundy and Newell (2007), (see also a review by Stavropolous & Carver, 2013).

Other Animals' Gaze Behaviour

Some non-human species demonstrate the ability to follow gaze to obtain information, often about potential food or predators. For example, rhesus macaques and chimpanzees follow gaze direction of conspecifics (Tomasello, Hare, & Fogleman, 2001) and chimpanzees can shift attention between tasks. Chimpanzees use gaze and head direction cues (Tomasello et al., 2007) and also exhibit checking-back behaviours (Bräuer, Call, & Tomasello, 2005; for a detailed review see also Carpenter & Call, 2013). In rhesus macaques we know from single-cell recordings that there is a neural network which supports gaze direction encoding (Perrett et al., 1985; Perrett, Hietanen, Oram, Benson, & Rolls, 1992). There is evidence that domestic dogs show sensitivity to human visual perspectives (Kaminski, Bräuer, Call, & Tomasello, 2009) and that domestic goats can follow human gaze (Kaminski, Riedel, Call, & Tomasello, 2005).

In order to accept the argument that humans have a unique ability to share attention it must be accepted that gaze following in non-humans is simply to ascertain the object of another's attention, rather than evidences intent to share attention for some social goal in itself. Leavens and Racine (2009) argue that apes do engage in joint attention but do not define joint attention as including shared knowledge of attentional state. Leavens and Racine (2009) cite evidence of wild apes using manual gestures and captive apes using finger pointing to communicate what they want to support their view. Carpenter and Call (2013) argue this evidence is unconvincing, particularly because the cited instances of apes using declarative pointing are usually trained responses in captivity and not for sharing attention for its own sake, as human infants do. Evidence seems to favour Carpenter and Call's (2013) viewpoint (see also Tomasello & Carpenter, 2005; Tomasello, Carpenter, Call, Behne, & Moll, 2005, for other papers offering the same opinion). Further support for this comes from Warneken, Chen, and Tomasello (2006) who compared humans and chimpanzees during cooperative tasks. Children aged 12 to 24 months showed motivation to cooperative with a human adult partner, whilst young chimpanzees only cooperated if obtaining food was a goal, rather than being motivated to cooperate and share attention for its own sake.

More recently, Goot, Tomasello, and Liszkowski (2014) have shown that apes will only point towards an object they want when they are constrained in cages and need their keeper to retrieve it for them and will always move towards the object first. Conversely, human infants will point from a distance towards objects they themselves could retrieve (Goot et al., 2014). This suggests again that sharing attention simply as a cooperative, social human act is likely to be uniquely human.

Finally, there are some very intriguing findings by Rosati et al. (2016) which reveal that a free-ranging population of rhesus macaques show similar age and sex-related differences in gaze following as humans. Observing gaze following in 481 macaques, Rosati et al.'s data (2016) indicate a decreasing propensity to follow gaze in old age (Kuhn, Pagano, Maani, & Bunce, 2015) and more gaze following in females than males (Bayliss, Pellegrino, & Tipper, 2005), apparently mirroring those same findings in humans. Rosati et al. (2016) do not, however, suggest there is a shared pattern with humans of social cognitive skills development in monkeys, noting the lack of language and theory of mind development which, in contrast, human gaze following often facilitates.

I, therefore, conclude that sharing attention is an exclusively human activity as an end in itself and, therefore, has great importance in shaping our culture and success as a species, particularly when considered as an integral part of our ability to engage in social cognition including empathy and the critical ability to possess theory of mind. This accords with the view originally offered by Tomasello et al. (2005), and the recent studies just described offered further support for this assertion.

Sense of Agency and Joint Attention

Another cognitive process which is likely involved in joint and shared attention, is sense of agency, a process involved in perhaps all our motor actions. Sense of agency is experienced when we cause or generate actions, and through them, feel that we control events around us (Gallagher, 2000). One study has suggested that having a successful response to a joint attention bid creates a sense of agency in the initiator. Pfeiffer et al. (2012) measured self-reports of feeling of relatedness between the participant's gaze towards an object and the response of either gaze aversion or gaze following. The results showed that greater sense of agency (relatedness) was reported when gaze was followed than when it was not followed. In one experiment, gaze was always followed with varying latencies and participants reported feelings of relatedness with 400 ms being the interval that produced the highest ratings of relatedness and a linear decrease thereafter up to

4000 ms. However, in another experiment where gaze was either followed or averted gaze resulted, even 4000 ms latencies resulted in feelings of relatedness, there being little effect on relatedness of latency. Therefore, the optimal temporal range within which a response to shared attention initiation feels naturalistic remains a subject for future further research to build upon these findings. Such information could help inform the interventions which seek to improve social skills for those with autism which were discussed earlier. This thesis will explore whether there is an *implicit* sense of agency over gaze leading, a previously unexplored area, given Pfeiffer et al. (2012) only explored *explicit* self-reported sense of relatedness. This is measured using the phenomenon termed temporal (or intentional) binding, whereby there is a subjective compression of time between a self-generated action and its outcome (see Moore & Obhi, 2012, for a review).

Electrophysiological Correlates of Joint Attention

EEG studies which looked at gaze processing will be reviewed first, involving the N330, the N2 posterior contralateral (N2pc), the early direction of attention negativity (EDAN), and the N170 components. Then, two studies which investigated gaze leading specifically, examining the N170 and P3, will be described.

N330. Greater occipito-temporal negativity (event-related potential, ERP, component, N330) has been demonstrated in response to incongruent gaze shifts away from an object, compared to congruent (Senju, Johnson, & Csibra, 2006). The suggested explanation was that the N330 reflected the greater effort required to process the violation of the expectancy that gaze would be shifted to an object. In addition, the N330 was believed to reflect activity in the posterior superior temporal sulcus (pSTS) because corresponding fMRI data showed increased activity in response to incongruent gaze shifts. Tipples, Johnston, and Mayes (2013) also found an enhanced negative occipito-temporal ERP (occurring slightly earlier at N300) for incongruent gaze shifts. In addition, Tipples et al. (2013) found an enhanced N300 when arrows provided the directional shifts of attention, suggesting a domain general mechanism for detecting and processing unexpected events, perhaps not limited to gaze shifts. Therefore, a little is already known about ERP correlates when participants observe a face looking towards or away from an object.

EDAN. The EDAN's role during reorienting attention from gaze shifts has resulted in mixed findings. The EDAN was found to be modulated in response to spatial cues of attention from arrows, but not from eye gaze (Hietanen, Leppänen, Nummenmaa, & Astikainen, 2008). Therefore, this was thought to support the theory of different systems for attentional orienting in response to gaze, compared to other stimuli. Brignani, Guzzon, Marzi, & Miniussi (2009), reported a reverse EDAN-like effect from eye gaze shifts and a more typical EDAN-effect from arrows. More research is needed to ascertain if the EDAN is modulated by gaze cueing similarly to other attentional shifts. However, even the EDAN's role in attentional orientating is controversial, specifically whether it reflects processing the stimulus itself or orienting attention based on the directional cue being given (see Velzen & Eimer, 2003; Woodman, Arita, & Luck, 2009, for further discussion on this point).

N170. The N170 has been the subject of a large body of work showing its involvement in face processing, but it has also has been implicated specifically in gaze processing (see Itier & Batty, 2009, for a review). This has resulted in mixed findings, with some studies showing greater N170 elicited for gaze aversion over direct gaze (e.g. Latinus et al., 2014), some the opposite effect (e.g. Conty, N'Diaye, Tijus, & George, 2007), others finding no modulation at all (e.g. Myllyneva & Hietanen, 2016). Therefore, this is very much an unresolved area for future research to try to address.

N2pc. The N2pc's role is not clearly established, only being shown to our knowledge in one study to date involving eye gaze, as it is more commonly found in visual search paradigms (Grubert & Eimer, 2015). This ERP component comprises greater negative activity at the posterior sites which are contralateral to the side on which the stimuli are presented, implicated in spatial attentional shifting (Galfano et al., 2011). Galfano et al. (2011) used the N2pc as an index of spatial attention reorientation to the target needed when incongruent gaze cueing occurred. Galfano et al. (2011) predicted, and found, greater N2pcs elicited from incongruent gaze cueing than congruent.

P3 for gaze leading. Only two studies, to my knowledge, have specifically examined the neural time course of processing responses to initiating joint attention. Caruana, de Lissa, and McArthur (2015) found an enhanced central parietal P3 ERP (reported as a "P350") occurred when participants' joint attention bids were ignored (an averted gaze shift resulted) over when successfully reciprocated. Caruana et al. (2015) found no such effect when another group of participants undertook a similar task which replaced eye gaze responses with arrows. This, therefore, offers evidence of a specific social evaluation of the outcome of a joint attention bid. In a second paper, Caruana, de Lissa, and McArthur (2017) found that the P350 was not modulated by averted gaze or congruent gaze shifts when participants were expressly told that they were engaging with a computer program rather than being told that the gaze shifts they observed were being controlled by a real human. More research is needed to build upon these preliminary findings. It can be concluded from reviewing all the eye gaze associated ERPs, that more work is needed to ascertain the electrophysiological correlates of gaze cueing and, especially, of initiating shared attention. This thesis will offer some new evidence for the latter.

Current Theories and Models of Shared Attention

Baron-Cohen (1997) theorised two evolved mechanisms to facilitate cognition that another person is looking at you and whether they are sharing attention with you. For these processes he hypothesised an Eye Direction Detector and a Shared Attention Mechanism respectively. Much of the empirical evidence from both neuroscience and behavioural studies since these theories were offered have, indeed, supported the existence of such neurocognitive mechanisms (see Frischen, Bayliss, & Tipper, 2007, for a review). Baron-Cohen (1997) proposed his hypothesised mechanisms as part of a ‘mind-reading’ system. The neuro-cognitive model of shared attention offered in this thesis (see General Discussion) also places theory of mind as one of the key outcomes of sharing attention.

Seminal single-cell recording work on macaques by Dave Perrett and others (Perrett et al., 1985; Perrett et al., 1992) revealed a hierarchical system in the monkey anterior superior temporal sulcus which codes, in order, direction of gaze, head and body orientations. As Nummenmaa and Calder (2009) pointed out, there is no equivalent evidence that a hierarchical system exists in humans but it would seem reasonable for such a system to exist given the eyes offer the best clues for social attention. Evidence that human neurons are dedicated to separate coding of gaze, head and body orientation have been shown repeatedly and Nummenmaa and Calder (2009) offer a succinct review of the adaptation paradigms used to explore this separate coding system.

Dave Perrett (Perrett & Emery, 1994; Perrett et al., 1992) theorised a Direction-of-attention-detector in a similar vein to Baron-Cohen’s Eye Direction Detector and a Mutual Attention Mechanism similar to Baron-Cohen’s Shared Attention Mechanism. One of the key findings Perrett’s work added was that the system was inhibitory, such that information from the eyes is always prioritised over head and body orientations (see Langton et al., 2000, for a more detailed discussion of this). However, there is evidence that, rather than being simply inhibitory, the system may allow integration of the information from eye and head orientation, providing an attenuated effect of head information if the eye information conflicts (Langton et al., 2000).

One behavioural consequence of initiating joint attention has been found to be better memory for the pictures participants gaze led to over those they, themselves, responded to in response to gaze cueing (Kim & Mundy, 2012). Another recent finding was that jointly attending to the same side of a computer screen with a social partner increased ratings on a social binding scale, whether or not there was a shared goal (Wolf, Launay, & Dunbar, 2016). This indicates that people feel connected when jointly attending and this could be built upon by exploring whether this sense of closeness is enhanced more by initiating the joint attention interaction, rather than responding to it.

Shteynberg (2015) reviewed behavioural shared attention studies, mainly from the field of social psychology. The review includes studies which look at effects of sharing attention online, encompassing any studies in which participants believe that they are jointly attending, and so goes beyond the much more narrow definition of shared attention in this thesis which is between two people who are in a face-to-face interaction. However, Shteynberg's (2015) review of behavioural studies does demonstrate that the recruitment of increased cognitive resources seems to be one result of sharing attention. Shteynberg's (2015) model lists five empirically demonstrated effects of sharing attention which are better memory, stronger motivation, more extreme judgments, higher affective intensity and greater behavioural learning and postulates a shared-attention mechanism which helps groups to coordinate and achieve mutual acts.

Neural Mechanisms of Joint Attention

The regions involved in detecting gaze presence are the amygdala (Adolphs, 2008; Adolphs & Spezio, 2006; Gamer, Schmitz, Tittgemeyer, & Schilbach, 2013; Kawashima et al., 1999), the hippocampus and lateral fusiform gyrus and the inferior occipital gyri (reviewed in Nummenmaa & Calder, 2008). More recently, the ventromedial prefrontal cortex has been shown to play a role in the driving of attention to the eye region as this is impaired in those with lesions to this region (Wolf, Philippi, Motzkin, Baskaya, & Koenigs, 2014). Following gaze detection, the encoding of gaze direction has been implicated in the Intraparietal Sulcus (Hoffman & Haxby, 2000), Frontal Eye Fields (O'Shea, Muggleton, Cowey, & Walsh, 2004), Superior Parietal Lobule and pSTS (Calder et al., 2007), and MT/V5 complex (Watanabe, Kakigi, Miki, & Puce, 2006).

Once a gaze shift is detected, the responder reorients attention towards the initiator's gaze cued location which involves the Inferior Parietal Lobule (Calder et al., 2007; Perrett et al., 1985; Perrett et al., 1992), the Bilateral Middle Frontal Gyri, the Bilateral Superior Temporal Gyri, the Bilateral Intraparietal Sulci (Thiel, Zilles, & Fink, 2004, 2005), and the Superior Colliculus (Furlan, Smith, & Walker, 2015). Whilst the initiator is already attending to the referent object, it has been demonstrated that the face of the responder has an attentional capture effect for the initiator and so reorienting, at least, covertly towards the responder is part of the process for the initiator (Edwards et al., 2015). This "gaze leading effect" is theorised to be a mechanism which facilitates the state of joint attention to move onto the higher-level cognitive state of share attention as it enables the initiator to monitor the response of the responder (Edwards et al., 2105). In addition, those who cooperatively follow our gaze leading produce less of a gaze cueing effect in us when we subsequently re-encounter them (Dalmaso, Edwards, & Bayliss, 2016). This may indicate that shared attention is affected by previous interactions and is not exclusively an automatic process, but subject to contextual influences (Dalmaso et al., 2016).

The neural mechanisms of joint attention have been the subject of several studies over the past six years with both distinct and common regions shown to be recruited when initiating or responding to joint attention (Caruana, Brock, & Woolgar, 2015; Redcay, Kleiner, & Saxe, 2012; Schilbach et al., 2010). These regions are summarised in Table 1. Most recently, Oberwelland et al. (2016) investigated the neural mechanisms of initiating and responding to joint attention in adolescents and children and confirmed the involvement of similar regions of activity and distinct regions recruited by self-initiated or other-initiated joint attention as shown in Table 1, specifically recruiting “social brain” areas and those used for processing shifts of attention. Furthermore, that study has begun to identify the developmental trajectory of joint attention neural mechanisms, hitherto not explored, making two key findings. Firstly, there was a trend towards decreasing precuneus activation from childhood to adolescence which the authors suggest may be consistent with children becoming less self-referential during adolescence as their social interactional skills develop towards maturity. Secondly, Oberwelland et al.’s (2016) data indicate a trend towards more precuneus activation during gaze leading than responding to joint attention in adolescence which may indicate an increasing awareness of gaze leading being self-generated and self-referential. These preliminary findings show promise to elucidate the development trajectory of “social brain” developmental changes during shared attention and opens the path towards more work in this new area of research.

Table 1

fMRI Evidence for Distinct and Common Brain Regions Activated During Initiating and Responding to Joint Attention Reported by Redcay et al. (2012).

	Initiating Joint Attention	Responding to Joint Attention
Distinct Regions	Bilateral middle frontal gyri	Posterior STS
	Bilateral intraparietal sulci	Ventral mPFC
	Dorsal anterior cingulate	Posterior Cingulate
Common Regions	Dorsal mPFC	
	Right Posterior STS	

Abbreviations: STS- Superior Temporal Sulcus, mPFC- Medial Prefrontal Cortex.

Taking these findings together over the past six years, the key regions seem to be the MPFC (when a social partner is perceived) and the pSTS (when a shift in attention is detected). Both regions are recruited together during joint attention but their distinct roles and how they might interact have not yet been precisely defined (see Carlin & Calder, 2013; Redcay & Saxe, 2013, for reviews).

Another revealing recent fMRI study went further and identified functional connections between the visual and dorsal attention networks as initiating joint attention develops in toddlers in a large sample of 116 one year olds and 98 infants of 24 months, 37 of these providing behavioural and imaging data at both age points (Eggebrecht et al., 2017). Infants were assessed for their initiating joint attention abilities. Then, brain functional connectivity was measured whilst the infants slept so that correlations between joint attention initiation abilities and brain functional connectivity between regions of interest identified by the work in adults described above (e.g. Redcay et al, 2012) could be examined. Broadly, the findings were that initiating joint attention abilities was most strongly associated with connectivity between the visual and dorsal attention networks and between the visual network and posterior cingulate default mode network (Eggebrecht et al., 2017).

All of these neuroimaging findings broadly support the Parallel and Distributed-Processing Model hypothesised by Mundy & Newell (2007) and Mundy, Sullivan, and Mastergeorge (2009) in which joint attention initiation is mainly served by the anterior attention network and joint attention response by the posterior attention network, but with an emphasis on the connections between these networks.

A further intriguing neural correlate is that shown by Schilbach et al. (2010) who demonstrated enhanced ventral striatum activity for initiating joint attention, suggesting this is a rewarding experience. This activity also correlated with self-reported subjective feelings of pleasantness. The greater the activity change in the ventral striatum, the greater the sense of pleasantness reported for looking at objects with another person. In this case the other person was an onscreen face but participants were told that the onscreen face was controlled by a real person. This type of research which seeks to examine online social interactions rather than offline, has become popular in recent years (see Pfeiffer, Vogeley, & Schilbach, 2013, for a review). A further study showed that gaze based behaviours with another person activated the ventral striatum, and it did not matter whether the participants believed their partner had a shared goal or not (Pfeiffer et al., 2014). Another study found increased striatum activity when initiating joint attention was reciprocated with gaze following, compared to an averted gaze response (Gordon, Eilbott, Feldman, Pelphrey, & Vander Wyk, 2013). Finally, the ventral striatum was activated more even when participants simply passively observed actors in a video clip engaging in a shared purpose than when the actors were simply acting in parallel (Eskenazi, Rueschemeyer, de Lange, Knoblich, & Sebanz, 2015). The growing literature reviewed show there is an extensive network of brain regions sub-serving joint attention. See Mundy (2017) for a comprehensive review of these neural mechanisms. These regions will be summarised further in a model of joint and shared attention that will be summarised in Chapter 6 (General Discussion).

Concluding Remarks

Initiating joint attention, leading to a shared attention interaction, is a key human process vital to typical social cognitive development. The atypicality in engaging in shared attention found in those with autism, specifically in its initiation, illustrate the difficulties in social interaction which can result if the mechanisms involved in shared attention are not optimal. More is known about the gaze responder during this interaction than the gaze leader, although work has begun to redress this imbalance and this thesis builds on this work.

Outline of Thesis

The experiments in Chapters 2 and 3 explore whether we feel a sense of agency over the gaze shift responses we cause in others (Experiments 1 to 5). Chapter 4's studies (Experiments 6 and 7) explore whether there are any effects upon facial recognition which result from having joint attention bids reciprocated or rejected. The data in Chapter 5 (Experiment 8) reveal the electrophysiological consequences of gaze responses to joint attention bids. Chapter 6 has a general discussion of the results from the empirical chapters and offers a novel, neuro-cognitive model of joint and shared attention, to encapsulate the processes at work for both the gaze leader and gaze follower and the associated neural mechanisms.

SECTION 2: Experimental Chapters

Chapter 2: Sense of Agency Over Responses to Gaze Leading

As prosocial beings, monitoring other's actions, and our control over those actions, is essential to successful human cooperation and social reciprocity (Tomasello & Vaish, 2013). One everyday interaction which is integral to social reciprocity is joint attention when one person follows another's eye gaze to a referent object (Emery, 2000). The experiments in this chapter investigate whether initiating joint attention ("gaze leading") elicits feelings of control over responses. There is a link between perception of time and perception of control over events because, if we are in control of an outcome, a compression of perception of time occurs between our action and its outcome (David, Newen, & Vogeley, 2008, for a review). This phenomenon is termed temporal binding (also known as intentional binding) and is believed to evidence a "sense of agency" (Haggard, Clark, & Kalogeras, 2002). Critically, this compression of time does not occur when we passively observe outcomes not caused by our actions (Haggard & Clark, 2003). To date, temporal binding research has been largely dedicated to investigating sense of agency for non-social motor actions, usually using hands, such as when making button presses to elicit an auditory tone (see Moore & Obhi, 2012, for a review). The experiments in this chapter, uniquely, as far as I am aware, investigated whether this distortion of time perception also occurs when controlling others' eye gaze during the social interaction of joint attention. Figure 2 summarises the process being investigated and the expected outcome if there is an implicit sense of agency over gaze shifts we elicit.

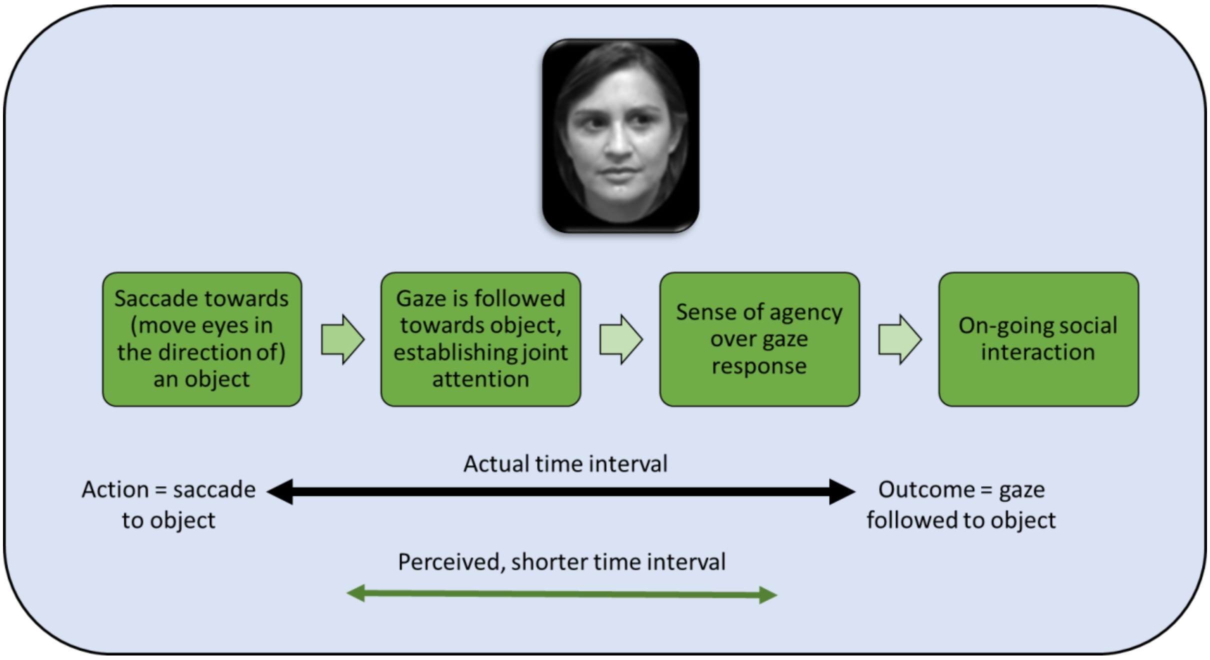

Figure 2. Graphical abstract summarising the expected temporal binding effect for causing gaze shifts in others in response to gaze leading.

I will begin by introducing the key construct of sense of agency, describing temporal binding paradigms and considering the cognitive and accompanying neural mechanisms which have been conceptualised to explain how a sense of agency may arise. Then, the key role eye gaze plays in controlling social attention and coordination will be described (see Emery, 2000; Frischen et al., 2007, for reviews). The process of controlling other's eye gaze occurs during "joint attention." A line of joint attention research will be described including findings about having one's own gaze followed. Next, research into sense of agency and joint attention for those with an autism spectrum condition (ASC) will be reviewed because there may be deficits in that population. The research question I explore in this chapter is whether temporal binding, offered as a proxy for an implicit sense of agency, results when we cause another to follow our gaze to an object. I also explore whether there is any correlation between the magnitude of any temporal binding effects and level of autism-like traits and also examine the relationship between implicit effects and explicit self-reported sense of agency.

To pre-empt the results, the explicit agency data demonstrates the greatest agentic attributions are made following active gaze leading tasks, and lower sense of agency is always reported for passive tasks involving no gaze leading. However, for implicit agency, the results are more complex, showing three main effects of interest. Firstly, temporal binding does result from gaze leading towards an object. Secondly, binding effects occur when our attention is already on an object, with no gaze leading having taken place, but these effects are attenuated compared to the binding effects following gaze leading. Finally, even low-level spatial shifts towards our object of gaze can elicit binding effects where there are no social clues to negate this self-agency attribution. This third main effect of interest was unexpected and makes the explanations for this data a little more complex than anticipated, but, I hope, more revealing and interesting at the same time.

Explicit and Implicit Agency

Sense of agency is our conscious experience of causing or generating actions, and through them, controlling events in our environment (Gallagher, 2000). Recently, David, Obhi, and Moore (2015) highlighted a rapidly expanding interest in sense of agency among the research community since 2002, with papers being published in this area exponentially. A key component of sense of agency is congruency between an action and its outcome. The hypothesis tested in this chapter's experiments is that if someone follows our gaze we will feel a sense of agency as we have caused them to respond. How to test and measure this in practice poses a challenge. Explicit measurement can be problematic because, as Gallagher

(2012) captured well, our agency is not normally something that we are explicitly aware of. Explicit sense of agency is somewhat limited to measurement through self-reported feelings of control over an action (David, Newen, et al., 2008; Sebanz & Lackner, 2007) which, in themselves, are limited by the agent's own ability for introspection (Barlas & Obhi, 2013). In addition, explicit measures have been criticised for their susceptibility to response bias and impression management (Obhi, 2012). Because of these difficulties, an alternative is to measure implicit (outside of awareness) sense of agency. Temporal binding experiments have sought to do this. Here, implicit agency via temporal binding effects is measured, alongside self-reported explicit agency.

Temporal Binding

Temporal binding is the subjective compression of a time interval between an action and its outcome (Haggard et al., 2002). Therefore, it is when a period of time is perceived to be shorter than it really is. Research into the subjective experience of time began as long ago as the 1880's with William Wundt's studies on attention. Wundt developed a "complication-clock apparatus" used for participants to report the onset of an auditory click relative to a clock hand location (Carlson, Hogendoorn, & Verstraten, 2006). Differences in the perception of the onset relative to the clock hand position were believed by Wundt to be explained by whether attention was on the clock or the auditory click (Moore and Obhi, 2012). The temporal binding paradigm itself first appeared with Haggard et al.'s (2002) introduction of a measure of sense of agency based upon the relationship between voluntary action and subjective time. Temporal binding is often called intentional binding (e.g. Desantis, Hughes, & Waszak, 2012). The term intentional binding implies that outcomes must be intended for binding to occur. There has been some debate about whether binding is due to causality (Buehner, 2012), or intentionality, or both (Cravo, Claessens, & Baldo, 2011). This is why, for clarity, the term used in this thesis is "temporal," rather than "intentional" binding, avoiding theoretical assumptions of intentionality.

Haggard et al. (2002) used a 'Libet clock' method where participants reported where the hands of a clock were when they performed actions and when there was an auditory tone, the outcome (see Figure 3). Haggard et al. (2002) found that participants perceived the time interval between their own button press (an intentional action) and the resulting tone to be shorter than it really was, whilst their perception of the interval between an action not caused by them (an unintentional action) and the outcome was estimated to be longer than it really was (see Figure 4).

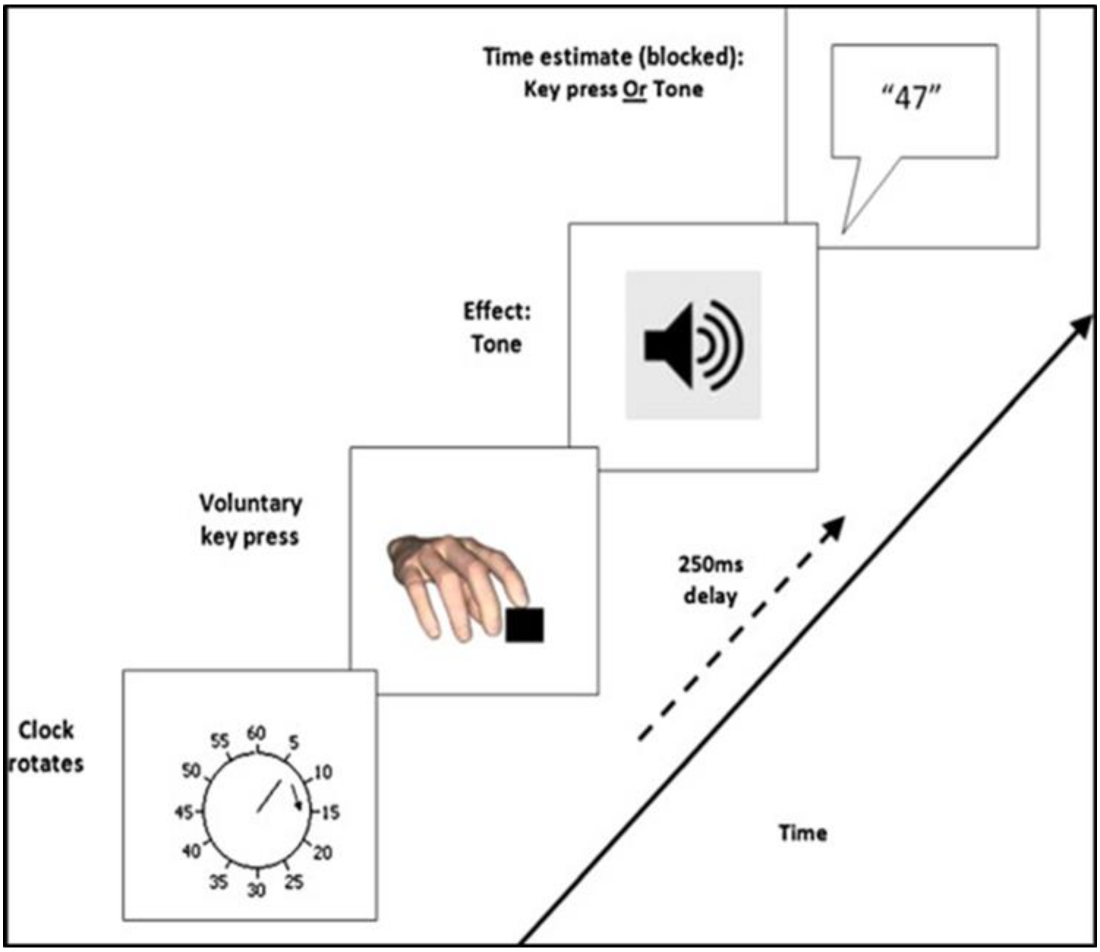

Figure 3. The Libet-clock method used by Haggard et al. (2002), figure taken from Moore and Obhi (2012). Participants pressed the key at a time of their choosing which produced a tone after a delay of 250 ms. Participants judged where the clock hand was when they pressed the key or when they heard the tone.

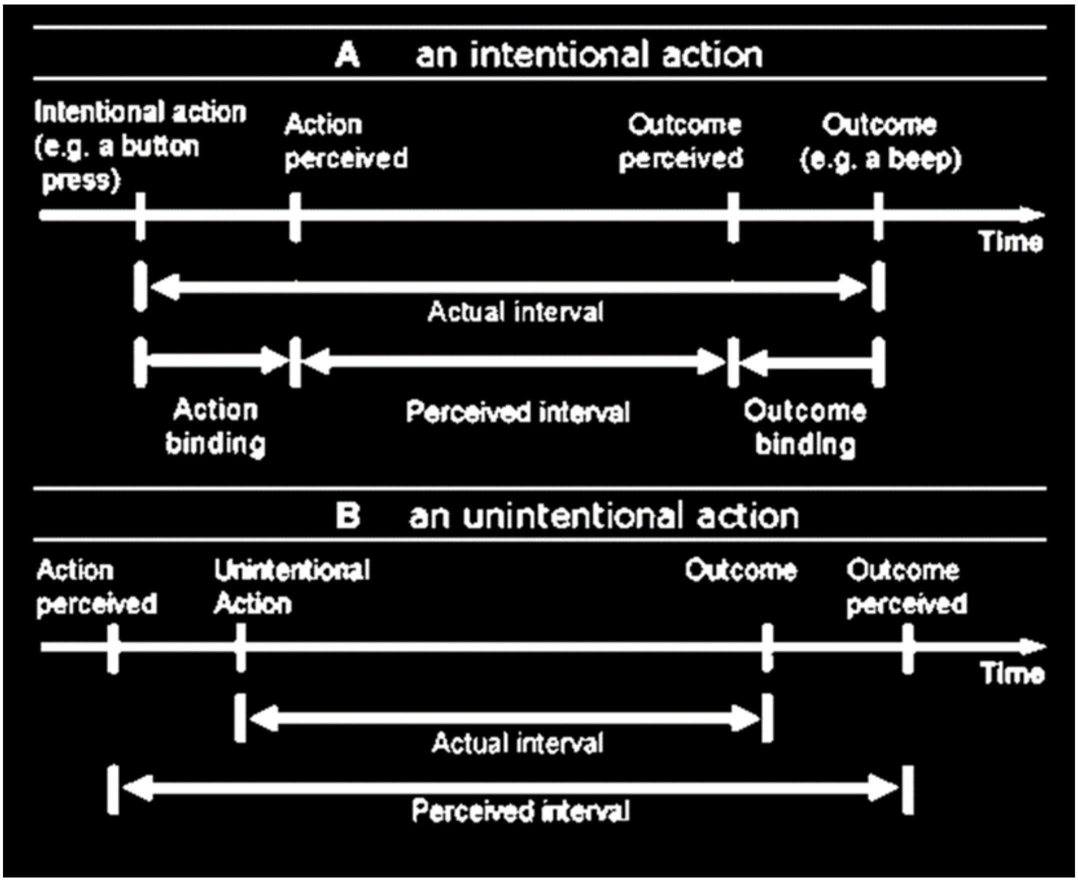

Figure 4. Illustration of the compression or extension of perceived time intervals according to whether an action was made by the participants or not, as first reported by Haggard et al. (2002), figure adapted from Limerick, Coyle, and Moore (2014).

Since Haggard et al.'s (2002) original paradigm, some concerns have been raised about the Libet clock method. For example, it has been argued that the binding effects reported may be an artefact of the Libet clock procedure itself, arising from fundamental subjective biases in judging timings (e.g. Danquah, Farrell, & O'Boyle, 2008), although Pockett and Miller (2007) tested seven possible factors which could challenge the reliability of the procedure and rejected them all. Whilst I do accept the Libet clock method is one reliable way of measuring binding, the method suits paradigms where auditory tones are the action outcomes. It is not as suitable a method when participants must pay careful visual attention to the onscreen stimuli as that would necessitate splitting attention between the visual stimuli and the Libet clock. One alternative method for measuring temporal binding involves participants pressing a space bar on a computer to replicate the time interval that participants have just experienced (e.g. Humphreys & Buehner, 2010; Poonian & Cunnington, 2013). This direct interval reproduction procedure is the method adopted here.

Temporal binding research has repeatedly provided evidence that a compression of perceived time occurs when participants actively cause events by their own actions, but not when passively estimating a time interval between an event they were not the agent of and its outcome (Moore & Obhi, 2012; for a review). Therefore, the theory put forward is that temporal binding is a measure of implicit sense of agency (Moore & Obhi, 2012). Temporal binding has been shown to result when participants press a button to cause an auditory tone (e.g. Humphreys & Buehner, 2009), a coloured disk to collide with another (Cravo, Claessens, & Baldo, 2009), or an onscreen colour change (Ruess, Thomaschke, & Kiesel, 2018). These are in the context of non-social situations, where it is established that action-outcome binding effects can occur. Implicit sense of agency has been investigated, then, for different button press outcomes (David et al., 2008; for a review), but not yet, to my knowledge, for outcomes from eye movements, nor for outcomes from a social interaction.

The experiments in this chapter are, therefore, theoretically novel for three reasons. Firstly, I am not aware that anyone has investigated temporal binding within the context of joint attention before. Secondly, the motor action of a saccade towards an object has not been tested to see if it can produce temporal binding in the same way other motor actions, usually button presses, have been shown to before. Eye saccades are unique in that they do not usually cause a physical consequence in the environment unless there is someone else to respond to the gaze leading within a social interaction, unlike other motor actions with our bodies (e.g. a push or a pull) which usually have a physical consequence upon the (non-social) world around us. Therefore, it is possible that saccades will not elicit the same temporal binding effects. Thirdly, although other studies have examined vicarious agency, that is, agency over another’s actions, this has often been within the context of performing the same action together with a partner (e.g. Obhi & Hall, 2011) or in paradigms aimed to “trick” the actor into perceiving another’s action as their own action (Wegner, Sparrow, & Winerman, 2004). These studies have been informative about when binding can occur when other agents are present but I believe these experiments are the first to examine implicit sense of agency when one person causes an onscreen face to respond to a bid for a social interaction.

Theoretical Models of Sense of Agency and Implicit-Explicit Agency Dissociation

There are two main, theoretically opposing, neurocognitive models proposed for sense of agency. These are the predictive position or the retrospective inference position (reviewed by Moore & Obhi, 2012). According to the predictive model, the sense of agency is produced when there is a match between the predicted and the actual sensory outcome from an action (Blakemore, Wolpert, & Frith, 2002). This idea has been encapsulated in a comparator model for explaining the perception of self-action (see e.g. Blakemore et al., 2002; Wolpert & Flanagan, 2001). This comparator model became a persuasive theory in which feed-forward mechanisms are used to predict the action goal and a feedback loop allows these predictions to be compared against the outcomes (see Frith, 2012, who accepts the model has had considerable empirical support, but advocates a more sophisticated model is needed). The retrospective model, however, conceptualises a comparison between the action’s idea and action’s effect and a sense of agency arises if they are similar (Chambon & Haggard, 2013). The predictive model assumes that sense of agency operates at a low, sensorimotor level, whilst the retrospective inference model conceptualises higher level, cognitive mechanisms may be employed (Barlas & Obhi, 2013).

The predictive model fits well with those who argue that intentionality of the outcome is needed for temporal binding to occur (e.g. Desantis et al., 2012), whilst the retrospective inference model fits best with those who argue that causality is driving the effect (Buehner, 2012). However, a series of studies have together provided evidence that the two models can be integrated, providing supporting data for both causality (retrospective inference) and intentionality (prediction) playing a part in temporal binding (Cravo et al., 2011; see Moore & Obhi, 2012, for a review). Moore, Wegner, and Haggard (2009) argued that different, and varied, agency cues are integrated to result in a sense of agency (e.g. consequences of actions and sensorimotor predictions). This also fits well with the extension of the comparator model offered by Synofzik, Vosgerau, and Newen (2008) and Synofzik, Vosgerau, and Voss (2013) to help reconcile this debate because it allows for both predictive and postdictive mechanisms to operate (see Figure 5).

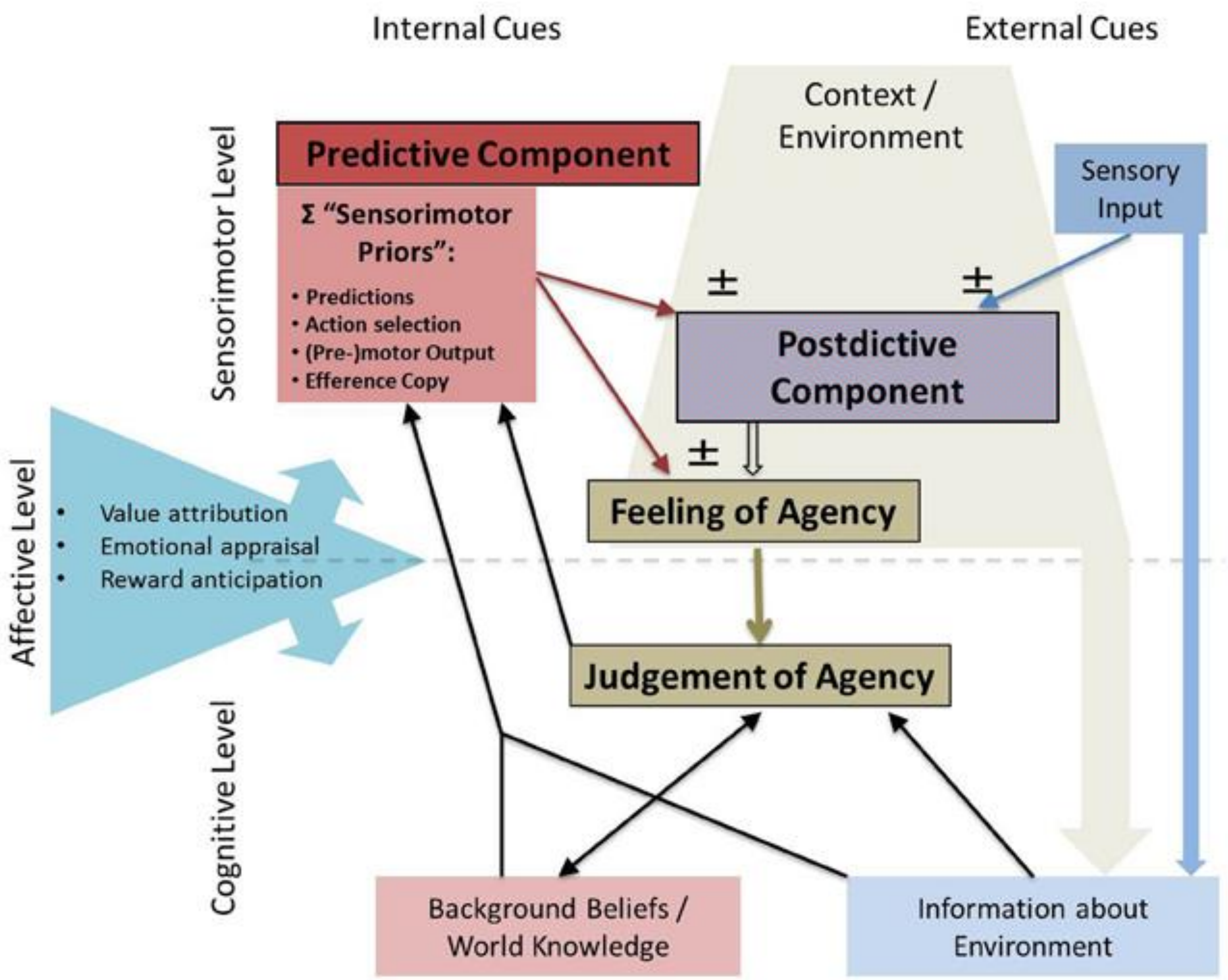

Figure 5. Account of optimal cue integration underlying the experience of agency, taken from Synofzik, Vosgerau, and Voss (2013).

This extension of the comparator model, to incorporate both predictive and postdictive processes and both feelings and judgements of agency, can also help account for why explicit and implicit agency may not always correlate. Ebert & Wegner (2010) provided evidence that changes in temporal binding, and, therefore, implicit agency, were related to changes in explicit self-reports of agency. However, Dewey and Knoblich (2014) reported both explicit and implicit (temporal binding) measures of agency where individual differences for both measures did not correlate. Therefore, Dewey and Knoblich (2014) argued that measures of implicit and explicit sense of agency are not necessarily tapping into the exact same processes. Synofzik et al. (2008) offered a dual conceptual framework where implicit agency operates at a sensorimotor level and can be understood as a "feeling of agency" in which an outcome is classified as self-caused, whilst explicit agency is made after higher level processes make a “judgement of agency.” These two processes can, therefore, be conceptualised as dissociable. Moore, Middleton, Haggard, and Fletcher (2012) tested this by exploring whether explicit and implicit agency were modulated differently by sequential patterns of action and outcome. Their results supported a model in which explicit and implicit agency can be thought of as dissociable, but, they argued, the two are not completely independent systems.

This dual process model was elucidated further into an elegant, optimum cue integration account in which sensorimotor priors form part of the predictive component and environmental information forms part of the postdictive component, together combining to result in how agency is experienced (Figure 5: Synofzik, Vosgerau, & Voss, 2013). Further support for a dissociation between explicit and implicit agency comes from a recent study by Saito, Takahata, Murai, and Takahashi (2015) who found no correlation between explicit measures and temporal binding. In sum, the balance of evidence points towards a dissociation between explicit and implicit agency, but how the two may, nevertheless, relate to one another is not precisely understood. The experiments in this Chapter aim to explore this debate about implicit and explicit agency further by asking participants to rate their feelings of explicit agency, in addition to undertaking the temporal binding tasks.

Neural Mechanisms of Sense of Agency

Attempts have been made to identify the neural mechanisms of sense of agency, although the putative regions are not yet clearly defined. The angular gyrus has been implicated in detecting that there is no agency when temporal congruence is lacking between an action and its outcome (Farrer et al., 2008). This has been partially confirmed by Chambon, Wenke, Fleming, Prinz, and Haggard (2012) who demonstrated angular gyrus activation increased as sense of control decreased when there was no match between prime and target, although no such activation increase occurred for compatible trials (see Figure 6).

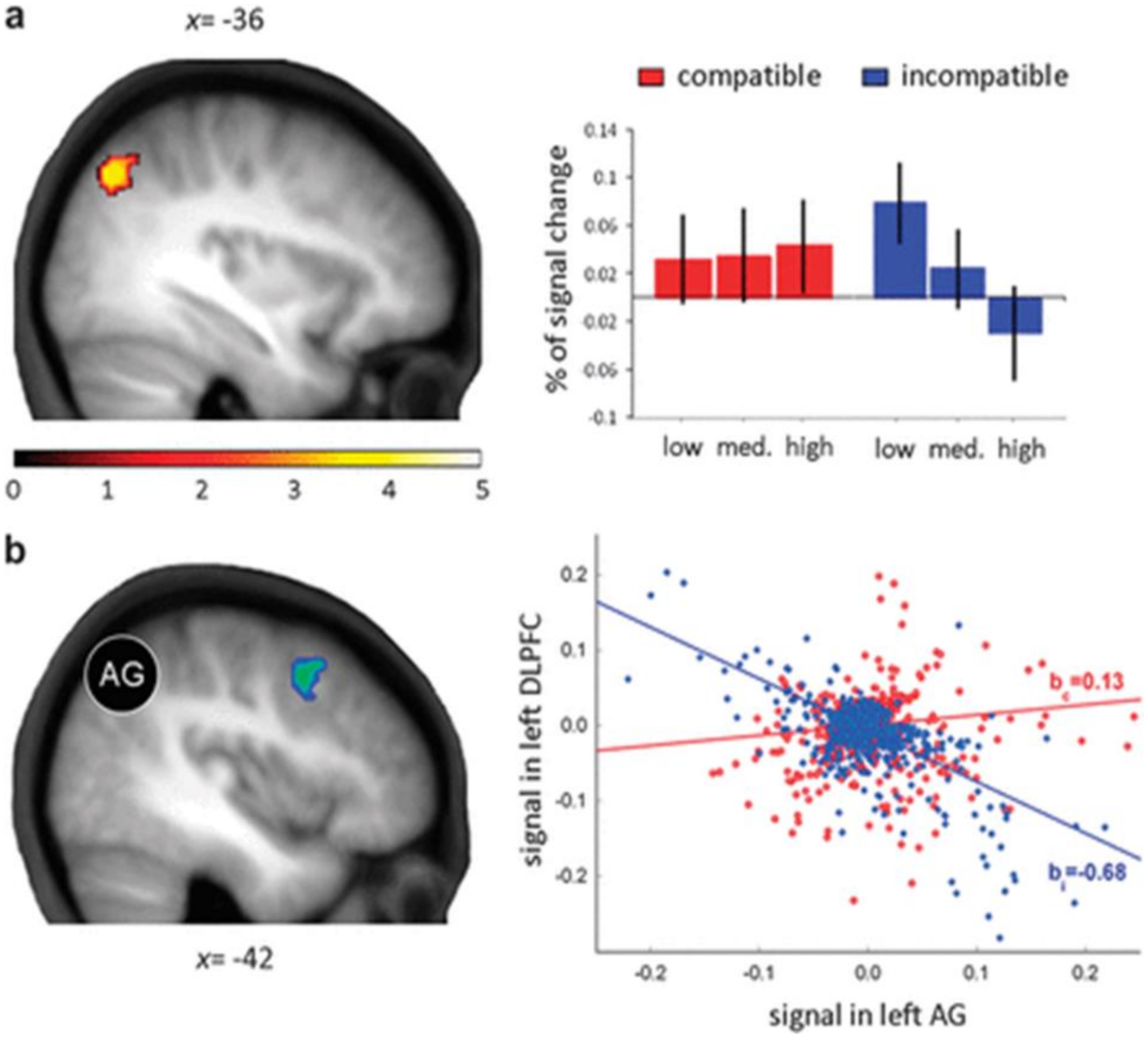

Figure 6. Increased activation of angular gyrus when reported sense of control decreased for incompatible (a mismatch between prime and target), but not compatible (prime and target matched) trials, taken from Chambon et al. (2012). This also shows (bottom right) incompatible trials decreased functional connectivity between the angular gyrus (AG) and left dorsolateral prefrontal cortex (DLPFC).

The dorsolateral prefrontal cortex (DLPFC) appeared to play a role in monitoring the fluency of action selection processes, whilst the angular gyrus detected agency violations (Chambon et al., 2012). This supports the theory that, at a neural level, there is a retrospective process of comparing the predicted outcome with the actual outcome (Chambon et al., 2012). A later study, Beyer, Sidarus, Fleming, and Haggard (2018), replicated these findings that the angular gyrus was modulated by agency, although Kühn, Brass, and Haggard (2013) found no association between angular gyrus activation and implicit sense of agency. Further support for the DLPFC playing a role in sense of agency when selecting between alternative actions has been very recently evidenced by Khalighinejad, Di Costa, and Haggard (2016). Kühn et al. (2013) found the left supplementary motor area activation positively correlated with the degree of temporal binding. A recent review by Merchant & Yarrow (2016) describes how motor actions can alter time perception in several contexts, including temporal binding effects, and identified the supplementary motor area as key to performing actions requiring timing. In sum, to date the regions implicated are the angular gyrus, the DLPFC and the left supplementary motor area but more research is needed for a complete picture to emerge. Haggard (2017) offers a further review of the neural mechanisms involved in sense of agency. Having reviewed the behavioural consequences and neural mechanisms of sense of agency, I now turn to the specific motor action of eye movements (saccades) used in gaze leading during joint attention and how sense of agency may be experienced when influencing another's direction of gaze.

Eye Gaze and Social Attention

Eye stimuli are considered to have special status within social attention because the morphology of the human eye allows easy detection of gaze signals (Kobayashi & Kohshima, 2001) and neural mechanisms have been developed to process gaze information (Langton et al., 2000). This is also consistent with the cooperative eye hypothesis which is that human eyes have evolved to be highly visible because of the need for enhanced social interactions (Tomasello et al., 2007). Eye gaze is used in everyday interaction for social evaluation and to detect intention to communicate (Hoehl et al., 2008). This gaze information is processed automatically and rapidly to help us understand the intentions of others, forming a key contributor to social cognition (see Emery, 2000; Frischen et al., 2007; Langton et al., 2000, for reviews).

Joint Attention

"Joint attention," during which there is a gaze leader and a gaze follower, is when eye gaze communication is shared between two people about an object (Emery, 2000). Joint attention plays a key role during development, particularly in the progression from the primary to the secondary "inter-subjective" phases of normal development (Mundy & Newell, 2007). "Inter-subjectivity" is the phenomenon of sharing experiences with one another (Bard, 2009; Trevarthen & Aitken, 2001). The mental processes during an infant's first year of life are termed the "primary intersubjective" phase during which a mother and child share attention in a dyadic fashion, that is, between each other (Bruner & Sherwood, 1976; Terrace, 2013). After one year, this phase progresses to the "secondary intersubjective" phase during which their attention becomes shared in a triadic fashion, that is, attention includes objects in the environment; this is the development of joint attention (Terrace, 2013). Mother-child joint attention is positively correlated with efficiency in word learning (Tomasello & Farrar, 1986) and the frequency of engaging in joint attention predicts language acquisition ((Morales et al., 2000; Mundy & Newell, 2007). In order to possess inter-subjectivity, the child must have a sense of self which must include a sense of agency over its environment. An implicit sense of self in infants from at least two months onwards is believed to develop before explicit manifestations of self-agency emerge in the second year of life (Rochat & Striano, 2000). This suggests a strong developmental link between joint attention and agency and so this chapter seeks to explore the relationship between them.

From an evolutionary perspective, agency and joint attention are of interest because sense of self may be unique to humans, although this is hotly debated (Terrace, 2013, see also a review byAnderson & Gallup, 2015). It is also argued that joint attention may be exclusively human (Call & Tomasello, 2005; Carpenter & Call, 2013; but see Leavens & Racine, 2009, for an alternative view). Therefore, the processes of agency and joint attention may potentially have co-evolved, whether or not other primates may share some capacity for sense of self or joint attention. Indeed, this accords with a theoretical stance from developmental psychology termed “natural pedagogy.” This hypothesises an evolutionary link between cognitive mechanisms for cultural learning (which I argue includes possessing a sense of agency) and the adaptive ability to communicate, in which joint attention plays a key part (see e.g. Csibra, 2010; Gergely, 2013).

Eye gaze research has tended to focus on the effects upon responders who follow gaze. Gaze cueing paradigms have repeatedly shown that when presented with a target, reaction times by the responder are faster to that target if preceded by a gaze cueing signal (Frischen et al., 2007). Gaze following has been shown to be an automatic, robust phenomena (Frischen et al., 2007, for a review), develops early in infancy (Farroni et al., 2004) and may be innate, although what ‘innate’ means can be debated (Brooks & Meltzoff, 2015). Some revealing studies have investigated brain activation associated with joint attention initiation (e.g. Schilbach et al., 2010), but there is much less research into the behavioural effects upon the initiator, compared to the responder. One study which did focus on the initiator, rather than the responder, found that participants spontaneously made faster saccades back to faces which had engaged in joint attention under non-speeded conditions (Bayliss et al., 2013). This process of initiating joint attention was termed “gaze leading.” These findings raised questions about what effects would occur under speeded conditions. This was investigated by

Edwards et al. (2015) who measured participant response times to targets presented on faces which did or did not follow eye gaze during a computer-simulated, eye-tracked joint attention interaction. Edwards et al.'s (2015) main finding was that faces which follow our eye gaze capture our attention because response times were faster to those faces than other faces. This led to the theoretical suggestion that this effect of gaze leading evidences an evolved mechanism to facilitate shared attention and the on-going social interaction. The aims of the experiments in this chapter are to explore this attentional capture effect further by investigating whether a sense of agency over the responder is experienced by the initiator.

One study has already suggested that having a successful response to a joint attention bid creates an explicit sense of agency in the initiator. Pfeiffer et al. (2012) collected self-reports of feeling of relatedness between the participant's gaze towards an object and the response of either gaze aversion or gaze following. The results showed that greater sense of agency (relatedness) was reported when gaze was followed than when it was not followed (see Figure 7).

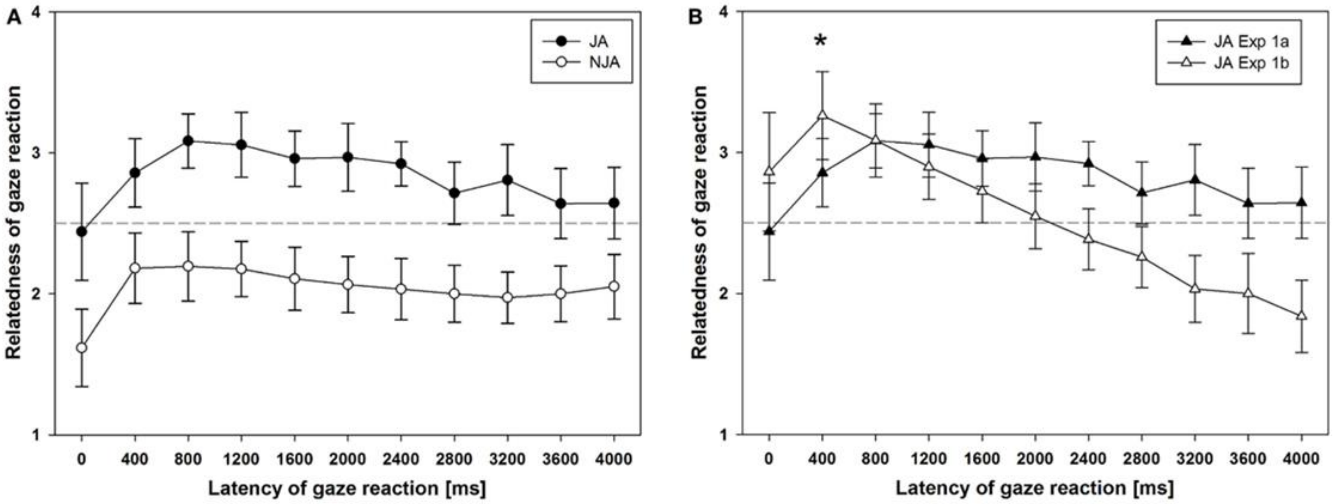

Figure 7. Graphs taken from Pfeiffer et al. (2012) showing self-reported relatedness of gaze reaction where gaze was always followed (Joint attention; JA) and when gaze was either followed (JA) or not followed (non-joint attention; NJA). Panel A shows self-reported relatedness of gaze reaction for a study where the response was equally likely to be JA or NJA. This shows little effect of latency on relatedness. Panel B shows another study where gaze was always followed to establish JA (showing greatest relatedness at 400 ms and a linear decrease thereafter) and also shows again the JA condition from Panel A which shows little effect of latency.

The neural mechanisms of joint attention have been the subject of several studies with both distinct and common regions shown to be recruited when initiating or responding to joint attention (Redcay et al., 2012). The main regions are the medial Prefrontal Cortex (mPFC; associated with social interaction) and the posterior Superior Temporal Sulcus (pSTS; associated with shifts of attention). Both regions are recruited during joint attention (Redcay & Saxe, 2013). In addition, Schilbach et al. (2010) and Gordon et al. (2013) have shown that the striatum may be involved in having joint attention bids responded to, suggesting this is a rewarding experience. Similarly, experiencing a sense of agency for a positive social outcome has been shown to be associated with increased ventral striatum activity (Decety & Porges, 2011).

Distortions of Perceived Time Associated with Eye Movements

Two phenomena which affect time perception have already been found to be related specifically to eye movements (see Merchant & Yarrow, 2016, for a review). One is saccadic suppression, whereby space and time are not perceived during a saccade. This has been demonstrated by Morrone, Ross, and Burr (2005) who found subjective time intervals were shortened by a factor of two during a saccade. Another process, believed perhaps to work in opposition to saccadic suppression, is chronostasis. This is the phenomenon whereby stimuli are perceived to have been presented for longer than actually presented following a saccade (Yarrow, Haggard, Heal, Brown, & Rothwell, 2001). It is persuasive that saccadic suppression would have an opposite process such as chronostasis to compensate for the loss of perceived time during saccades (Merchant & Yarrow, 2016). This was tested by Knöll, Morrone, and Bremmer (2013), across four experiments, who showed that chronostasis does not exclusively occur at saccadic targets and may be a more general mechanism during visual perception. Nevertheless, both processes of saccadic suppression and chronostasis are distortions of time found in eye vision research and it is interesting to explore whether the further time distortion of temporal binding will occur in the paradigm introduced in this chapter.

ASC and Sense of Agency

ASCs are heterogeneous disorders, defined by the DSM-5 (American Psychiatric Association, 2013) as persisting deficits in social communication and social interaction across multiple contexts, alongside restricted, repetitive patterns, interests, or activities. Two studies to my knowledge have investigated sense of agency for those with an ASC with conflicting results. David et al. (2008) found no sense of agency deficits, but just investigated self-reported explicit sense of agency. Sperduti, Pieron, Leboyer, and Zalla (2014), exploring implicit sense of agency, did find reduced temporal binding effects for those with an ASC compared to the typically developing population. Therefore, the limited empirical evidence appears contradictory but may relate to the dissociation between explicit and implicit agency. A self-report questionnaire of level of autism-like traits (the Autism-Spectrum Quotient Questionnaire; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) was used in the experiments in this chapter alongside the temporal binding experiments to enable exploration of any association between autistic-like traits and temporal binding effects in the typically developing population.

Further studies which examined time reproduction abilities, rather than agency, for those with an ASC found intact reproduction abilities compared to controls, suggesting time perception is not impaired, and may, in fact, be more accurate (Stewart, Griffiths, & Grube, 2015; Wallace & Happé, 2008). It is not clear whether this would be found to be the case in studies like those presented in this chapter where joint attention is also a feature of the paradigm because there are deficits associated with those with ASCs for joint attention, which I now turn to.

ASC and Joint Attention

Importantly, interventions to improve joint attention in those diagnosed with an ASC have been found to significantly improve language development, play skills and social development (Goods, Ishijima, Chang, & Kasari, 2013; Reichow & Volkmar, 2010, for a review). Therefore, any insight into links between autism-like traits and joint attention initiation could prove helpful to inform interventions like these. Edwards et al. (2015), across three experiments, found a negative correlation between the gaze leading effect (attentional orienting towards faces which had just followed gaze) and level of autism-like traits. Therefore, the higher the autism-like traits, the less faces who followed gaze captured attention. Therefore, in these experiments, I explore whether a similar correlation might be found between magnitude of temporal binding and autism-like traits when joint attention initiation is made.

To summarise the predicted results, I hypothesised that having participants' self-initiated eye movements result in an outcome would produce temporal binding, just as other motor actions have been shown to. The temporal binding effect was measured in milliseconds (ms) and was calculated by the difference between the actual time interval and participants' replicated reproduction of the interval. In addition, the percentage reproduction was calculated to compare participants' time reproductions with a perfect reproduction of one hundred per cent. I also expected that there would be greater temporal binding effects when participants' eye gaze was followed to an object (Active Gaze Leading conditions) than when no saccades to the object were made (Passive conditions). If this hypothesis is supported, this may evidence an implicit sense of agency for the initiator in causing the gaze following outcome in the responder.

I predicted there would be a negative correlation between autism-like traits and temporal binding effects in the Active Gaze Leading condition. Therefore, the higher the level of participant's autism-like traits, the lower the magnitude of the temporal binding effects when participant's gaze is followed to an object. I maintained an open hypothesis about whether the magnitude of temporal binding effects across conditions would correlate with level of autism-like traits because the evidence for reduced sense of agency for those with an ASC is limited. Finally, explicit ratings of agency are predicted to be greatest in the Active Gaze Leading tasks over the Passive tasks with an open hypothesis about whether explicit and implicit agency would correlate.

Experiment 1

This study was designed to test whether gaze leading (using a horizontal saccade) towards an object to engage in joint attention would result in a sense of agency over gaze shift responses. This was tested using a temporal binding paradigm in which participants replicated the time interval between an object appearing (immediately after which they saccaded to it) and an on-screen face looking towards the object. It was hypothesised that having gaze leading followed would result in temporal binding, evidencing a sense of agency over the on-screen face's gaze shifts. Conversely, during passive tasks in which participants made no saccades towards the object, no temporal binding was predicted to occur. There were two passive tasks; one with a face, identical to the Active Gaze Leading task, and one control task using a phase scrambled version of the face, with a spatial shift towards the object in place of the gaze shift. In addition, correlations between the magnitude of any binding effects and level of autism-like traits and between implicit and explicit agency were examined. Throughout all three experiments reported here, Active Gaze Leading tasks involve gaze leading towards an object, whilst Passive tasks involve no active eye movements. Passive Face tasks require participants to look at the face throughout, whilst Passive Object tasks require participants to look at the object throughout.

Experiment 1 Method

Participants. After ethical approval was obtained from the University of East Anglia Psychology Ethics Committee, 32 Psychology undergraduate student participants (mean age = 20.59 years, *SD* = 6.40; 2 males), gave written, informed consent and were granted course credits in return for participation. The sample size was guided by previous eye tracked experiments (Edwards et al., 2015) which tested 32 participants. Also, a power analysis with a medium effect size and a power of 0.80 (Cohen, 1988) calculated using G*Power3, (Faul, Erdfelder, Lang, & Buchner, 2007) produced a required sample of 28. All participants reported normal or corrected-to-normal vision. Participants were positioned comfortably in a chin rest and had rest breaks between three blocks of trials.

Design. The within-subjects design had three blocked conditions of 56 trials each. The conditions were Active Gaze Leading, Passive Face and Passive Phase Scrambled. The Active Gaze Leading condition is illustrated at Figure 8. The dependent variable was the participant’s percentage reproduction of the inter-event interval. The inter-event interval was the time between an object’s appearance and a subsequent gaze shift (Active Gaze Leading and Passive conditions) or a spatial shift (Passive Phase Scrambled condition) towards the object. There were also two correlational designs to examine any associations between, firstly, level of autism-like traits and, secondly, explicit ratings of control and degree of temporal binding.

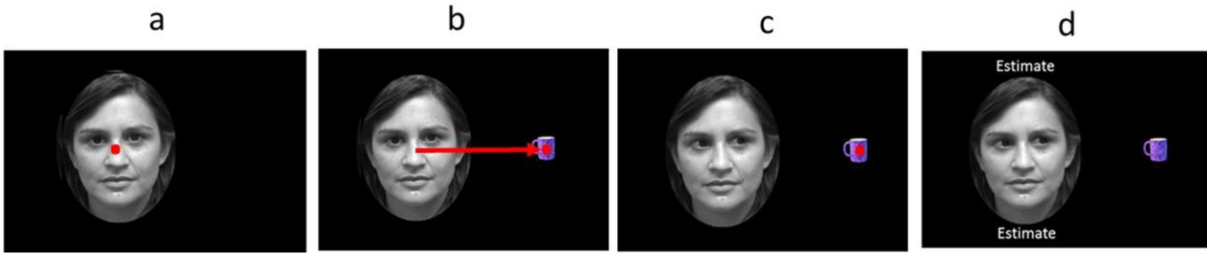

Figure 8. Trial sequence for the Active Gaze Leading task. Circles and the arrow were not displayed but represent where participants were instructed to fixate and the saccade from the face to the object, respectively. Participants looked at the face (a), displayed for 1000ms. Participants made a saccade (b) to the object as soon as it appeared. After a random inter-event interval of 400ms to 2300ms, gaze onset (c) occurred. After 1000ms, estimate instruction appeared (d) until response. Participants pressed and released the space bar to replicate the inter-event interval. The inter-event interval is the time between the object appearing and the gaze onset.

Stimuli. The female face stimulus was a grayscale photograph with a neutral expression (280 × 374 pixels) taken from Bayliss, Bartlett, Naughtin, and Kritikos (2011). The photographs of the face included three variations: eyes open looking straight ahead, eyes closed and looking right (25°). The object stimuli set consisted of 8 objects commonly found in the kitchen (220 × 78 pixels), taken from Bayliss, Paul, Cannon, and Tipper (2006). The face was presented on the left hand side of the display with the centre of the nose 13.5 cm from the left hand side of the display and 7.5cm from the top of the display. The objects were presented to the right of the face stimuli with the centre of the object 11.5cm from the centre of the nose on the face stimulus. There was also a phase scrambled version of the face stimulus, used in the Passive Phase Scrambled task, comprising a rectangle (280 x 374 pixels) with two smaller rectangles (37 x 26 pixels) placed where the eyes would be on the face. The smaller rectangles were phase scrambled images of the face stimulus eye regions. All stimuli appeared on a black background and were presented using E-Prime 2.0 software. Examples of stimuli are illustrated in Figure 8.

Apparatus and materials. Right eye position was tracked with an infrared eye tracker (Eyelink 1000, SR Research, Ontario, Canada; spatial resolution of 0.1^0, 500 Hz). A head and chin rest was used to maintain head stability. Viewing distance was 70cm from eyes to an 18” computer monitor (resolution 1024 × 768 pixels). A standard keyboard was used for participants’ manual responses. The Autism Spectrum Quotient Questionnaire was used as a measure of the level of autism-like traits (AQ; Baron-Cohen et al., 2001), presented using E Prime. Participants rated how strongly they agreed or disagreed with each item (e.g. I prefer to do things the same way over and over again) on a four point Likert scale ranging from definitely agree to definitely disagree. To measure self-reported explicit agency, participants rated, after each task, how much control they felt they had over the onscreen face’s eye movements or the rectangles shifting on a scale from 1 to 8 (with 1 representing no control at all and 8 representing a lot of control).

Procedure. The eye tracker was calibrated to each participant to enhance tracking accuracy using a standard nine point calibration at the start of each block. The participants completed a practice block of six trials followed by three blocks of 56 trials, one for each task. The three tasks, which were counterbalanced, are termed Active Gaze Leading, Passive Face and Passive Phase Scrambled. Example trial sequences are illustrated at Figure 8.

In the Active Gaze Leading task, each trial began with the presentation of a female face with direct gaze on the left side of the screen. Participants were instructed to look at the face (presented for 1000 ms) until one of eight objects appeared to the right of the face which became the participant’s cue to rapidly shift their gaze to reorient to it. The object was displayed for a random inter-event interval of 400-2300 ms, after which the face’s gaze shifted 25° to the right to look at the object. This inter-event interval was the time interval the participants were asked to replicate, with the start of the time interval being when the object appeared and the end of the interval being when the gaze shift occurred. The random sampling of the time interval meant the time interval for any one trial could be anywhere between 400 and 2300 ms to the millisecond, as is usual in the temporal binding literature.

Next, after 1000 ms, the participant was prompted by the on-screen instruction "Estimate" presented on a black background with a white font (Courier, 18pt) above and below the face to press and hold down the spacebar with their preferred hand to replicate the time interval between the object's appearance and the face's gaze shift towards it. Participants were given no feedback about their responses.

An inter trial interval (a blank screen) was displayed for 1000 ms. The stimulus set of objects for all trials was presented randomly. Participants were told they must fixate on the object as soon as it appeared to trigger the face to follow their gaze. They were given no further instructions about fixation during each trial beyond fixating the object. Participants did not know anything about the hypotheses.

The Passive Face task display was identical to that of the Active Gaze Leading task, except that the onscreen face had closed eyes (instead of direct gaze) until the gaze shift to the right occurred. The Passive Phase Scrambled task had the same display as the other two tasks except that, in place of the face, was a rectangle comprised of the phase scrambled face, with two smaller, phase scrambled rectangular regions, which were used for a spatial shift towards the object, instead of a gaze shift. The phase scrambled rectangles, positioned in the place the eyes would have been, shifted 2 mm to the right after the inter-event interval. The size of the 2 mm spatial shift was chosen as this matched the same spatial shift as occurred for the pupils of the eyes when gaze shifted in the other two tasks.

In the Passive Face task participants were instructed to fixate on the face throughout each trial, not to look at the object and that they needed to replicate the time interval but they were not causing the gaze shift to occur. The time interval participants were asked to reproduce was the same as the Active Gaze Leading task. In the Passive Phase Scrambled task the same instructions were given as the Passive Face task, except participants fixated the rectangle throughout and the inter-event interval was the time between the object's appearance and the shift of the smaller rectangles to the right. After each task participants rated their self-reported degree of control felt over the face's eye movements or the rectangles shifting. After all three tasks, participants completed the Autism Quotient Questionnaire on the computer.

Experiment 1 Results

Reproduction error and percentage reproduction. Mean reproduction error was calculated for each participant in each condition which is the reproduced time interval minus the actual time interval (see Table 2). Trials in which participants' estimates were 3*SD*s above or below their individual means were removed (0.41% of trials). The reproduced time intervals were divided by the actual time intervals to calculate mean percentage reproduction (see Figure 9). Therefore, 100% reproduction represents perfect accuracy, anything greater than 100% is over-reproduction, and less than 100% is temporal compression (under-reproduction). Greenhaus-Geisser corrected degrees of freedom are reported when applicable. Confidence intervals and standard errors around the means are based on 1000 bootstrap samples. Confidence intervals around effect sizes have been calculated using ESCI (Exploratory Software for Confidence Intervals; Cumming & Calin-Jageman, 2016).

Table 2

Mean and Standard Deviations (in parentheses) for Reproduction Errors (ms) and Explicit Agency Ratings (from 1 to 8) in all conditions and the Agency Rating Difference.

	RE									Agency Ratings Difference			
		A Direct Gaze	A Closed Eyes	PF Direct Gaze	PF Closed Eyes	PO Direct Gaze	PO Closed Eyes	SR	SO	A and PF	A and PO	A and SR	A and SO
E1	RE	-211.53 (436.28)	-	-	2.34 (399.20)	-	-	-77.85 (398.60)	-	-213.87 (339.79)	-	-133.69 (253.42)	-
	Explicit Rating	4.44 (2.09)	-	-	2.25 (1.61)	-	-	2.03 (1.43)	-	2.19 (2.02)	-	2.41 (1.95)	-
E2	RE	-	-280.49 (263.94)	-58.32, (305.37)	-	-136.91 (289.20)	-	-	-	-222.17 (206.74)	-143.57 (195.32)	-	-
	Explicit Rating	-	3.97 (1.79)	2.59 (1.50)	-	2.72 (1.75)	-	-	-	1.38 (2.12)	1.25 (1.93)	-	-
E3	RE	-248.59 (296.61)	-	-	-	-	-137.20 (300.89)	-	-207.98 (345.57)	-	-111.39 (211.98)	-	-40.61 (220.81)
	Explicit Rating	4.06 (2.00)	-	-	-	-	2.16 (1.57)	-	2.00 (1.72)	-	1.91 (2.08)	-	2.06 (2.31)

Abbreviations: RE = Reproduction Error in milliseconds (ms)

Notes: the Agency Rating Difference is calculated by subtracting the mean RE or explicit rating for the Passive task from the RE or explicit rating for the Active Gaze Leading task. The conditions are: A is Active Gaze Leading (a gaze leading saccade was made from face to object), PF is Passive Face (fixation was on the face throughout), PO is Passive Object (fixation was on the object throughout), SR is Phase Scrambled Rectangle (fixation was on the scrambled rectangle throughout) and SO is Phase Scrambled Object (fixation was on the object throughout).

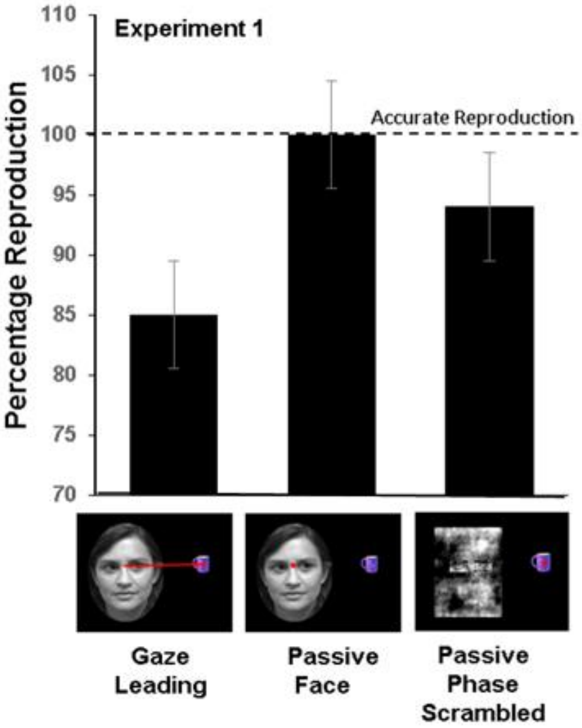

Figure 9. Mean percentage reproductions for Experiment 1. In the Gaze Leading task, participants looked first at the face, and then at an object as soon as it appeared. In the Passive Face or Passive Phase Scrambled tasks, participants looked at the face or scrambled face throughout. The images show how the face/scrambled stimulus was displayed when gaze onset occurred. Circles and the arrow were not displayed but represent where participants were instructed to fixate (and the saccade from the face to the object for the Active task). Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by (Loftus & Masson, 1994).

First, in order to establish whether each condition produced temporal compression (reliable under-reproductions of the time between object and gaze onset), or relatively accurate reproductions, single sample *t*-tests were performed for each of the three conditions using percentage reproduction. This showed that temporal compression was only statistically significant in the Active Gaze Leading condition. Here, participants reproduced *M*=84% of the veridical time interval, 95% CI [73, 96] (*SD*=32%), *t*(31) =2.76, *p*=.01, d_z=0.69, 95% CI [0.18, 1.19]. In the two passive conditions, reproduction errors (REs) were low and did not differ statistically from 100% reproduction (Passive Face Fixation condition: *M*=100% reproduction, 95% CI [91, 112], *SD*=30%, *t*(31)=0.09, *p*=.926, d_z=0.02, 95% CI [-0.51,0.47]; Passive Phase-scrambled, *M*=94% reproduction, 95% CI [82, 100], *SD*=30%, *t*(31)=1.09,

p=.286, d_z=0.27, 95% CI [-0.22, 0.76]. There was a main effect of task, F(1.53,47.42)=10.91, MSE=207, p<.001, η_p^2=0.260, and follow-up contrasts showed that the proportional temporal compression effect in the Active Gaze Leading condition was greater than in both the Passive Face Fixation, t(31)=3.73, p=.001, d_z=0.52, 95% CI [0.21,0.82] and Passive Phase Scrambled Fixation conditions t(31)=3.17, p=.003, d_z=0.32, 95% CI [0.10,0.52]. Therefore, the hypothesis that having participants' deliberately-initiated saccade *followed* would result in greater temporal compression than passive conditions (where no saccades were made) was supported.

Secondary measures, manipulation checks and participant subset analyses. Mean self-reported explicit ratings of agency were greater for the Active Gaze Leading (*M*=4.44, *SD*=2.09) 95% CI [3.68, 5.19], than both the Passive Face Fixation (*M*=2.25, *SD*=1.61) 95% CI [1.67, 2.83] and Passive Phase Scrambled Fixation (2.03, *SD*=1.43) 95% CI [1.52, 2.55] conditions; *t*s>6, *p*s<.001, d_zs>1**.** This shows that participants felt a degree of explicit agency in the Gaze Leading condition, supporting the inference that the temporal binding effect presented here reflects a sense of agency.

I considered potential concerns that something about performing a saccade per se might explain the data. Saccades can, indeed, affect time perception; a substantial amount of work has demonstrated an expansive effect (chronostasis; see review by Merchant & Yarrow, 2016), which if present in these data would of course increase participants' estimates (i.e. this effect, if present, would work in opposition to the predicted and demonstrated effects). However, two studies have noted an opposing compressive effect (Morrone et al., 2005; Yabe, Dave, & Goodale, 2017). These opposing effects are small and might be of similar magnitude, potentially cancelling each other out were they to be present in this (rather different) task. Nevertheless, it is still possible that any temporal compression effects of making saccades alone may account for an element of the temporal binding effects shown here so it is prudent to explore the data for any relationships between saccade metrics and time reproductions. It was found that in the critical Active Gaze Leading condition, neither mean saccadic reaction time (sRT, *M*=220ms, *SD*=41ms) nor saccade duration (*M*=81ms, *SD*=44ms) correlated with temporal binding, *r*=-.05, *n*=32, *p*=.80, *r*=.004, *n*=32, *p*=.98, respectively.

Further exploration was carried out to check for those participants whose saccades may have landed on the object *after* the onscreen face's gaze shift as this may have affected their perception of their sense of agency over that outcome. This was possible in this design if participants were particularly slow to execute a saccade on a trial with one of the shortest possible temporal intervals (here, the shortest possible interval was 400 ms). Such occurrences were present in nine participants, and on a maximum of three trials for a given participant (and a total of 0.7% of active trials). The explicit and implicit data were reanalysed, excluding all nine of these participants, and it was found that the data pattern was very similar without them. Their mean explicit ratings are similar to those who never experienced a gaze shift before their saccade landed (*M*=4.5, *SD*=2.22 and M = 4.41, *SD*=2.15, respectively). Temporal compression was only statistically significant in the Active Gaze Leading condition. Here, participants reproduced *M*=84%, 95% CI [74,95] (*SD*=30%), of the veridical time interval *t*(22)=2.49, *p*=.02, d_z=0.73, 95% CI [0.13,1.3]. In the two passive conditions, reproduction errors were low and did not differ statistically from 100% reproduction, Passive Face Fixation condition: *M*=103%, 95% CI [93,113] *SD*=23%, *t*(22)=0.597, *p*=.556, d_z=0.18, 95% CI [-0.75,0.40]; Passive Phase-scrambled, *M*=98%, 95% CI [87,109], *SD*=25%, *t*(22)=0.31, *p*=.763, d_z= 0.09, 95% CI [-0.49,0.67].

To check whether passive tasks were compromised by saccades occurring contrary to the fixation instruction, erroneous saccades were also examined; on only 0.28% of trials were saccades made in error to the object during the Passive Face task and on 0.11% of trials in the Passive Scrambled condition. These few trials are unlikely to have had a critical impact on the data. Thus, overall, saccade metrics cannot parsimoniously explain the observed time underestimation in the Active task at the trial or participant levels.

As this is the first attempt to my knowledge using a temporal binding paradigm with saccades as the action, it is useful to examine whether the data share another commonality often observed in manual tasks in order to inform comparability across effectors. Previous temporal binding research using interval replication or estimation methodologies show stronger effects with longer intervals (Humphreys & Buehner, 2009; Wen, Yamashita, & Asama, 2015). In order to determine whether the data share this latter characteristic of the temporal binding phenomenon, performance of each participant on the longer 50% of intervals they estimated was compared with the shorter 50% of intervals they estimated. In order to establish whether this pattern is present in the data, the *reproduction error* (RE) is used as the measure, calculated in milliseconds as the participants' reproduction of the temporal interval between two events minus the veridical temporal interval (rather than the percentage reproduction used in the main analysis). The split was applied by taking all the trials with intervals below the median for the shortest half, and all those above for the longest half. The median time interval was 1350 ms and the mean was 1341 ms (*SD*= 550). The mean temporal interval for the shortest half was 865 ms (*SD*= 280) with a mean RE of 394 ms (*SD*= 705), meaning there was no under-reproduction for the shortest half. The mean temporal interval for the longest half was 1815 ms (*SD*= 548) with a mean RE of -819 ms (*SD*= 548), meaning there was under-reproduction in the longest half. The shortest and longest halves were significantly different, $t(31)=10.27$, $p<.001$, $d_z=1.75$. This corroborates the notion that the observed data reflects a temporal binding effect, rather than some form of previously unreported saccade-induced temporal discounting effect that would most likely be either proportional to saccade metrics, or in fact be stronger for short intervals, not weaker (given the timescale of saccades, and the timescale of previously observed interactions between saccades and time perception).

Correlation analysis of explicit and implicit agency findings. Any correlations between explicit and implicit agency were examined for each of the three conditions and found no significant correlations. There was a statistical trend towards a negative correlation in the Active Gaze Leading condition; this showed a relationship (but not significant), so that the higher the self-reported explicit agency, the higher the implicit agency measured by binding effects, $r(32) = -0.34$, $p = .054$ (two-tailed).

There was a medium, negative correlation between the differences between Active Gaze Leading and Passive Face. Greater implicit or explicit agency in the Active Gaze Leading task over the Passive task is termed here the "Agency Effect." Therefore, the greater the Agency Effect for the implicit task measured by temporal binding, the greater the explicit Agency Effect, $r(32) = -0.38$, $p = .03$ (two-tailed). Similarly, there was a medium, negative correlation between the Agency Effect for the Active Gaze Leading task and Passive Scrambled. Again, the greater the implicit Agency Effect, the greater the explicit Agency Effect, $r(32) = -0.41$, $p = .02$ (two-tailed).

Correlation Analysis of AQ. The mean AQ score was 16.59 (*SD* = 5.58) and the range was 3 to 26. This compares well with a meta-analysis mean of 16.94 from 73 studies in non-clinical populations (Ruzich et al., 2015) and a mean of 16 obtained by the authors of the AQ from a social science population, as here (Baron-Cohen et al., 2001). The hypothesis was that a higher level of autism-like traits would be associated with less of a difference in binding in the Active compared to Passive conditions. Therefore, a Pearson's correlation analyses was performed on AQ score and Agency Effects for Passive Face task, *r*(32) = -0.13, *p* =.469 (two-tailed) and Passive Scrambled task, *r*(32) = -0.14, *p* = .445 (two-tailed). The hypothesis was not supported. There was also no correlation between AQ score and binding effects collapsed across all conditions, *r*(32) = 0.05, *p* = .787 (two-tailed).

Experiment 1 Discussion

The results demonstrate that having a voluntary saccade towards an object reciprocated results in the classic temporal binding effect which suggests an implicit sense of agency over the onscreen face's response during a joint attention interaction. This contrasts with no temporal binding in the Passive Face and Passive Phase Scrambled tasks. Therefore, it seems to be the action of making a saccade which drives the effect in this experiment and produces the implicit sense of agency. This supports the growing body of temporal binding research which has demonstrated the same effects but for other motor actions, usually button presses, and within a non-social context (see Moore & Obhi, 2012, for a review). The significantly greater binding in the Passive Phase Scrambled task, compared to the Passive Face task, is surprising, although it may be due to the closed eyes in the Passive Face task at the outset offering the social cue that a gaze shift is unlikely to be triggered by the participant, whilst in the Passive Scrambled task, no such social cue was available.

After Experiment 3, I will discuss further the potential effects of low level spatial shifts upon implicit agency. However, for now, the null binding effects in both Passive tasks compared to the Active Gaze Leading task do support the hypothesis that it is the action of gaze leading which drives the binding effects. The explicit ratings of agency show the manipulation of explicit sense of agency was successful with participants rating more sense of control over onscreen events in the Active Gaze Leading task than both the Passive tasks. There was a correlation between the Agency Effects (more binding for Active Gaze Leading over Passive tasks) between implicit and explicit measures. However, what this could mean for the relationship between explicit and implicit agency will be discussed in the Chapter Discussion because the results here do not replicate in the next two experiments or when combining the data for all three experiments.

There were no correlations between AQ and magnitude of binding and magnitude of explicit agency. The lack of correlation in a non-clinical sample between AQ and explicit binding support previous findings in an ASC sample of no difference in explicit agency to typically developing individuals, suggesting this may possibly be intact in those with an ASC (David et al., 2008). However, the findings of no correlation between AQ and magnitude of temporal binding contrast with those of Sperduti et al. (2014) who found less binding in an ASC sample than the typically developing sample. It may be that this study failed to find any correlation because the effect is only found in those with a clinical diagnosis. Future research is needed to test further whether implicit agency over gaze responses is intact for those with an ASC diagnosis.

Experiment 2

Experiment 2 Method

A new group of participants of the same number (N=32) and type as Experiment 1 was recruited in the same manner and, again, took part in return for course credits. Their mean age was 19.69 years (SD = 2.42, four males). The findings in the Phase Scrambled conditions in Experiments 1 had been interesting. However, the most crucial effect to seek to replicate was the novel finding, that an intentional saccade towards an object, if followed, compresses our perception of time between the object appearing and the subsequent gaze shift. In Experiment 1 the Active Gaze Leading condition had the onscreen joint attention responder maintain direct gaze towards the participant at the outset. Therefore, it was prudent to check whether a lack of eye contact between the participant and onscreen partner before the saccade was made could eliminate the temporal binding effects demonstrated before. The Active Gaze Leading condition in the second experiment was kept the same as before with the exception of the onscreen stimulus having closed eyes at the outset, instead of direct gaze. This enabled a check on whether it was the direct gaze, rather than the saccade, driving the binding effects, given it is known that direct gaze is a powerful social cue (Hamilton, 2016; for a review). Similarly, because there were closed eyes at the outset for the Passive Face task in the previous experiment, direct gaze from the onscreen face was used at the start, for

Experiment 2’s Passive Face task. No saccade towards the object was made and participants kept fixation on the face throughout.

One other important factor that needed to be tested relates to the fact participants began their interaction looking at the face before they began gaze leading in the Active Gaze Leading task. In everyday social exchanges we often do deliberate or intentional saccades to objects to direct another’s attention to it. However, we also sometimes find our gaze is followed by others even when we did not intend this to occur. I term this “incidental agency” such as when someone notices you are looking at something and so follows your gaze to see whether your attentional focus is also of relevance to them. To test whether incidental implicit agency also results in such an interaction a grey fixation dot (Courier, 18pt) was added where the object was due to appear in a Passive Object task. In this task, therefore, a situation in which gaze following was caused was simulated but not in the same intentional way as the Active Gaze Leading task. No saccade was made towards the object and participants maintained fixation throughout on the fixation dot placeholder and then the object, once it appeared, in the same place. Therefore, in this Passive Object task, the onscreen gaze response occurred when participants were already looking at the object, not having first performed a gaze leading saccade to it.

The procedure and task for participants was the same in all other respects for Experiment 2 as the previous experiment. The Active Gaze Leading task was identical to Experiment 1 except for the addition of a fixation dot as a place holder for the object to keep consistency with the displays in the Passive tasks. For clarity, in the Active Gaze Leading task the instructions were exactly the same as Experiment 1 so participants looked at the face and then at the object once it appeared.

Experiment 2 Results

Reproduction error and percentage reproduction. Mean reproduction error was calculated for each participant in each condition which is the reproduced time interval minus the actual time interval (see Table 2). Trials in which participants’ estimates were 3*SD*s above or below their individual means were removed (0.28% of trials). The reproduced time interval was divided by the actual time interval to calculate mean percentage reproduction (see Figure 10).

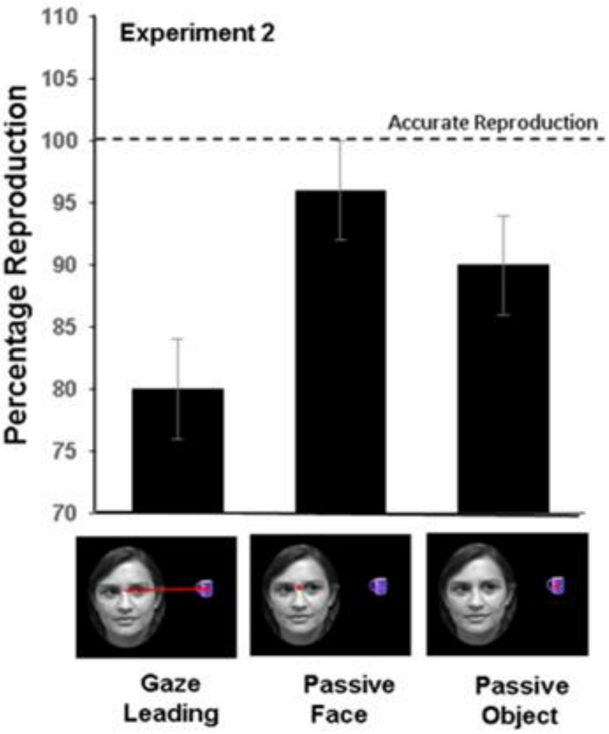

Figure 10. Mean percentage reproductions for Experiment 2. In the Gaze Leading task, participants looked first at the face, and then at an object as soon as it appeared. In the Passive Face task, participants looked at the face throughout. In the Passive Object task, participants looked at the placeholder/object throughout. The images show how the face was displayed when gaze onset occurred. Circles and the arrow were not displayed but represent where participants were instructed to fixate (and the saccade from the face to the object for the Active task). Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus & Masson (1994).

The same processing and analysis was performed on the data as in Experiment 1. First, in order to establish whether each condition produced temporal compression (reliable under-reproductions of the time between object and gaze onset), or relatively accurate reproductions, single sample *t*-tests for each of the three conditions on the percentage reproductions were performed. This showed that temporal compression was statistically significant in the Active Gaze Leading condition. Here, participants reproduced the temporal gap by *M*=80%, 95% CI [73,86] (*SD*=19%), *t*(31)=6.18, *p*<.001, d_z=1.55, 95% CI [0.98, 2.10]. In the Passive Face condition, reproduction did not differ statistically from 100% reproduction (Passive Face Fixation condition: *M*=96%, 95% CI [88, 104], *SD*=23%, *t*(31)=1.00, *p*=.327, d_z=0.25, 95% CI [-0.24,0.74], but did in the Passive Object Fixation condition, *M*=90%, 95% CI [82, 98], *SD*=22%, *t*(31)=2.70, *p*=.01, d_z=0.67, 95% CI [0.17, 1.18]. There was a main effect of task, *F*(2,62) =21.45, *MSE*=0.221, *p*<.001, η_p^2=0.409, and follow-up contrasts showed that the proportional temporal compression effect in the Active Gaze Leading condition was greater than in both the Passive Face Fixation, *t*(31)=6.02, *p*<.001, d_z=0.79, 95% CI [0.46, 1.11] and Passive Object conditions *t*(31)=4.17, *p*<001, d_z=0.51, 95% CI [0.23, 0.77].

Secondary measures, manipulation checks and participant subset analyses. As in Experiment 1, greater explicit agency was reported following the Active Gaze Leading (3.97, *SD*=1.79) 95% CI [3.32, 4.61] than both the Passive Object Fixation (2.72, *SD*=1.57) 95% CI [2.09, 3.35] and Passive Face Fixation (2.59, *SD*=1.50) 95% CI [2.05, 3.13] conditions (*t*s>3.6, *ps*<.001, d_zs>0.7). In the critical Active Gaze Leading condition, mean saccadic reaction time was 219ms (*SD*=57ms), and mean saccade duration for the gaze leading saccade was 79ms (*SD*=69). There were only 0.6% of trials where the onscreen face gaze shift occurred before the participant's saccade was completed. The same check as Experiment 1 was performed, by reanalysing the data with the nine participants excluded who experienced a gaze shift onscreen before their saccade was completed. This was for only an average of 1.22 trials per participant. These nine participant's mean explicit ratings were not different to the rest of the sample (M = 3.66, SD=1.87 and M = 4.01, SD=1.75, respectively). The data showed a remarkably similar pattern. The Active Gaze Leading condition revealed temporal compression – participants reproduced 76%, 95% CI [68,84], *SD*=19% of the veridical time interval, *t*(22)=6.12, *p*<.001, d_z=1.81, 95% CI [1.11,2.48]. The Passive Face Fixation condition did not produce temporal compression (*M*=92% reproduction, 95% CI [82,101] *SD*=23%, *t*(22)=1.77, *p*=.091, d_z=0.52 95% CI [-0.07,1.11]. However, the Passive Object Fixation task did reveal reliable under-reproductions, of about one third less than that in the active condition; *M*=84% reproduction, 95% CI [76,93] *SD*=19%, *t*(22)=3.87, *p*=.001, d_z=1.14, 95% CI [0.51,1.76].

In the critical Active Gaze Leading condition, mean sRT was 219ms (*SD*=57ms), and did not correlate with the temporal compression effect, *r*=-.06, *n*=31, *p*=.77. (There was no eye tracking data due to technical reasons for one participant so *n*= 31 for these analyses of sRT and saccade durations). Mean saccade duration for the gaze leading saccade was 79ms (*SD*=69) and did not correlate with temporal compression either, r=.26, n=31, p=.16.

Saccades to the object in error were made on only 0.33% of trials during the Passive Face task. In the Passive Object task of Experiment 2, saccades in error away from the object to the face were made on only 0.06% of trials. Therefore, passive tasks were not compromised by erroneous saccades, just like Experiment 1, as these were so small in number. The same split half analysis of binding by temporal interval was carried out as Experiment 1. The median time interval was 1376 ms and the mean was 1375 ms (*SD*= 541). The mean temporal interval for the shortest half was 905 ms (*SD*= 271) with a mean RE of -10 ms (*SD*= 352), meaning there was very little under-reproduction for the shortest half. The mean temporal interval for the longest half was 1844 ms (*SD*= 267) with a mean RE of -557 ms (*SD*= 465), meaning there was under-reproduction in the longest half. The shortest and longest halves were significantly different, $t(31)=14.53$, $p<.001$, $d_z=2.57$, showing larger effects with the longer intervals, again supporting the notion that these are, indeed, temporal binding effects.

Correlation analysis of explicit and implicit agency findings. Any correlations between explicit and implicit agency were examined for each of the three conditions and between the Agency Effects for the Active Gaze Leading task compared to the Passive Object task. All correlations were non-significant, *r*s< 0.20, *p*s>.28.

Correlation analysis of AQ. The mean AQ score was 15.06 (*SD* = 6.35) and the range was 5 to 33. Just like Experiment 1, no significant correlations were found between AQ and RE for all tasks, *r*s< 0.17, *p*s>.35.

Experiment 2 Discussion

The findings of temporal binding resulting from having an Active Gaze Leading towards an object followed, replicated but now there is another, novel finding of binding in the Passive Object task, albeit attenuated compared to the Active Gaze Leading task. The findings for Passive Face (no binding) replicated that of Experiment 1. I now summarise the results from Experiments 1 and 2 together. If an Active Gaze Leading was made, binding was greatest. When no saccade was made, some binding resulted, but only if participants were already looking at the object. There was no binding in passive tasks if participants were looking at the face throughout. The same effects or lack of effects resulted whether or not participants experienced direct gaze or closed eyes with the participant before the gaze response. This will be discussed further in the Chapter Discussion as it relates to theoretical models of predictive or retrospective mechanisms underpinning the implicit agency process.

There were no reliable correlations between implicit and explicit agency and in Agency Effects between the two types of agency measures. This is evidence, like the last experiment, that implicit and explicit agency may be dissociable (Synofzik et al., 2008, 2013). The explicit agency ratings for both Passive tasks were low in both experiments, compared with higher ratings for the Active Gaze Leading task. Just like the previous experiment, no correlations between implicit or explicit agency and AQ were found.

Experiment 3

Experiment 3 Method

In this experiment a further, new group of participants of the same number (N=32) and type were recruited in the same way and given credits for participation, as before. Their mean age was 21.34 years (SD = 5.72, three males). This experiment was an exact replication of Experiment 1 for the Active Gaze Leading task, being the main novel finding to seek to replicate. In the final experiment the aim was to check any outstanding questions not yet tested in the previous experiments for Passive Object tasks.

For the Passive Object task in this experiment, whether the attenuated incidental implicit agency effect would replicate that found in Experiment 2 was explored, but with the onscreen partner having closed eyes at the outset, instead of the direct gaze used in Experiment 2. The Passive Phase Scrambled Object task in this third experiment was the same as the Passive Phase Scrambled task in Experiment 1, with the important difference that participants fixated the fixation dot and then the object throughout, instead of the rectangle throughout, to test whether an attenuated sense of agency would occur when there is a low-level shift in peripheral vision towards the object of gaze, but no gaze leading saccade is made. The prediction was that participants would not attribute agency in this case as they would not expect to cause rectangles to shift, unlike eye gaze shifts which they would expect to control. To pre-empt the results, this is the most unexpected finding across the three experiments as there was binding in this task, but explicit agency ratings remained low. I will put forward possible explanations for this potentially revealing finding. The procedure was the same as Experiments 1 and 2 with participants reproducing the same time intervals as before, being the time between the object appearing and the gaze or rectangle shift onset.

Experiment 3 Results

Reproduction error and percentage reproduction. Mean reproduction error was calculated for each participant in each condition which is the reproduced time interval minus the actual time interval (see Table 2). Trials in which participants' estimates were 3*SD*s above or below their individual means were removed (0.24% of trials). The reproduced time interval was divided by the actual time interval to calculate mean percentage reproduction (see Figure 11).

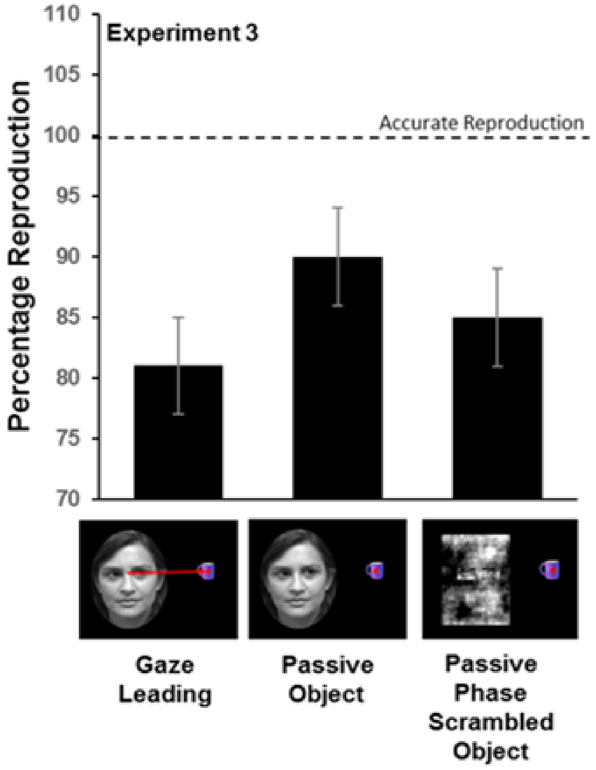

Figure 11. Mean percentage reproductions for Experiment 3. In the Gaze Leading task, participants looked first at the face, and then at an object as soon as it appeared. In the Passive Object and Passive Phase Scrambled Object tasks, participants looked at the placeholder/object throughout. The images show how the face/scrambled stimulus was displayed when gaze onset occurred. Circles and the arrow were not displayed but represent where participants were instructed to fixate (and the saccade from the face to the object for the Active task). Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus & Masson (1994).

The same processing and analysis was performed on the data as in Experiments 1 and 2. First, in order to establish whether each condition produced temporal compression (reliable under-reproductions of the time between object and gaze onset), or relatively accurate reproductions, single sample *t*-tests were performed for each of the three conditions on the percentage reproductions. This showed that temporal compression was statistically significant in all three conditions. In the Active Gaze leading condition, participants reproduced the temporal gap by *M*=81%, 95% CI [74, 89] (*SD*=22%), *t*(31)=-4.84, *p*<.001, d_z=1.21, 95% CI [0.63, 1.78]. In the Passive Object condition, participants reproduced the temporal gap by *M*=90%, 95% CI [82, 98] (*SD*=22%), *t*(31)=2.64, *p*=.013, d_z=0.66, 95% CI [0.14, 1.17]. In the Passive Phase Scrambled Object Fixation condition, reproduction was *M*=85%, 95% CI [75, 94], *SD*=25%, *t*(31)=3.45 *p*=.002, d_z=0.86, 95% CI [0.32, 1.39]. There was a main effect of task, *F*(2,62) =5,15, *MSE* =0.055, *p*=.009, η_p^2=.142, and follow-up contrasts showed that the proportional temporal compression effect in the Active Gaze Leading condition was greater than the Passive Object Fixation, *t*(31)=3.02, *p*=.005, d_z=0.38, 95% CI [0.11,0.64], but no different from the Passive Phase Scrambled Object condition, *t*(31)=1.09, *p*=.283, d_z=0.13, 95% CI [0.11,0.37].

Secondary measures, manipulation checks and participant subset analyses. As in Experiments 1 and 2, greater explicit agency was reported following the Active Gaze Leading (4.06, *SD*=2.00) 95% CI [3.34, 4.78], than both the Passive Phase Scrambled Object Fixation (2.00, *SD*=1.72) 95% CI [1.38, 2.62] and Passive Object Fixation (2.16, *SD*=1.57) 95% CI [1.59, 2.72] conditions (*t*s>5.03, *ps*<.001, d_zs>1.80). In the critical Active Gaze Leading condition, mean saccadic reaction time was 180ms (*SD*=43ms), and mean saccade duration for the gaze leading saccade was 71ms (*SD*=69ms). There were only 0.3% of trials where the onscreen face gaze shift occurred before the participant's saccade was completed. The same check as Experiments 1 and 2 was carried out, by reanalysing the data with the five participants excluded who experienced a gaze shift onscreen before their saccade was completed. This was for only an average of 1.20 trials per participant. These five participant's mean explicit ratings were not different to the rest of the sample (M = 4.07, SD=1.98 and M =4.06, SD=2.00, respectively). The data showed a remarkably similar pattern. The Active Gaze Leading condition revealed temporal compression – participants reproduced 81%, 95% CI [73, 89], *SD*=21% of the veridical time interval, *t*(26)=-4.66, *p*<.001, d_z=1.27, 95% CI

[0.63, 1.90]. The Passive Object Fixation condition produced temporal compression (*M*=89% reproduction, 95% CI [80, 97] *SD*=22%, *t*(26)=-2.70, *p*=.012, d_z=0.73 95% CI [0.16, 1.30]. The Passive Scrambled Object Fixation task also revealed reliable under-reproductions; *M*=83% reproduction, 95% CI [73, 92] *SD*=24%, *t*(26)=-3.78, *p*=.001, d_z=1.03, 95% CI [0.42, 1.63].

Just like Experiments 1 and 2, saccade metrics revealed that in the critical Active Gaze Leading condition, neither mean saccadic reaction time (*M*=223 ms, *SD*= 113 ms) nor saccade duration (*M*=78 ms, *SD*= 55 ms) correlated with temporal binding, *r*=-,21, *n*=32, *p*=.26, *r*=-.26, *n*=32, *p*=.16, respectively. In the Passive Phase Scrambled Object task, there were no saccades made away from the object to the scrambled rectangle. In the Passive Object task, saccades in error away from the object to the face were made on only 0.06% of trials. Therefore, passive tasks were not compromised by erroneous saccades, just like Experiments 1 and 2. The same split half analysis of binding by temporal interval was carried out as Experiments 1 and 2. The median time interval was 1290 ms and the mean was 1325 ms (*SD*= 546). The mean temporal interval for the shortest half was 853 ms (*SD*= 177) with a mean RE of 2 ms (*SD*= 390), meaning there was no under-reproduction for the shortest half. The mean temporal interval for the longest half was 1800 ms (*SD*= 303) with a mean RE of -503 ms (*SD*= 507), meaning there was under-reproduction in the longest half. The shortest and longest halves were significantly different, *t*(31)=8.38, *p*<.001, d_z=0.40, showing larger effects with the longer intervals, again supporting the notion that these are, indeed, temporal binding effects.

Correlation analysis of explicit and implicit agency findings. Any correlations between explicit and implicit agency for each of the three conditions was examined. There was a medium, positive correlation between implicit Active Gaze Leading task and explicit Active Gaze Leading task, *r*(32) = 0.41, *p* = .02 (two-tailed). Therefore, the greater the implicit agency, the less explicit agency was reported. This is in the opposite direction to that found trending in Experiment 1. There were no significant correlations between implicit and explicit findings for the Passive Object task or the Passive Phase Scrambled object task, *rs< 0.23, ps>.23*. Analyses were carried out to explore any correlations like those found in Experiment 1 between the differences in the Agency Effect between implicit and explicit findings, but no significant correlations were found, *r*s< 0.13, *p*s>.09. Combining data across all three experiments still revealed no significant c3orrelations between binding and explicit ratings.

Correlation analysis of AQ. The mean AQ score was 14.13 (*SD* = 8.48)) and the range was 1 to 33. Just like Experiments 1 and 2, there were no correlations between AQ and the Agency Effect, nor between AQ and degree of binding in any task, *r*s< 0.20, *p*s>.285. Combining the data across the three experiments in case they lacked power to detect a correlation still failed to reveal any significant correlations between AQ and binding or AQ and Agency Effects.

Experiment 3 Discussion

The findings from Experiments 1 and 2 of temporal binding effects in the Active Gaze Leading condition were replicated. This supports the hypothesis that making eye movements which cause an outcome produces perceived compression of time, like other motor actions. Therefore, an implicit sense of agency results from having gaze leading responded to. The attenuated binding effect for incidental agency found in Experiment 2 also replicated. Participants apparently felt some implicit agency when they were already looking at an object in the Passive Object task, whilst they did not self-report corresponding high levels of explicit agency perhaps because higher level processes led to a low agency judgement. This supports the balance of evidence that implicit and explicit agency are dissociable (e.g. Moore et al., 2012, Saito et al., 2015). Implicit agency for incidental agency is not surprising in the context of social interaction because our gaze is not only followed after a deliberate saccade to an object, but also incidentally when others notice our object of attention. Indeed, we are aware our gaze may sometimes be followed when we do not want it to be followed.

However, unexpectedly, this binding effect was even stronger, in the Passive Phase Scrambled Object task, when no saccade was made but a spatial shift occurred towards the incidental object of gaze. This observation could merely reflect a carry-over effect from the active task blocks (given the repeated measures design) so those participants who completed the Scrambled Passive Object block first were examined. However, those participants under-reproduced the interval by a similar amount to the sample as a whole, 86% reproduction, compared to 85% reproduction for the whole sample. I speculate that implicit agency resulted for one of three reasons, or a combination of them. Firstly, when there is a face we have a social context within which to make agency attributions. However, the situation is ambiguous in the Passive Phase Scrambled Object task. This lack of context makes it a safer option to over-attribute agency when there is a spatial shift towards our direction of gaze. It is adaptive to assume we caused an outcome for which we may be responsible and have to take further action. This explanation also fits with recent findings reported by Desantis, Waszak, and Gorea (2016) who found that participants over-attribute self-agency when they are in an ambiguous situation.

Secondly, the two rectangles shifting sideways resemble eyes moving on a face and so may be processed as such with face-like stimuli having been shown to produce gaze cueing effects, when perceived as faces (the “pareidolia” phenomenon, Takahashi & Watanabe, 2013). Interestingly, there were no order effects between the three blocks in Experiment 3, meaning participants were just as likely to experience implicit agency whether they experienced the Phase Scrambled Object task before or after the face tasks. Thirdly, the gaze cueing literature has often, but not always, found shifts of spatial attention result from arrows in a similar way to eyes (see Frischen et al., 2007, for a review). It may be that low level spatial shifts towards our object of gaze from non-faces, here the shifting rectangles akin to arrows, produce an implicit sense of agency within a non-social context, just as have been shown here to emerge in the social context of the faces.

A lack of reliable correlations between binding and explicit agency ratings across the three experiments points towards a dissociation between implicit and explicit agency and this will be discussed further in the Chapter Discussion. Finally, there were no correlations with AQ and binding, just like Experiments 1 and 2, suggesting no individual differences in implicit agency according to level of autism-like traits in a typically developing sample.

Chapter Discussion

Here, using a temporal binding paradigm in a novel context, a sense of agency resulting from initiating joint attention was investigated. In three experiments whether eye saccades, like other motor actions in previous research within a non-social context, can result in temporal binding, was tested and it was found that they can within a simulated social interaction context. Furthermore, incidental implicit agency effects were examined when no saccades are made but attention is already on an object when a gaze (or spatial shift) shift occurs, and an attenuated sense of implicit agency was found in those circumstances. I also explored whether implicit and explicit agency are associated or independent systems, finding evidence that they can be viewed as dissociable. Finally, no correlations between autism-like traits and binding and Agency Effects were found.

Temporal binding findings. This data demonstrates, for the first time, that having an eye movement responded to does induce a perception of a compression of time between its initiation and an outcome; the classic temporal binding effect (Moore & Obhi, 2012; for a review). I argue, as many have before me, that this is a putative marker of an implicit sense of agency (e.g. Kühn et al., 2013). Sense of agency is particularly important within the context of joint attention interactions. This is because controlling other’s gaze takes place within everyday social interactions and enables us to communicate our “theory of mind” to others (Premack & Woodruff, 1978), helping us achieve our goals (Baron-Cohen, 1997). Therefore, being able to monitor how we impact the behaviour of others is essential during these interactions. The lack of binding in Passive Face tasks, in contrast to binding in the Passive Object tasks, demonstrates that it matters where we are looking when we experience agency over another’s eye movements. If we are looking at an object already, we may attribute some agency to an observed congruent eye shift and we have demonstrated the same effect occurs when observing a spatial shift akin to an eye gaze shift, congruent with our location of gaze.

The lack of binding in passive conditions shows that the mere presence of a social stimulus does not interfere greatly with accurate timing of intervals per se. However, the weaker but reliable binding effect in the Passive Object Fixation task of Experiments 2 and 3 is worth further discussion. This observation could merely reflect a carry-over effect from the active task blocks (given the repeated measures design). However, upon examination of those participants who completed the Passive Object task first, the binding effect was still present, or even more so, as it was for the whole sample. In Experiment 2, there was a mean of 87% reproduction for those who experienced Passive Object first, compared to the whole sample’s mean of 90%. In Experiment 3 the mean reproduction for those who experienced Passive Object first was 84%, compared with the whole sample mean of 90%, so carry-over effects are an unlikely explanation for the effects found in that experiment too. Therefore, a more interesting (but speculative) suggestion would be that object-oriented attention in the presence of a face gazing at the same object might affect time estimation, even in the absence of a recently preceding action. It could be the case that if we are looking at an object already, we may attribute some agency to an observed congruent eye shift; but the effect is stronger if we have *recently* saccaded to that object (as in the Active Gaze Leading condition). This chimes with work highlighting the critical importance of objects in joint attention (Bayliss & Tipper, 2006; Bayliss et al., 2013; Lobmaier, Fischer, & Schwaninger, 2006).

Taking this evidence together, there may be a hierarchical system which attributes the greatest sense of implicit agency after intentional gaze leading and then an attenuated sense of implicit agency if a gaze shift is detected when we are already directing our gaze towards an object incidentally. It makes adaptive sense to monitor gaze in this way and to detect and attribute our causal role in these shifts of attention. Indeed, these data show that even a low level spatial shift towards our object of gaze will elicit implicit, although not explicit, sense of agency. As social, communicative and collaborative beings, we need to know what impact our own actions may be having on other's actions, whether or not we intended to influence them. An implicit, interpersonal agency effect during joint attention is one mechanism which can help us achieve that.

I now consider what these findings offer for the theories about a predictive or retrospective position, or a combination of both, for how sense of agency occurs. The data in this chapter do not support an exclusively predictive position. Because binding occurred even when the onscreen eyes were closed at the outset, this implies some retrospective inference has been made as we do not expect someone who is not looking at us to respond to our gaze signals. This supports a growing body of research which has shown retrospective processes are involved (Moore & Haggard, 2008; Moore, Lagnado, Deal, & Haggard, 2009). This is also consistent with the finding that angular gyrus activation shows, at a neural level, that there is a comparison of predicted outcome with actual outcome (Chambon et al., 2013; Farrer et al., 2008).

Moore and Obhi's (2012) review of evidence on the intentionality or causality debate, concluded that both are necessary for binding to occur but either one alone is not enough. Causality seems to have played the main role in the findings for the Passive Object tasks because intentionality was not manipulated in that task. Participants were told they were not causing the gaze shift in the Passive Object task but binding still resulted, albeit weaker than in the Active Gaze Leading tasks where participants were told they were causing the gaze shift. This contrasted with the explicit low agency ratings for the Passive Object task. This supports the position of Buehner (2012) who argued causality is the main driver of binding.

Further research using this paradigm but manipulating intentionality could shed further light on the relative contributions of causality and intentionality. For example, participants could be instructed to choose their preference between two objects by looking at the one they liked best, with the intention of making the on-screen face follow their gaze to that choice. If temporal binding continued to occur when gaze was not followed in that paradigm, then it would be hard to argue that intentionality had driven the effect.

Explicit and implicit agency. Taking all three experiments together, no reliable correlations were found between implicit and explicit agency. This points towards a dissociation between the two types of agency as previously proposed (Moore et al., 2012; Saito et al., 2015; Synofzik et al., 2008). The data does, nevertheless, support the idea that there is some relationship between the two types of agency, rather than them operating completely independently. This is because, whilst individuals' degree of binding and degree of self-reported feelings of control did not correlate, nevertheless, the largest sense of explicit agency was reported in the Active Gaze Leading task compared to the lower ratings in the Passive tasks in all three experiments, just as the binding effects were largest in the Active tasks compared to the Passive tasks. The findings lend support to the existence of a dichotomy between what is experienced on a low, sensorimotor level during implicit agency and what is judged to be under our control on an explicit level, also supporting an extended version of the comparator model (see Moore et al., 2012; Synofzik et al., 2008, 2013). This is because, on an implicit level, binding occurred in the Passive Phase Scrambled Object task when participants were already looking at the object but explicitly in that task they made self-reported judgments that they felt little control over causing the rectangles to shift.

Autism-like traits and binding and agency effects. There were no correlations between magnitude of temporal binding overall and level of autism-like traits. This accords with David et al. (2008) who found intact explicit sense of agency in an action monitoring task for those with a diagnosis of autism. The data also lend support to the idea that time perception may be intact for those with an autism diagnosis (Wallace & Happé, 2008). However, in a temporal binding paradigm similar to the one employed here, Sperduti et al. (2014) did find reduced temporal binding in their sample with an autism diagnosis, compared with control.

Similarly, there was no association with autism-like traits and the magnitude of the Agency Effects. I had predicted that there would be a correlation here because of previous findings in my lab of weaker attentional orienting towards faces who followed gaze for those with higher levels of autism-like traits (Edwards et al, 2015). However, Edwards et al. (2015) demonstrated an attentional effect of gaze leading and, in the light of the findings here, sense of agency effects during joint attention appear distinguishable from attentional effects and not to be modulated by autism-like traits.

Limitations and Future Directions

I cannot rule out that the temporal binding effect demonstrated in these three experiments may not be a measure of implicit sense of agency, although much research has offered evidence that it is (See Moore & Obhi, 2012, for a review). It is also noted that these experiments were limited to demonstrating temporal binding within a computer simulated joint attention context, rather than a naturalistic interaction.

Another methodological limitation is that the direct interval replication method was used to measure temporal binding. Some studies have used other binding measures such as instructing participants to make evaluations of intervals on an analogue scale in milliseconds (e.g. Kühn et al., 2013). These studies could be repeated using analogue scale reports in place of direct interval replication to explore whether that method is more or less sensitive to the binding effect. In addition, as Hughes, Desantis, and Waszak (2013) point out, by employing the direct interval replication method we cannot isolate the relative contributions of a shift in the perception of the motor action or of the outcome effect to the effect of temporal binding. The same point can be made about any interval estimation procedure. To explore these relative contributions, reporting the position of the clock hands when the action or the outcome occurs would be needed as employed for the Libet clock method (Haggard et al., 2002), but this would be difficult in this paradigm where visual attention needs to be on the stimuli and not on the clock. There is also an argument put forward by Stetson, Cui, Montague, and Eagleman (2006) that there is a recalibration of order judgments rather than a perceived shifting backwards in time of the outcome and my paradigm cannot explore this interesting possibility.

Another possible future direction would be to manipulate the inter-event interval further as random intervals between 400 and 2300 ms were employed. It would be instructive to explore if longer intervals remove the effect altogether because the interaction may feel less naturalistic and so implicit agency may not result. It is possible that using very short intervals may also eliminate the effect if that feels less naturalistic. There is likely to be a “Goldilocks” time interval which feels just right. To explore the apparent dissociation further between implicit and explicit agency, this paradigm could be used but with the important difference that participants are told in all tasks that they are not causing the gaze or spatial shifts, even in the Active Gaze Leading tasks. I speculate that similar implicit Agency Effects would be found in the Active Gaze Leading task but that the explicit ratings for that task would decrease significantly.

There are a host of boundary conditions that remain untested in order to establish the conditions necessary and sufficient to produce indices of implicit agency in social contexts. One important future condition to test is to establish whether the observed gaze response needs to be congruent with the participant’s saccadic action, or can be any response (e.g. to avert gaze, or to change emotional expression, for example). I speculate that possibly an incongruent gaze shift might elicit binding if we feel we have caused another to look away from our direction of interest. Whether this would be binding of the same magnitude as a congruent gaze shift (or no binding at all) would be interesting for future studies to explore. The current results identify just one instance in which temporal binding can occur following a causal eye movement. Although determining the specificity of this effect is, of course, important for understanding the nature of the mechanisms involved, if future work were to demonstrate that the effect does generalise widely, this would not necessarily reduce the direct importance of this mechanism for understanding how social cognition is supported by such basic sensorimotor mechanisms.

One potential complication for the interpretation of these findings is that in both active and passive conditions, participants must detect the onset of the object in their periphery (while they are looking at the face). However, in the active tasks, the onset of the responding gaze shift is to be detected in their periphery because the participant is now looking at the object having performed a saccade, while in the passive conditions, the participant detects the gaze shift at their point of fixation, having not moved their eyes. This difference could have affected the speed of detection of the gaze shift across conditions. However, were participants to be slower to detect the gaze shift in their peripheral vision in the active task, this would have extended their time estimations, which means that the binding effects may have, if anything, been artificially relatively reduced. Despite this difference potentially working against the predictions, medium (Experiment 1) and large (Experiments 2 and 3) binding effect sizes emerged.

Future work could employ a gaze-contingent design to explore agency in social gaze interactions. The present work did not take this approach. If the action of the participant had been yoked more directly to the stimulus changes by using gaze-contingent stimuli, perhaps participants would report a greater explicit sense of agency than found here, and the temporal binding effects might have also been more stable. I did not employ a gaze contingent design here because I wished to avoid the introduction of a confound. Specifically, in the Active Gaze Leading task the to-be-estimated time interval would have included three periods of temporal lag that would not be present in the Passive conditions, making them not comparable without off-line adjustment. These lag periods are the saccade latency, the saccade duration and the eye-tracker uptake time to detect good fixation upon the object in order to cause the gaze shift. By not using gaze contingent stimuli, my chosen design afforded direct comparison of the actual time intervals across conditions. Nevertheless, it is clear that future studies should employ gaze contingent designs that circumvent the issues noted above to overcome this limitation of the present research. This would allow for even more robust tests of hypotheses regarding the temporal dynamics of social gaze.

Finally, only one female face was used for the gaze stimulus. The task is to hold the time of object appearance and gaze onset in working memory in order to reproduce a time interval. This task requires good concentration and so keeping the face stimulus constant throughout was thought to be best to avoid distraction from that task. However, this means the binding effects demonstrated are in predominately female samples using a female onscreen face. A potential fruitful line of research could follow up these findings by exploring effects for onscreen male faces compared to female faces. I speculate that there may well be revealing differences in a similar way to the gaze cueing sex differences demonstrated by Bayliss et al. (2005), where females were shown to be more responsive to gaze cueing.

Conclusions

These three experiments show, to the best of my knowledge uniquely, that temporal binding can occur when a gaze response is perceived to result from deliberate eye saccade bids for joint attention. Moreover, this implicit agency effect is within the context of a simulated joint attention interaction. I offer this as evidence for an implicit sense of agency for initiating joint attention interactions. Similarly, implicit agency can result when detecting a gaze shift towards our object of gaze, even if not intentionally caused. Finally, an incidental sense of agency may be felt even for low level spatial shifts towards our object of gaze, in an ambiguous, non-social situation. Given that this is a unique study into temporal binding within joint attention, the findings need replication, but show promise to elucidate the cognitive processes at work which produce a sense of agency during gaze interactions. There were no associations with level of autism-like traits and the magnitude of the temporal binding effect, suggesting time perception and agency may possibly be intact in the ASC population. The findings also support the theory that implicit and explicit agency are dissociable mechanisms. Future work is needed to explore this fascinating human experience of how we monitor our control over the world around us using our eyes and the effects of these motor actions on time perception.

Chapter 3: Further Investigation of Sense of Agency Over Responses to Gaze Leading

In the three experiments in Chapter 2, I hypothesised that that there would be greater temporal binding when participants' eyes were followed to an object (Active Gaze Leading conditions) than when no saccades to the object were made (Passive conditions). The data was consistent with this hypothesis, providing evidence that an implicit sense of agency is generated in the gaze leader when their saccade is followed, establishing joint attention. However, these findings left two key, unanswered questions. Firstly, was it simply the action of making a saccade which would always induce temporal binding rather than the context of the nature of the interaction with an onscreen face? Secondly, I had found attenuated temporal binding effects of about half the magnitude of the gaze leading effects when participants were already fixating object when the gaze shift response occurred. This seemed to evidence a form of incidental agency when some agency is felt over a congruent gaze shift towards our object of attention, even without a gaze leading saccade. This begged the question of whether a saccade was even necessary to induce temporal binding. Therefore, two further experiments were designed to test these questions. To pre-empt the results from both experiments, the data suggests that a gaze leading saccade may be necessary, but may not be sufficient by itself, to elicit the binding effects found in gaze leading conditions in Experiments 1-3.

There are only two studies I am aware of which have investigated agency and/or temporal binding for oculomotor actions. Firstly, Yabe and Goodale (2015) found temporal binding occurred between the intention to perform a saccade and the saccade itself, demonstrated by a shifting backwards in time of the perceived visual cue to saccade. This demonstrates that, in principle, a saccade may elicit binding effects between an intention to saccade and the performance of the saccade. Secondly, Grgič, Crespi, and De'Sperati (2016) found an explicit sense of gaze agency can arise from causing (via saccades) auditory beeps to co-occur with bouncing balls moving on a screen. This study evidenced the ability to become explicitly self-aware of controlling a gaze-based, non-social interface. What the first experiment in this chapter aimed to do was explore oculomotor agency during an interaction with a non-social stimulus with the same type of saccadic metrics as the social interaction simulated in the three experiments in Chapter 1.

Experiment 4

In Experiment 4, participants completed an interval reproduction task under two conditions manipulated within-subjects. In the “Saccade” task, participants attempted to reproduce the time interval between a fixation cross’ appearance, to which the participants were to immediately saccade, and an initial fixation cross’ enlargement. As typical for temporal binding paradigms, performance in the active “Saccade” condition was compared with a ‘passive’ condition in which no action (here, a saccade) is made by the participant. In the “No Saccade” condition participants fixated the first cross throughout but still reproduced the time interval between the second cross appearing and the first cross enlarging. The spatial distance between the crosses was matched to the distance between the face and object in Experiments 1-3, to ensure meaningful comparison. In both experiments, I have reported how I determined sample size, all data exclusions (if any), all manipulations and all measures.

Experiment 4 Method

Participants. Thirty-two participants (mean age=19.69 years; 2 were men) completed the study in return for course credit. The target sample size was determined by matching that of the samples in the three previous experiments where medium or large effect sizes for the gaze leading binding effects were found. Participants reported normal or corrected-to-normal vision. Ethical approval was granted by the School of Psychology Ethics Committee, University of East Anglia. All participants were drawn from the Psychology undergraduate programme, were naïve to the aims of the study and gave written informed consent. One participant whose data revealed the essential instruction to attempt to reproduce the time intervals had not been followed was excluded. Henceforth, n=31.

Stimuli. The stimuli were a white first fixation cross (Courier, 24pt) initially (but enlarging to Courier, 48pt) presented 5 cm left-of-centre onscreen centrally and a second white fixation cross (Courier, 24pt) presented centrally onscreen and 11.5 cm to the right of the first cross. Stimuli appeared on a black background and were presented using E-Prime 2.0 software (see Figure 12).

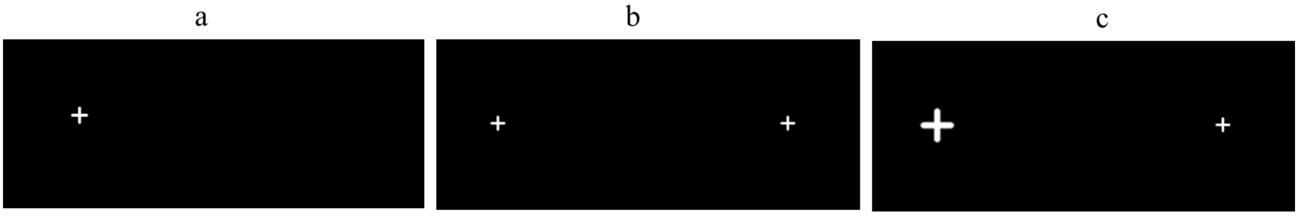

Figure 12. Trial sequence for the Saccade task. Participant looked at the first cross (a), displayed for 1000ms. Participant made a saccade (b) to the second cross as soon as it appeared. After a random inter-event interval of 400ms to 2300ms, the first cross enlarged (c). After 1000ms, estimate instruction appeared until response. Participants pressed and released the space bar to replicate the inter-event interval. The inter-event interval is the time between the second cross appearing and the first cross enlarging.

Apparatus and materials. Right eye position was tracked with an infrared eye tracker (Eyelink 1000, SR Research, Ontario, Canada; resolution 0.1^0, 500 Hz). A chin rest was used to maintain head stability. Viewing distance was 70cm from eyes to a 45 cm monitor (resolution 1024×768 pixels). A standard keyboard was used for manual responses. A 1-8 scale was used for participants' self-reported feelings of agency in each condition, with 8 representing the highest feeling of agency.

Design. The within-subjects design had two blocked conditions of 56 trials each. Block order was counterbalanced across participants. The conditions were Saccade and No Saccade. The dependent measures were the reproduction error (RE), calculated in milliseconds as the participants' reproduction of the temporal interval between two events minus the veridical temporal interval, and the percentage reproduction, calculated as the veridical temporal interval divided by the reproduced time interval. The inter-event interval was the time between a second fixation cross' appearance and a subsequent enlargement of the first fixation cross. The temporal interval between the second cross' appearance (rather than saccade onset) and the first cross enlarging was used to allow direct comparison between conditions (as no saccades were made in the No Saccade condition).

Procedure. Each experimental block commenced with a standard nine point eye tracking calibration, then 8 practice trials, then 56 experimental trials (see Figure 12). In the Saccade task, each trial began with the presentation of the first fixation cross on the left side of the screen. Participants were instructed to look at the cross (presented for 1000 ms) until a second fixation cross appeared on the right. This sudden onset was the participant's cue to immediately saccade to the second fixation cross. After a randomly selected inter-event interval of 400-2300ms following the onset of the second cross, the first cross enlarged to twice its original size. After 1000ms, the word “Estimate” appeared (white font, Courier, 18pt) above and below the crosses. This prompted the participant to manually press and hold down the spacebar for a duration that to their best ability replicated the time interval between the second cross’ appearance and the first cross’ enlargement. Participants were given no feedback about their responses. Finally, after releasing the spacebar, the display cleared to black for 1000ms.

The design relied on the low variance of saccadic RT and spatial acuity in this very simple eye movement task to ensure that the minimum temporal gap of 400ms was greater than the vast majority of saccades. Moreover, timing the temporal gap from a single fixed onset that occurred in both conditions (the second cross onset) afforded a straightforward and direct comparison between conditions. Finally, in the construction of the task, the fact that merely making a saccade can compress perceived temporal durations (saccadic suppression) was considered. However, this effect has been shown to operate on a much smaller scale than the effects anticipated here (Morrone et al., 2005), and is likely offset by chronostasis (Merchant & Yarrow, 2016).

The baseline control condition, in which I predict accurate temporal reproduction intervals was the No Saccade task. This was identical to the Saccade condition, except that the participant maintained fixation throughout on the first fixation cross. After each task (at the end of a 56 trial block) participants self-reported their degree of felt control over the first cross enlarging. The instruction was “Please rate how much control you felt over the first cross enlarging from 1 to 8, 1 meaning no control at all to 8 meaning a lot of control.”

Experiment 4 Results

Reproduction error and percentage reproduction. Mean reproduction error was calculated for each participant in each condition which is the reproduced time interval minus the actual time interval (see Table 3).

Table 3.

Reproduction errors for each condition for Experiments 3 and 4. Confidence intervals are based on 1000 bootstrap samples.

Condition	*M*	*SD*	Explicit Rating
E4 No Saccade	60 [-70, 190]	355	2.10 (*SD*=1.64)
E4 Saccade	-120 [-270, 28]	406	2.13 (*SD*=1.45)
E5 Passive Face Fixation	87 [-41, 216]	357	2.50 (*SD*=2.00)
E5 Passive Object Fixation	-27 [-124, 71]	271	3.03 (*SD*=1.51)

Abbreviations: M = Mean. SD = Standard Deviation.

Trials in which participants' estimates were 3*SD*s above or below their individual means were removed (0.58% of trials). The reproduced time interval was divided by the actual time interval to calculate mean percentage reproduction (see Figure 13). Confidence intervals around effect sizes are reported, using ESCI (Exploratory Software for Confidence Intervals) to calculate these (Cumming & Calin-Jageman, 2017). First, in order to establish whether each condition produced temporal binding (reliable underestimations of the time between second cross onset and first cross enlarging), or relatively accurate reproductions, two single sample *t*-tests were performed for each condition. This showed that temporal binding was not statistically significant in either the Saccade or No Saccade conditions. In the Saccade condition participants reproduced 92%, SD= 22%, 95% CI [74, 89] of the temporal interval, *t*(30)=1.45, *p*=.159, d_z=-0.37, 95% CI [0.14, 0.87]. In the No Saccade condition, participants reproduced 105%, SD= 27%, 95% CI [95, 115] of the temporal interval, and this did not differ statistically from zero, *t*(30)=1.02., *p*=.315, d_z=0.26, 95% CI [0.76, 0.25]. Next, to test for differences between conditions, a paired sample t-test revealed percentage reproductions were greater for the No Saccade task than the Saccade Task, *t*(30)=4.48, *p*<.001., d_z=0.45 95% CI [0.22, 0.67].

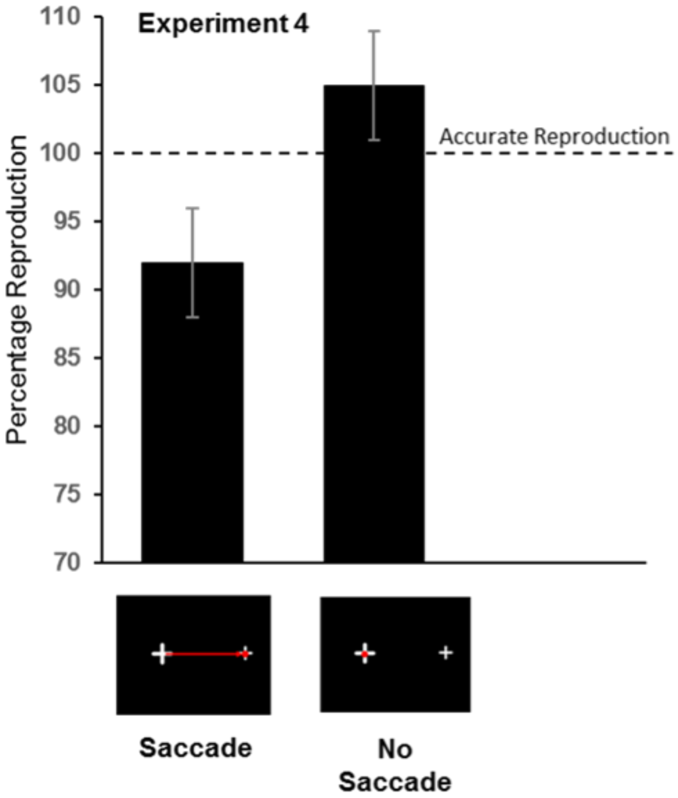

Figure 13. Mean percentage reproductions by condition for Experiment 4. In the Saccade task, participants looked first at a fixation cross, and then at a second fixation cross, as soon as it appeared. In the No Saccade task, participants looked at the first cross throughout. The images show how the crosses were displayed when the second event occurred. Circles and the arrow were not displayed but represent where participants were instructed to fixate (and the saccade from the first cross to the second cross for the Saccade task). Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus and Masson (1994).

Secondary measures and manipulation checks. Mean self-reported explicit ratings of agency were low and similar for the Saccade task (2.13, *SD*=1.45) 95% CI [1.60, 2.66] and the No Saccade task (2.10, *SD*=1.64) 95% CI [1.50, 2.70], *t*(30)=0.71, *p*= .865. Mean saccadic reaction time in the Saccade task was 200 ms (*SD*=43) and mean saccade duration was 84 ms (*SD*=91). In the No Saccade task, saccades in error to the second fixation cross were made on only 0.95% of trials.

Experiment 4 Discussion

There were no binding effects for the No Saccade task, showing that we do not under-reproduce time intervals between events which involve no action on our part. This is as expected and consistent with the temporal binding literature (see Moore & Obhi, for a review, 2012) and Experiments 1-3 in Chapter 2. Moreover, it has been shown in the Saccade condition that the mere action of making a saccade which is followed by an outcome, devoid of social context, does not cause us to reliably bind our saccade and the outcome together in time. Therefore, the gaze leading effects found in Experiments 1-3 seem to require some form of social context in order to elicit an implicit sense of agency for a gaze response following a saccade to an object. Saccades alone are not sufficient to drive the gaze leading effects previously demonstrated. I note there was some numerical, although non-significant, under-reproduction of the veridical time interval in the Saccade condition and the potential implications of that will be discussed further in the Chapter Discussion.

Experiment 5

In Experiment 5, I wanted to see if the findings of attenuated binding when participants were already fixating the object when gaze shift occurred (incidental agency) would replicate. This had been found in Experiments 2 and 3 in the Passive Object condition. However, another question was whether the binding effects can be boosted to those found in the gaze leading task simply by informing participants that their fixation on the object was causing the gaze shift. In Experiments 2 and 3, for the Passive Object task, participants were told they were not causing the gaze shift and attenuated binding resulted of about half the magnitude of the gaze leading task when a saccade was made. Now that Experiment 4 has established that saccades alone do not drive reliable binding effects outside of a social context, could stronger binding effects result with no saccade at all within a social context but by manipulating belief of causation explicitly? To examine this the Passive Object task from Experiments 2 and 3 was repeated, but participants were now told they were causing the gaze shift. If this resulted in strong binding, then, within a simulated interpersonal context, a gaze leading saccade may not be necessary to elicit reliable binding, if participants feel they are, nevertheless, causing the gaze shift towards their incidental object of interest.

Experiment 5 Method

A new sample of participants (*n*=32; mean age=20.47 years, three were men) was recruited from the same population as Experiments 1-4 and took part in return for course credits. The stimuli were the same as the Passive Object tasks in Experiments 2 and 3. The blocked design was manipulated within subjects and counter-balanced. The Passive Face condition was the same as main Experiment 1 with participants fixating the face throughout. The Passive Object Fixation task was the same as Experiments 2 and 3 with participants fixating the placeholder dot or the object throughout (see Figure 14) , but it was emphasised to them they were causing the gaze shift in the Passive Object Fixation task, but not in the Passive Face Fixation task. To be clear, in the Passive Object Fixation task, the onscreen gaze response occurred when participants were already looking at the object, not having first performed a gaze leading saccade from the onscreen face to it.

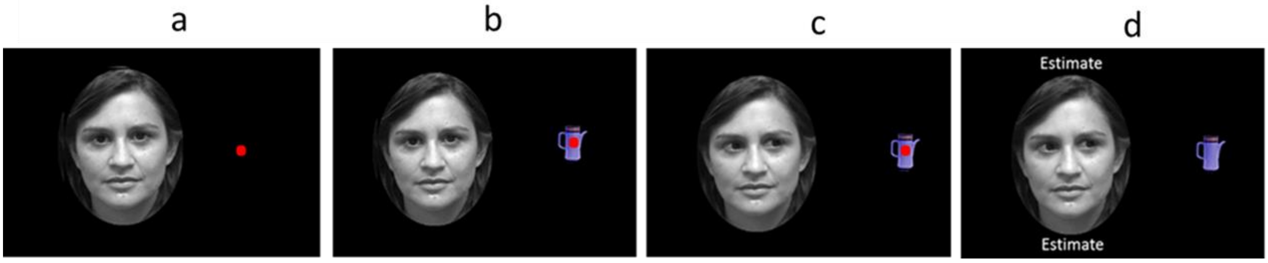

Figure 14. Trial sequence for the Passive Object task. Red circles were not displayed but represent where participants were instructed to fixate. Participant looked at a fixation dot placeholder (a), whilst the face was displayed for 1000ms. Participant fixated on the object (b) when it appeared in the same place as the place holder. After a random inter-event interval of 400ms to 2300ms, gaze onset (c) occurred. After 1000ms, estimate instruction appeared (d) until response. Participant pressed and released the space bar to replicate the inter-event interval. The inter-event interval is the time between the object appearing and the gaze onset.

Experiment 5 Results

Reproduction error and percentage reproduction. Mean reproduction error was calculated for each participant in each condition which is the reproduced time interval minus the actual time interval (see Table 3). The reproduced time interval was divided by the actual time interval to calculate mean percentage reproduction (see Figure 15). Trials in which participants' estimates were 3*SD*s above or below their individual means were removed (0.36% of trials).

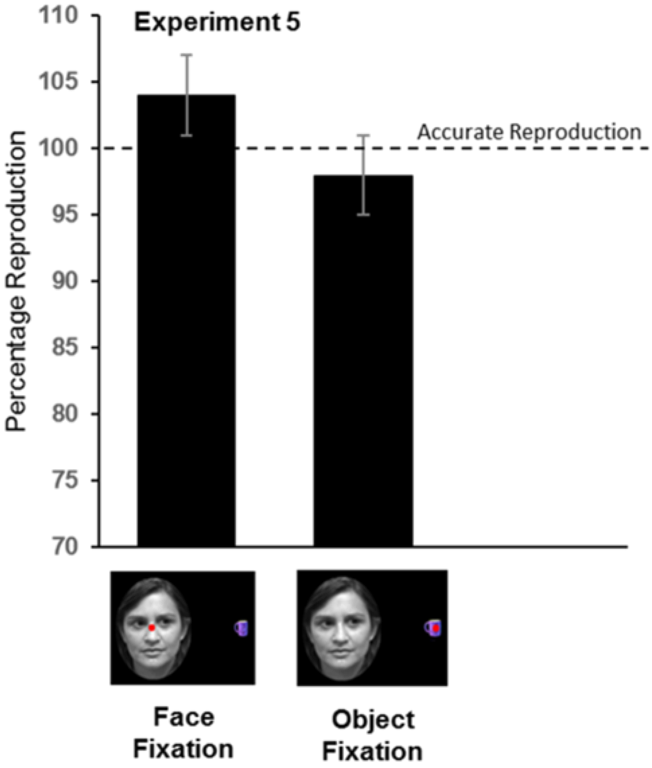

Figure 15. Mean percentage reproductions by condition for Experiment 5. In the Passive Face Fixation tasks, participants looked at the face throughout. In the Passive Object Fixation task, participants looked at the placeholder/object throughout. The images show how the face was displayed when gaze onset occurred. Circles were not displayed but represent where participants were instructed to fixate. Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus and Masson (1994).

The same processing and analysis was performed on the data as Experiment 4. The Passive Object Fixation condition, unlike in Experiments 2 and 3, revealed no significant temporal binding: 98% interval reproduction (*SD=22*), 95 % CI [90, 106], *t*(31)=0.45, *p*=.642,

d_z=0.12, 95% CI [-0.37, 0.61]. As in Experiments 1-3, the Passive Face Fixation condition did not produce binding, with an interval reproduction of 104% (*SD*=25, 95% CI [95, 113], *t*(31)=0.92, *p*=.366, d_z=-0.23, 95% CI [-0.72, 0.27]. The conditions differed, with greater under-reproductions in the Passive Object task than the Passive Face Fixation task *t*(31)=2.15, *p*=.039, d_z=0.25, 95% CI [0.01, 0.48], representing a small effect.

Secondary measures and manipulation checks. Greater explicit agency was reported following the Passive Object Fixation task (3.03, *SD*=1.51) 95% CI [2.49, 3.58], than the Passive Face Fixation task (2.5, *SD*=2.00) 95% CI [1.78, 3.22], although there was no statistical significance between these ratings, *t*(31)= 1.87, *p*=.071. This suggests the manipulation of causation was not as successful as when a gaze leading saccade precedes the gaze shift response as in Experiments 1-3. In the Passive Face Fixation task, saccades in error to the object were made on only 0.50% of trials. In the passive Object task, erroneous saccades to the face were made in only 0.06% of trials.

Experiment 5 Discussion

The null binding effect for the Passive Face Fixation tasks from Experiments 1-3 replicated. However, the attenuated binding effects found in Experiments 2 and 3 for the Passive Object Fixation task, did not replicate with no reliable binding, even though those attenuated effects were attempted to be boosted by explicit manipulation of belief of a causal link between fixation on the object and the subsequent gaze shift. This suggests that saccades might be necessary to find consistent binding effects and that manipulating causality alone may not drive the effects. However, there was a smaller under-reproduction in the Passive Object condition, compared to the Passive Face condition, where a small over-reproduction was made, albeit representing a small effect. This suggests there is some difference, nevertheless, in the two tasks which leads participants to reproduce smaller time intervals when their gaze is fixated on a referent object. Perhaps never experiencing a gaze shift towards their direction of gaze following a gaze leading saccade, like they did experience in Experiments 1-3, accounts for why no significant under-reproductions were found in this experiment. Nevertheless, there is some suggestion of greater implicit agency in the Passive Object task over the Passive Face task. Order effects were checked and revealed those half of the sample who completed Passive Face Fixation first had a mean RE of -26.53 ms, compared with the whole sample mean of 38 ms. Those who experienced Passive Object Fixation first had a mean RE of 68 ms, compared with the whole sample mean of -68 ms. This shows no clear order effects.

Chapter Discussion

These two further experiments aimed to answer some unresolved questions arising from Experiments 1-3. Firstly, Experiment 4 is evidence that performing a saccade alone cannot account for the temporal binding effects in the Active tasks in Experiments 1-3. Previous temporal binding studies, employing a traditional button press for the action have not investigated whether the motor action button press alone causes temporal compression in the way that has been investigated here for the saccade action. Furthermore, the field has not considered that participants are making saccades all the time during a binding experiment and that those saccades may elicit some temporal compression not controlled or accounted for in the design. This opens up a new question for the temporal binding field generally. Although no reliable binding in the Saccade task was found, the numerical under-reproduction of ~120ms, together with the previous findings of saccadic temporal compression reported by Morrone et al. (2005), suggest there may be circumstances in which reliable binding might occur for saccades alone if other paradigms or manipulations are employed. For example, if participants were told they were causing the cross to enlarge in Experiment 4, binding might have resulted.

Binding within the context of a simulated social interaction has been demonstrated in Experiments 1-3. However, that does not mean I believe only social contexts would elicit binding from saccades. We have seen from Experiment 3 (Chapter 2) that binding occurred in a non-social context when the situation was unfamiliar and ambiguous (in the Phase Scrambled Object Fixation task). My notion is that any outcome which participants believed they caused *or may have caused* by their eye movement could elicit binding. This is a potentially rich avenue for future studies to explore.

The second question arising from the experiments in Chapter 2 was whether an incidental sense of agency can result from a gaze shift towards one’s object of gaze, without being preceded by a deliberate gaze leading saccade from the face to the object. This was found to be the case in Experiments 2 and 3, albeit with attenuated effects compared to the active gaze leading effects. In Experiment 5, this effect did not replicate, despite participants being informed they were causing the gaze shift. This casts doubt upon the original suggestion in Chapter 2 that there could be a hierarchical system which elicits greater binding following a deliberate joint attention bid and lesser binding for incidental congruent gaze shifts. It may be that only reliable binding was found in Experiments 3 and 4 for Passive Object Fixation tasks because of carry-over effects from completing the Active Gaze Leading task. However, having examined those participants who completed the Passive Object task first in Experiment 2, it was found their reproductions were of a similar magnitude (87%) as those for all participants in that task (90%). Similarly, those who completed the Passive Object task first in Experiment 3 made 86% reproductions, compared to reproductions of 85% for the whole sample. Carry-over effects are, therefore, an unlikely explanation for those results. Future studies are needed, therefore, to seek to resolve the mixed findings for incidental joint gaze, particularly given that there was a difference between conditions in Experiment 5, with numerical under-reproductions for the incidental joint gaze task, but over-reproductions when there was no joint gaze upon an object.

Conclusion

Taken together with the findings from Experiments 1 to 3, Experiments 4 and 5 show that, whilst gaze leading saccades may be necessary to drive binding effects within a social context, they are not sufficient by themselves without a social context.

Chapter 4: Examining Effects of Gaze Responses on Face Recognition

Face recognition is a highly important part of our social interaction (Haxby, Hoffman, & Gobbini, 2000). At a very basic level, we need to know, in an instant, whether a face is familiar to us or a stranger we have never met before. Our eye gaze interactions form part of how we process faces (Haxby, Hoffman, & Gobbini, 2002) and so may affect how we encode a new face. The studies in this chapter aim to explore any effects the responses we receive to gaze leading may have on face recognition. I will outline what is known about face processing and recognition generally and then some differences and similarities between typically developing and autism populations. Finally, I will summarise some recent work which has begun to investigate the effects upon memory of gaze-based interactions and how the studies in this chapter were designed to build upon those findings.

Face Recognition

Processing faces for identification, whether for group classification such as gender or for individual identity recognition, is an essential part of our function as typical social beings (Ellis, 1975). It is critical for us to distinguish whether a face is familiar or a stranger. Indeed, we appear to have an innate potential for becoming experts at face processing, evidenced by a very early preference, even in new born infants, for attending to faces for longer (Goren, Sarty, & Wu, 1975) and further away (Johnson, Dziurawiec, Ellis, & Morton, 1991) than other stimuli. In addition, there is evidence we can discriminate between the face of our mother and a stranger as early as one month of age (Sai & Bushnell, 1988). This potential for developing face processing expertise may develop from a preference for the top-heavy configurations that faces have, rather than faces themselves (see Turati, 2004, for a review) but, nevertheless, face processing expertise is evident during early infancy in the typically developing population. This ability is supported by specialised neural regions for face processing in the ventral visual stream, including the fusiform face area (FFA; Kanwisher & Yovel, 2006), and the occipital face area (OFA; Pitcher, Walsh, & Duchaine, 2011). However, importantly, the literature has recently begun to emphasise how this expertise which we have for recognising faces only applies to familiar faces and that unfamiliar face recognition ability is more limited (Kramer, Young, & Burton, 2018; Young & Burton, 2017a, 2017b). For a contrary view, see Rossion (2018).

There is a considerable body of evidence that we process faces holistically (see a meta-analysis by Richler & Gauthier, 2014). Individual differences have been identified in facial recognition ability along a continuum, from the impaired ability found in the condition known as prosopagnosia (McNeil & Warrington, 1993) at one end, through to the exceptional ability to recall a face found amongst so-called "super-recognisers" at the other (Russell, Duchaine, & Nakayama, 2009). One commonly used measure of face recognition ability is the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006) which seeks to identify scores for recognition ability, ranging from low to high ability, with normative scores for comparison.

Face Recognition in Autism

Face recognition and processing abilities in those with autism has been subject to quite a large body of studies. A behavioural review by Weigelt, Koldewyn, and Kanwisher (2012) found that people with autism do not perform as well as typically developing samples on facial recognition tasks, although there is no evidence of qualitative differences in the way that faces are processed, for example, the face inversion effect (Yin, 1969) has mostly been found to be intact for those with autism. Performance on the Cambridge Face Memory Test (CFMT) is apparently worse for those with autism than typically developing (Kirchner, Hatri, Heekeren, & Dziobek, 2011; O'Hearn, Schroer, Minshew, & Luna, 2010). Hedley, Brewer, and Young (2011) found a similar pattern of results for those with autism on the CFMT but offered additional, more nuanced evidence that performance is on a continuum for those with and without autism as some autistic individuals can out-perform typically developing individuals.

Other research has examined autistic-like traits (AQ: Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) within the typically developing population and found no significant correlation between face recognition scores as measured by the CFMT and AQ scores overall (Rhodes, Jeffery, Taylor, & Ewing, 2013). However, Rhodes et al. (2013) did find that for males there was an association between face-selective recognition ability and AQ sub-scale scores for social skills. Higher overall AQ scoring males displayed lower adaptive coding of identity. Women, however, had greater adaptive coding ability, the greater their AQ scores. In summary, it seems then, given the above reviewed evidence, that there are both differences and similarities in face processing abilities between those with and without an autism diagnosis and the picture is even more nuanced when considering the broader phenotype.

Joint Attention and Gaze Behaviour Effects on Memory

An essential component of face processing includes coordinating gaze behaviours with others during joint and shared attention (see Frischen, Bayliss, & Tipper, 2007, for a review). Some research has demonstrated that learning from observed gaze behaviours of faces influences subsequent gaze behaviours with those same faces (see Capozzi, Becchio, Willemse, & Bayliss, 2016; Dalmaso, Edwards, & Bayliss, 2016). Therefore, there are some preliminary indications that joint attention can interact with memory encoding and subsequent behaviours. There is evidence that gaze behaviours during joint attention can impact upon memory for the *object* of joint attention. For example, Kim & Mundy (2012) have demonstrated that gaze leading towards pictures enhanced recognition memory for those stimuli, compared to when responding to gaze cued pictures, even controlling for picture viewing time. Working memory effects have also been shown by Gregory and Jackson (2017) who demonstrated that being gaze cued to stimuli can enhance working memory for those stimuli, compared to when there was no joint attention on the items to be recalled. Gregory and Jackson (2018) went on to show that these effects do not emerge if the cue face has a barrier placed between it and the object. Therefore, the working memory enhancement in a gaze follower from gaze cueing seems to rely on perception of a shared perspective with the gaze leader.

There is some limited evidence that looked-at *faces* are rated as more trustworthy than faces which are not looked at (Kaisler & Leder, 2016). However, another study found no effects on how faces were evaluated according to whether they were looked at by faces with positive or negative emotional expressions (Landes, Kashima, & Howe, 2016). There has not been any previous research examining how face recognition may be affected by responses to joint attention initiation. Therefore, these studies in this chapter aimed to explore whether different gaze responses to joint attention initiation result in any differential memory encoding for those faces encountered. This was measured by ability to recognise if a face had been seen before or was a new face, not previously encountered.

The hypothesis was an open one. On the one hand, it is possible that gaze leading being reciprocated would enhance face recognition, given previous findings that gaze leading being followed increases attention to the faces who follow over those who respond with averted gaze (Edwards et al., 2015). However, it is also entirely possible that these attentional effects will not translate into enhanced recognition because it may be the uncooperative faces which are encoded stronger in memory because not being followed is the unexpected outcome. The social norm transgressors who avert gaze may, therefore, be preferentially encoded over the co-operators.

To explore whether there are any effects on face recognition, and the direction of those effects, a gaze leading task was used where participants encountered faces which either followed their gaze to an object or averted gaze. This gaze response was gaze contingent upon the participant's fixation on the referent object and was followed by a surprise face recognition task, and then individual differences measures were taken (AQ and CFMT).

Experiment 6

Experiment 6 Method

Participants. After ethical approval was obtained from the University of East Anglia Psychology Ethics Committee, 35 Psychology undergraduate student participants (mean age = 19.43 years, *SD* = 2.03 ; 3 males), gave written, informed consent and were granted course credits in return for participation. All participants reported normal or corrected-to-normal vision. The sample size was determined by an a priori power calculation conducted using G*Power (Faul et al., 2007) which indicated 34 participants would be required to achieve 80% power at .05 alpha, anticipating a medium effect size. I note here, however, that there is no indication from previous literature of what effect size to expect, because I believe this exploratory study, to my knowledge, may be the first of its kind.

Stimuli. The face stimulus set comprised of 80 smiling photographs (560 × 760 pixels) of 40 females and 40 males. Of these, 35 females and 35 males were taken from the Karolinska Directed Emotional Faces Set (Lundqvist, Flykt, & Öhman, 1988) and the remaining five females and five males were taken from the NimStim face set (Tottenham et al., 2009). Smiling faces were used for the experiments in this chapter to enhance participant engagement with the faces as the research question here was about remembering the faces (not necessary in Experiments 1-5 where one neutral face was used). The photographs of the faces included three versions: the original image displayed with direct gaze and two further versions which had been photo-shopped so that the eyes looked right (for the joint attention gaze response) or looked left (for the averted gaze response). The object stimuli set consisted of 8 objects commonly found in the kitchen (220 × 78 pixels), taken from Bayliss et al. (2006). The face was presented 4 cm to the left of the centre of the screen. The faces were scaled to appear approximately life-sized onscreen. The objects were presented to the right of the face with the centre of the object 15 cm from, and in line with, the bridge of the nose. Examples of stimuli are illustrated in Figure 16.

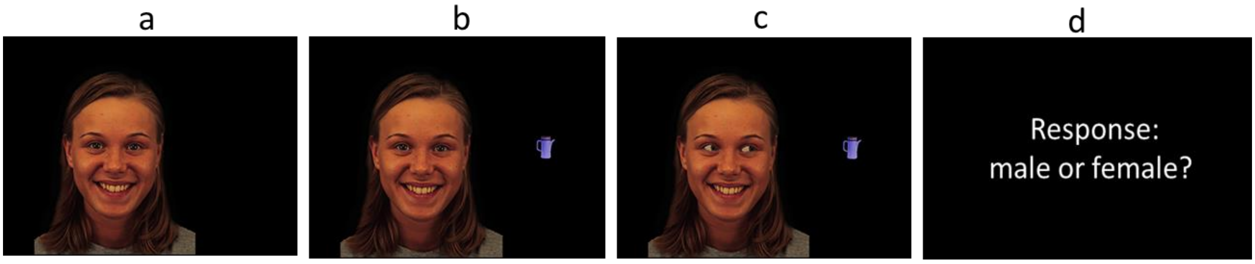

Figure 16. Trial sequences and examples of stimuli for the gaze leading task. (a) Participants fixated the face displayed with direct gaze. (b) After 2000 ms of good fixation on face was detected, an object appeared. Participants saccaded to the object as soon as it appeared. (c) After 300 ms of good fixation on the object was detected, gaze onset occurred, either toward the object (joint attention) or averted (non-joint attention). Participants saccaded back to the face as soon as they noticed the gaze shift. Gaze was displayed for 3000 ms. (d) The stimuli cleared and the participant prompted to identify the face as male or female, displayed until response.

Apparatus and materials. Right eye position was tracked with an infrared eye tracker (Eyelink 1000, SR Research, Ontario, Canada; spatial resolution of 0.1^0, 500 Hz). To maintain head stability, a head and chin rest was used. Viewing distance was approximately 70 cm from eyes to an 18” computer monitor (resolution 1024 × 768 pixels). A standard keyboard was used for participants’ manual responses to the gaze leading task and a ‘Black Box’ four-button response box was used for participant’s finger press responses to the old/new face recognition task. The Autism Spectrum Quotient Questionnaire was used as a measure of the level of autism-like traits (AQ; Baron-Cohen et al., 2001), presented using E Prime. Participants rated how strongly they agreed or disagreed with each item (e.g. I prefer to do things the same way over and over again) on a four point Likert scale ranging from definitely agree to definitely disagree. The Cambridge Face Memory Test (CFMT: Duchaine & Nakayama, 2006) was used as a standardized assessment of participant face recognition ability. The CFMT presents six unfamiliar male faces from three different views and then tests recognition in a three-alternative forced-choice task.

Design. The within-subjects condition had two levels: congruent response (joint attention face) and averted/incongruent gaze response (non-joint attention face). The dependent variables were the number of correctly recalled hits for faces recognised, the level of confidence about the recollection (sure or not sure), reaction time (RT) to categorise the gender of the faces (gaze leading task) and RT to report whether they recollected seeing the face before (old/new face recognition task). There were also two correlational designs to examine any associations between recall and RT for the gaze-led faces and, firstly, level of autism-like traits and, secondly, scores on the CFMT (Duchaine & Nakayama, 2006).

Procedure. Participants were positioned comfortably in the chin rest for a short practice block of the gaze leading task (6 trials) and then for the gaze leading block of 40 trials. The eye tracker was calibrated to each participant to enhance tracking accuracy using a standard nine point calibration at the start of the gaze leading, eye-tracked block. An example trial sequence for the gaze leading task is illustrated at Figure 16. Each face was presented with direct gaze until 2000 ms of fixation on the face was detected.

Next, an object appeared to the right of the face with participants being instructed to immediately perform a saccade to the object. Upon detection of 300 ms of fixation on the object, the onscreen face responded either with a congruent gaze shift towards the object (joint attention condition) or an averted gaze shift away from the object (non-joint attention condition). This was a gaze contingent design (unlike Experiments 1-5), meaning the fixation on the object triggered the onscreen gaze shift. The time between the object appearing and the onscreen gaze shift, therefore, comprised of the saccade latency (approximately 200 ms based on mean saccadic reaction times in Experiments 1-4), the saccade duration (approximately 80 ms based on the mean saccade duration in Experiments 1-4) and the 300 ms required for fixation on the object to be detected and any uptake time by the eye tracker to detect fixation. Therefore, the total time between object appearance and gaze shift would be around 580 ms. Participants were instructed to saccade back to the face as soon as the onscreen face's gaze shifted. Gaze onset was displayed for 3000 ms before the stimuli cleared and participants were prompted to press the 'f' key with their left hand to report the gender if the face was female and the 'm' key with their right hand for male, displayed until response. There was an inter-trial interval of a blank screen for 1000 ms before the next trial began. Participants were told about the gaze contingency, that is, that the onscreen face shifted gaze once the eye tracker detected their fixation on the object. Participants saw 20 faces (10 males, 10 females) who followed and 20 faces (10 males, 10 females) who never followed their gaze. The faces were counterbalanced four ways with 20 faces in each counterbalanced set. The stimulus set of objects for all trials was presented randomly.

Following completion of the gaze leading task, participants completed a surprise old/new face recognition task in which they were asked to recollect whether they had seen (old) or had not seen (new) the face in the previous task, being shown 40 faces (20 males, 20 females) they had seen and 40 faces (20 males, 20 females) they had not seen (foils), presented randomly. Each face was displayed centrally for 4000 ms following the presentation of a central fixation cross (white font, Courier, 18pt) for 1000 ms. Next, participants were prompted to use the four button response box and their right hand to report whether the face was old or new, using their thumb to report old and their little finger to report new. Following their old/new response, they were asked to rate their level of confidence in that recollection, using their thumb to report they were sure they had seen the face before and their little finger to report they were sure they had not seen the face before. They used their index finger for faces they thought they had seen but were not sure, and their ring finger for faces they thought they had not seen, but were not sure. Participants were instructed to respond as quickly, but as accurately, as possible and were given no feedback about their responses for either the gaze leading or the face recognition tasks. An inter trial interval (a blank screen) was displayed for 1000 ms after each old/new response. Participants were naïve to the hypotheses. After the face recognition task, participants completed the AQ followed by the CFMT on the computer.

Experiment 6 Results

Data processing. Trials in which participant RTs were +/- 3 SDs from their individual means were removed (1.71% of trials for RT for gender in the gaze leading task and 2.86% of trials for old/new RT). Confidence intervals around means and confidence intervals around effect sizes are reported, calculated using ESCI (Exploratory Software for Confidence Intervals; Cumming & Calin-Jageman, 2016).

Recognition. Overall accuracy in the face recognition task was 70%. Mean correct identification of old faces (hits) out of 20 for joint attention faces was 12.77, 95% CI [11.74, 13.77] (SD=2.91), and for non-joint attention faces was 13.11, 95% CI [11.84, 14.39] (*SD*=3.70). Overall accuracy for the correctly rejected foils was 75% with 30.01 (*SD*= 7.24) mean correct rejections out of 40. There was no significant difference between hits for joint attention and non-joint attention faces, *t*(34)= 0.737, *p*= .466, d_z=0.10 95% CI [-0.17, 0.28]. There was no significant difference between mean number of high confidence hits for joint attention faces, 7.66, 95% CI [6.59, 8.73] (*SD*=3.11), and for high confidence non-joint attention faces, 7.97 (*SD*=3.92), *t*(34)=0.486, *p*=.630, $\mathrm{d_z}$=0.09, 95% CI [-0.27, 0.45].

Reaction time for gaze leading task. Mean RT to report the gender of joint attention faces was 645 ms, 95% CI [580, 710] (*SD*=189) and 658 ms, 95% CI [591, 725] (SD=196) for non-joint attention faces, and there was no significant difference between them, *t*(34)= 0.652, *p*= .519, $\mathrm{d_z}$=0.07, 95% CI [-0.13, 0.26].

Reaction time for face recognition task. Mean RT to report old/new was 650 ms, 95% CI [568, 732] (*SD*=239) collapsed across all conditions, 665 ms, 95% CI [580, 750]

(*SD*=247) for foils, 657 ms, 95% CI [556, 758] (*SD*=294) for joint attention faces and 628 ms, 95% CI [558, 698] (*SD*=204) for non-joint attention faces (see Figure 17). An ANOVA (face condition: foils, JA faces, non-JA faces) showed no effect of face condition for RT to recall whether a face had been seen before in the gaze leading task, $F(2, 68)= 1.471$, $p=.237$, $\eta_p^2 =.041$.

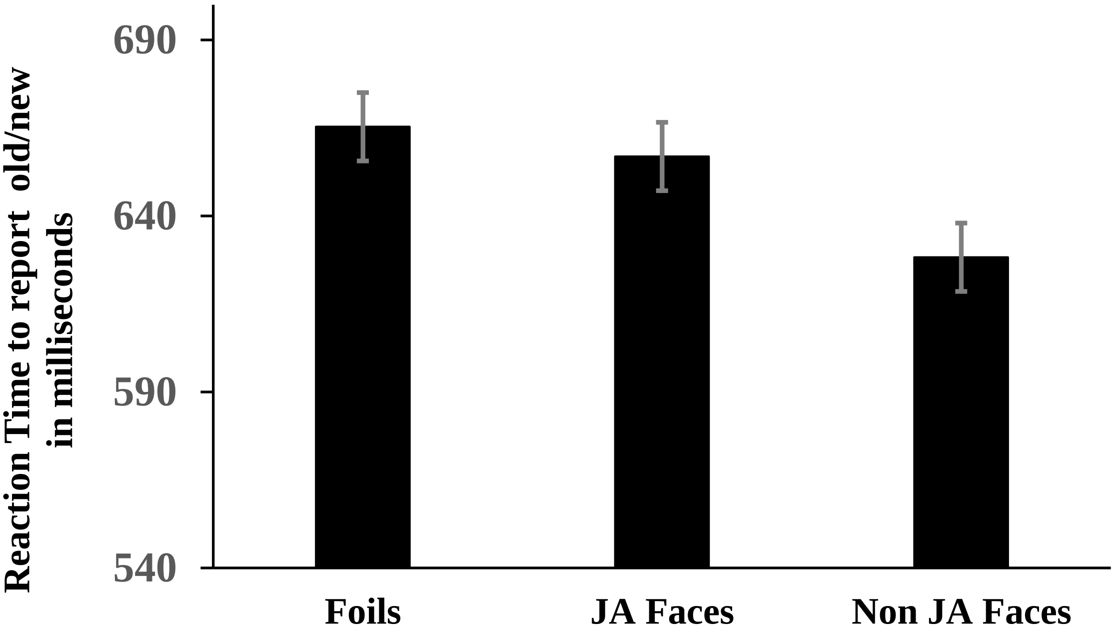

Figure 17. Mean Reaction Times to report old/new for foils, joint attention (JA) and non-joint attention (non JA) faces in face recognition task for Experiment 6. Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus & Masson (1994).

AQ correlational analyses. Mean AQ score was 15.60 (*SD*=8.53), which is about normative for social science university students (see Baron-Cohen et al., 2001, who reported a social science student mean of 16). The range was 3 to 39. There were no correlations between total AQ scores and accuracy overall, number of joint attention or non joint attention hits (whether with high or low confidence), or the difference between them, rs<0.167, ps>.338. Analyses were carried out on the AQ sub-scales and can be found in Appendix C.

CFMT correlational analyses. Mean CFMT performance was 75% (*SD*=12) accuracy, which is a little below the normative score of 80% (*SD*=11) (Duchaine & Nakayama, 2006). No correlations between any variables of interest in any condition (RT for gender, RT for old/new, recollection accuracy or AQ) and performance on the CFMT were found, rs<0.21, ps>.22.

Experiment 6 Discussion

This study shows that the congruency of gaze shift following gaze leading (averted gaze or joint attention) appears to have no consequences for how well those faces are subsequently recognised, nor for how quickly basic information about a face (in the form of gender classification here) is reported. This may be because there are simply no effects upon how faces are encoded in memory by gaze response, or it may be because this particular experimental design limited participants to encountering each face just once, and that may not be enough to lead to differential encoding. Alternatively, this null finding may offer evidence that it is equally important to remember all faces, whatever their social response to gaze leading. This will be discussed further in the Chapter Discussion.

There were no correlations found between autism-like traits (AQ) and subsequent face recognition performance following gaze leading, and no correlations with any variables of interest and face recognition ability, measured by scores on the CFMT. There was a medium correlation between total score for autism-like traits and RT to report the gender of the face. Those with higher AQ were slower to report gender after averted gaze responses. The next experiment, Experiment 7, being higher powered, was designed to enable a more reliable examination of these preliminary, promising finding of individual differences, alongside a change to how often the faces were encountered.

Experiment 7

Having found that recognition of faces was unaffected by responses to gaze leading, it was possible that only presenting each of the 40 faces in the gaze leading task once was insufficient to enable potential effects to emerge. In addition, the lack of correlations with CFMT and task performance may require greater power afforded by an increase in sample size in the next experiment and enable further exploration of the significant AQ correlation in Experiment 6. Therefore, two changes to the design of Experiment 6 were made. Firstly, participants were presented with the same faces from Experiment 6 twice, instead of once, doubling the number of trials in the gaze leading task. Secondly, 59 new participants were recruited for this study. One participant did not complete the tasks due to difficulties with eye-tracking, hence *n*=58 (mean age 19.93, *SD* = 2.54; 10 were male). Other than those changes, the procedure was identical in Experiment 7 to Experiment 6.

Experiment 7 Results

Data processing. Trials in which participant RTs were +/- 3 SDs from their individual means were removed (1.46% of trials for RT for gender in the gaze leading task and 2.16% of trials for old/new RT). As for Experiment 6, confidence intervals around means and confidence intervals around effect sizes are reported.

Recognition. Overall accuracy in the face recognition task was 81%, an increase of 11% from Experiment 1, which is expected because participants encountered each face twice, rather than just once. Mean correct identification of old faces (hits) out of 20 for joint attention faces was 15.40, 95% CI [14.69, 16.10] (*SD*=2.67), and for non-joint attention faces was 15.10, 95% CI [14.40, 15.81] (*SD*=2.67). Overall accuracy for the correctly rejected foils was 88% with 35.38 (*SD*= 3.01) mean correct rejections out of 40. There was no significant difference between hits for joint attention and non-joint attention faces, *t*(57)= 0.794, *p*= .430, d_z=0.11 95% CI [-0.38, 0.16]. There was no significant difference between mean number of high confidence hits for joint attention, mean 11.24 hits 95% CI [10.28, 12.21] (*SD*=3.67), and for high confidence non-joint attention, mean 11.43 hits 95% CI [10.60, 12.26] (*SD*=3.15), *t*(57)=0.392, *p*=0.696, d_z=0.06, 95% CI [-0.22, 0.33].

Reaction time for gaze leading task. Mean RT to report the gender of joint attention faces was 588 ms, 95% CI [542, 633] (*SD*=174) and 577ms, 95% CI [533, 621] (*SD*=167) for non-joint attention faces, with no significant difference between them, *t*(57)= 1.611, *p*= .113, d_z=-0.10, 95% CI [-0.22, 0.02].

Reaction time for face recognition task. Mean RT to report old/new was 486ms, 95% CI [447, 525] (*SD*=148) collapsed across all conditions, 493ms, 95% CI [450, 535] (*SD*=161) for foils, 494 ms, 95% CI [450, 539] (*SD*=169) for joint attention faces and 464 ms, 95% CI [427, 501] (*SD*=141) for non-joint attention faces. These times are much faster than Experiment 6 RTs. This is expected, given participants saw each face twice, rather than once. Mean RTs are shown in Figure 18.

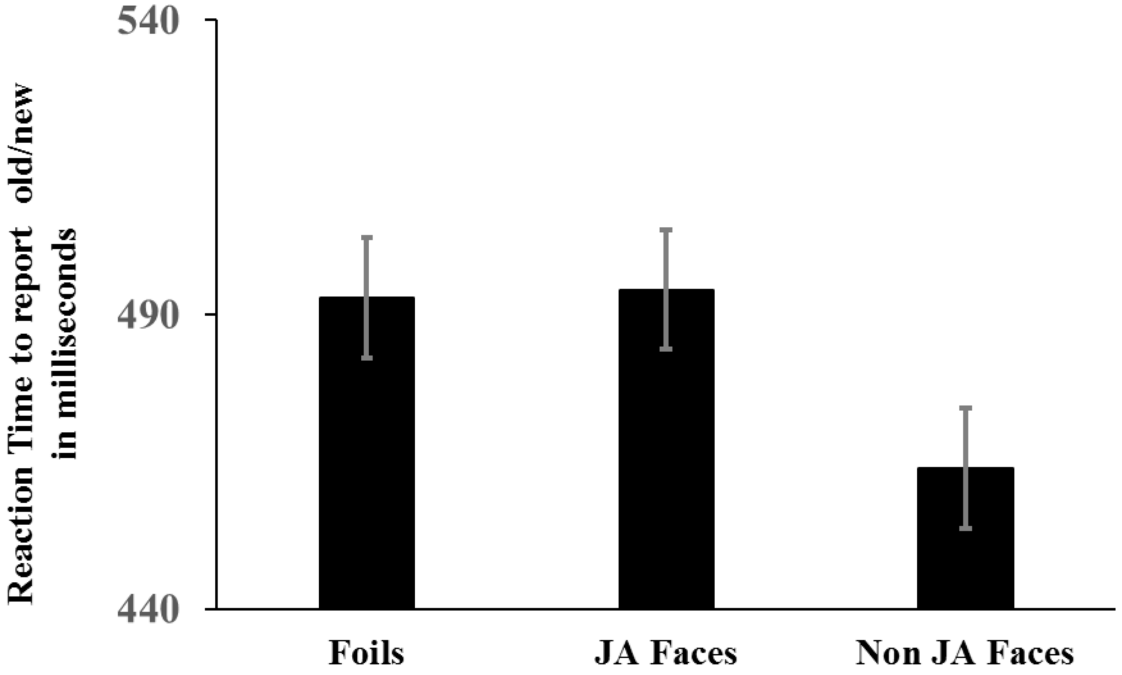

Figure 18. Mean Reaction Times to report old/new for foils, joint attention (JA) and non-joint attention (non JA) faces in face recognition task for Experiment 7. Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus & Masson (1994).

An ANOVA (face condition: foils, JA faces, non-JA faces) showed no effect of face condition for RT to recall whether a face had been seen before in the gaze leading task, $F(2,114)= 2.961$, $p=.056$, $\eta_p^2 =.049$. It is notable, however, that RT to JA faces was almost exactly the same as RT to foils. Significant differences between RT to JA and non-JA faces were checked by analysing separately correct or incorrect responses, but the differences were not reliable in either case, *t*s<1.54, *p*s>.07. For completeness, given the pattern of the RT data, the datasets from Experiment 6 and 7 were combined, and an ANOVA performed (face condition: foils, JA faces, non-JA faces) which showed an effect of face condition on RT, $F(1.87, 171.97)= 4.301$, $p=.017$, $\eta_p^2 =.045$, Greenhaus-Geisser-corrected, with a small to medium effect. A follow up contrast showed that RT to JA faces were slower (*M*=555 ms *SD*= 236) 95% CI [506,604] than NJA faces (*M*=525 ms, *SD*=185) 95% CI [488,564], $t(92)=2.189$, $p=.031$, $d_z=-0.14$ [-0.27, -0.01], but this was only a small effect. In contrast, reaction time to foils (*M*=558 ms, *SD*=214) was strikingly similar to those for joint attention faces (*M*=555 ms *SD*= 236). Figure 19 shows the mean RT collapsed across both experiments.

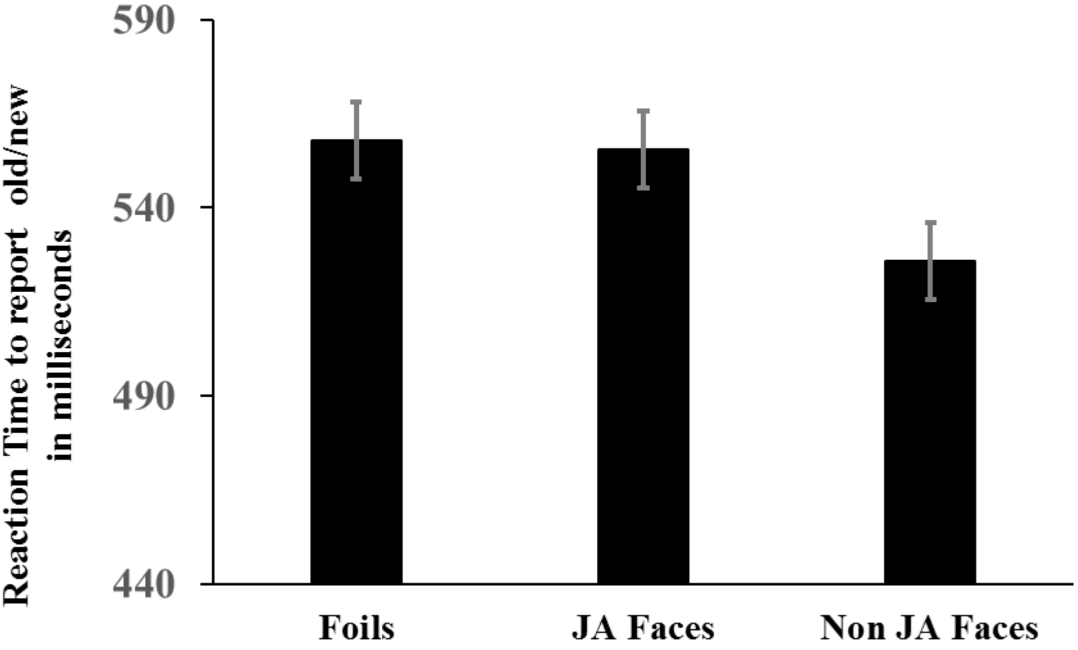

Figure 19. Mean Reaction Times to report old/new for foils, joint attention (JA) and non-joint attention (non JA) faces in face recognition task collapsed across Experiment 6 and Experiment 7. Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus & Masson (1994).

AQ Correlational analyses. Mean AQ score was 14.21 (*SD*=5.90) and the range was 5 to 30. There were no correlations between total AQ scores and accuracy overall, number of joint attention or non joint attention hits (whether with high or low confidence), or the difference between them, *r*s<0.174, *p*s>.192. There were no correlations between AQ and RT to report old/new in the face recognition task, whether to foils, joint or non-joint attention faces or difference between joint and non-joint attention faces, *r*s(58)<0.193, *p*s>.148. A correlational analysis of the AQ sub-scales can be found at Appendix C.

CFMT correlational analyses. Mean CFMT performance was 79.91% (*SD*=11.45) accuracy, which is normative (Duchaine & Nakayama, 2006). No correlations between any variables of interest in any condition (RT for gender, RT for old/new, number of hits or AQ) and performance on the CFMT were found, *r*s<0.242, *p*s>.067. However, there was a positive correlation between CFMT scores and accuracy overall in the old/new face recognition task, *r*(58)=0.272, *p*=.039. The higher the CFMT score, the higher the accuracy score for recognition. This which would be expected as the recognition task should reveal individual differences which correlate with performance on the CFMT, an established measure for this ability.

Experiment 7 Discussion

Experiment 7 confirmed, like Experiment 6, that recognition of faces is not modulated by previous gaze leading responses. Participants were just as accurate at remembering whether they had seen a face, whether or not the face had responded with joint attention or with averted gaze. Reaction times to report gender of the faces and in the old-new task were not modulated by gaze response either. However, there was a pattern of slower reaction times for the old-new task to recollect joint attention faces over averted gaze faces. This did not reach statistical significance, however. These findings will be discussed further in the Chapter Discussion. The AQ correlation found in Experiment 6 did not replicate in Experiment 7. The lack of consistency across experiments in these correlational findings makes it hard to draw any reliable inferences from them.

Chapter Discussion

Across two experiments, it has been shown that the gaze response to gaze leading has no effect upon subsequent recognition of faces. This may be because it is equally important to know whether we have encountered a face before, whatever their responses to our own gaze behaviours. The ability to recognise new faces better than we can recognise other novel stimuli has been a robust finding in social cognition for years (see seminal paper by Ellis, 1975). This ability speaks to the fundamental importance of face recognition in human relationships, although recently an important distinction has been made between our apparently automatic face expertise for familiar faces, compared to unfamiliar faces, where there are far greater individual differences (Young & Burton, 2017a). Whilst gaze leading has been shown to influence subsequent recall of stimuli which are the object of the joint attention bid (Kim & Mundy, 2012), this may not be the same for recognition of the faces we have gaze led.

The findings of no effect upon old-new recognition is consistent with Bell, Buchner, and Musch (2010). Bell et al. (2010) employed a trust game with cooperators and defectors and found enhanced old-new recognition for the faces of *both* groups of people. Similarly, the fairly high rates of old-new accuracy in both experiments (70% for Experiment 6 when faces were encountered once and 81% for Experiment 7 when faces were encountered twice) for gaze followers (who are co-operators) and non-followers (who are the defectors/norm violators here), could show the equal importance of encoding both types of people in memory.

The pattern of slower reaction times in the old-new task for joint attention faces over averted gaze faces is worthy of brief discussion, albeit with the full acknowledgement that this was not a statistically significant effect in either Experiment 7 or 8 alone, only reaching significance across both experiments combined, and with a very small effect size. Furthermore, no corrections have been made for multiple comparisons, meaning the small, significant effect needs to be treated with caution. It is also interesting that reaction times to foils were just as slow as those to joint attention faces. If such an effect could be more reliably demonstrated, then this would demonstrate perhaps better encoding in memory, evidenced by faster reaction times, for norm violators (averted gaze) over the co-operators (joint attention). This would not be consistent with Bell et al. (2010) where both violators and cooperators are remembered equally well in a trust game, but it is a difference context to this design so cannot necessarily be expected to produce the same outcome. Future research should seek to test this further.

Limitations and Future Directions

Perhaps presenting the faces just once (Experiment 6) or twice (Experiment 7) for a short time was not enough to reveal memory encoding differences. Alternatively, the paradigm may have been too subtle in asking participants to report the gender of the faces, which ensured they paid attention to the face, but was not very effortful. Maybe if participants had been asked to report the eye colour of the faces, this would increase the saliency of the eyes and have been a better task to precede the surprise memory test. Another change could be to deploy an n-back task (Kirchner, 1958) which would allow participants greater opportunity to encode the faces through repeated, sequential exposure, and test any effects on working memory. Perhaps the time for recall following straight after the encoding task does not lead to recall differences, but there may be longer-term memory effects. Therefore, postponing the task until the next day or longer might reveal differences in recall of followers and non-followers. Finally, these studies show high accuracy for recognition of all unfamiliar faces following gaze leading (70% for Experiment 6 when faces were gaze led just once and 81% for Experiment 7 when faces were gaze led twice). It may be that the act of a gaze leading interaction itself with those faces caused such high accuracy in the old-new task, which this design has not explored. Therefore, an unanswered question is: would lower accuracy result from simply being presented with the faces once or twice, with no gaze leading? To explore this possibility, a future study could add a control condition in which no gaze leading interaction took place.

Conclusions

These two experiments investigated whether congruency of gaze response following gaze leading to a referent object modulated subsequent recognition of those faces. There were no such effects. It may simply be that this paradigm has failed to tap into this phenomenon and more encounters than one or two might lead to effects. However, it may be that there is never any difference in face encoding according to gaze response because it is equally important to recognise a new face, whatever gaze behaviours have been encountered. Future work should test whether faces presented for the same amount of time, but not gaze led, are remembered as well as gaze-led faces.

Chapter 5: Follow my Lead: Event-related Potentials Elicited by Responses to Joint Attention Initiation

Having established behavioural consequences of gaze leading, this thesis aimed to examine the neural underpinnings of joint attention initiation. Saccades are one of the fastest actions humans can perform. They are are ballistic and generated and executed at sub-second speeds. For example, a ten degree of visual angle saccade ranges in velocity between 420 to 520 degrees per second (Bahill & Stark, 1979). Therefore, I took advantage of the high and continuous temporal resolution afforded by electroencephalography (EEG) to investigate the event-related potentials (ERPs) elicited by gaze responses to gaze leading. The aim was specifically to explore whether the N170 ERP component showed any differences when gaze leading was reciprocated with a congruent gaze shift, compared with an incongruent/averted gaze shift response. In other words, can I establish evidence that the neural system rapidly detects the outcome of a successful joint attention bid?

Over the past eight years, there have been a growing number of *f*MRI papers which have identified there are some distinct and some overlapping regions when initiating joint attention, compared to responding to joint attention (e.g. Gordon, Eilbott, Feldman, Pelphrey, & Vander Wyk, 2013; Oberwelland et al., 2016; Redcay, Kleiner, & Saxe, 2012; Schilbach et al., 2010). Mundy (2017) offers a thorough review of these findings and the brain regions associated with joint attention. The neuro-imaging work has implicated the posterior Superior Temporal Sulcus and the Medial Pre-frontal Cortex as the main regions involved in processing joint attention (see Mundy, 2017, for a review).

Schilbach et al. (2010) found participants rated gaze interactions as more pleasant when they engaged in gaze leading, compared to responding to other's gaze leading, providing self-report evidence of the rewarding nature of gaze leading. The *f*MRI research has also implicated the putative reward system is involved in successful gaze leading with the ventral striatum shown to be activated when successfully leading a social partner's gaze to an object (Schilbach et al., 2010). Similarly, Gordon et al. (2013) found increased activation in the subthalamic regions of the striatum and the ventral tegmental area for congruent gaze shifts in response to joint attention initiation, but not for incongruent responses. We may be motivated, therefore, to cause others to align their locus of regard with our own because of its rewarding nature in addition to the mutual cooperation and communication it can facilitate.

Gaze following (responding to joint attention) and eye contact have also been the subject of many ERP studies (see, for example, Itier & Batty, 2009; Myllyneva & Hietanen, 2016; Senju, Johnson, & Csibra, 2012.; Tipples, Johnston, & Mayes, 2013). Only two recent papers have examined gaze *leading* associated event-related potentials (Caruana, de Lissa, & McArthur, 2015, 2017). The experiment in this chapter aimed to follow these findings and ascertain in an exploratory fashion whether, at an electrophysiological level, there is a differentiation between gaze being followed and not being followed within a simple, simulated gaze interaction. Another aim was to examine whether autism-like traits were associated with any effects.

This review will begin by reviewing the N170 gaze research, as this is the main candidate for investigating ERPs elicited when interacting with faces, and then, to provide some further ERP context, evidence will be reviewed for the involvement of the following components in gaze processing: P3, N300, and N2pc. Then, there will be a brief review of the atypicalities found for gaze leading amongst those with an autism diagnosis.

N170

The N170 event-related potential (ERP) component is a negative-going evoked potential associated with face processing, emerging over parietal-occipital scalp sites around 170ms after face stimulus onset (e.g. Bentin, Allison, Puce, Perez, & McCarthy, 1996). The N170 is usually associated with face perception and thought to be face-sensitive as it is generally greater for faces than other objects (see Eimer, 2011; Rossion, 2014, for reviews). The N170 has also, however, been studied specifically in relation to eye gaze which has produced a set of conflicting results. Some studies have found greater N170 amplitude for averted gaze over direct gaze (e.g. Itier, Alain, Sedore, & McIntosh, 2007; Latinus et al., 2014; Watanabe, Miki, & Kakigi, 2002), whilst others have found no differences (e.g. Itier & Batty, 2009; Myllyneva & Hietanen, 2016; Schweinberger, Kloth, & Jenkins, 2007). Others still have reported greater N170 for direct over averted gaze (Conty et al., 2007). Greater N170 for direct over averted gaze was also reported by Pönkänen, Alhoniemi, Leppänen, and Hietanen (2011), but only when participants viewed a live face, rather than a face on a computer. Carrick, Thompson, Epling, and Puce (2007) found no modulation of the N170 when viewing different types of gaze averted faces. Similarly, Caruana et al. (2015) found no modulation of the N170 by congruency of gaze responses to gaze leading.

This mixed array of findings makes it difficult to predict in which direction gaze response to gaze leading may modulate the N170. However, the literature does provide evidence that eye gaze is processed during face processing and can modulate the N170 in some contexts. Finally, as the present study is an investigation of shared attention initiation, requiring the evoking stimulus to be presented in the peripheral visual field, it is important to note that the N170 to faces can be delayed when presented in periphery (e,g. Rigoulot et al., 2011). The N170 is a strong candidate component to be sensitive to not just observed averted gaze, but the social context in which it is presented.

P3

The P3 ERP component has often been shown to reflect discrimination of less frequent targets from more frequent targets, using the oddball paradigm (see Polich, 2007, for a review). However, the P3 has been the subject of a diverse body of other research and found to be modulated by the cognitive demands of the task at hand (e.g. Kok, 2001) and to be associated with memory recall function (see Polich, 2007, for a review).

Caruana et al. (2015) found a greater and later P3 (a more specific time-locked waveform, the “P350” was reported) was elicited when gaze leading elicited an incongruent gaze response from a virtual computer-generated face, compared to a congruent gaze response within the context of a virtual game. Caruana et al. (2015) also found that when arrows provided the directional shifts in response to gaze leading, no P350 congruency effect resulted. Taken together, the authors suggested that the P350 reflected an evaluation of the social significance of the gaze response. Caruana et al. (2017) went on to compare P350 and N170 elicited when onscreen gaze responses to gaze leading were thought by participants to be controlled by a human, compared with being controlled by a computer. Participants who were told a human controlled the eye movements (using the same dataset as Caruana et al., 2015), demonstrated a larger P350 for incongruent gaze shifts compared to congruent. There was no such difference for those who were told a computer was generating the gaze shift responses. There was a smaller N170 for those who thought a computer, rather than a human, was in control but no N170 differences according to gaze congruency. Caruana et al. (2017) suggested the findings of P350 modulation in the human attribution condition may reflect the activation of mentalising processes not needed during non-human interactions.

N300

Firstly, note the important distinction that the following reviewed studies exploring the N300 component examine participants being gaze cued, rather than performing the gaze cueing (gaze leading), as participants do in the current study. Greater occipito-temporal negativity (ERP component reported as “N330”) has been demonstrated in response to incongruent gaze shifts away from an object than congruent (Senju et al., 2006). The suggested explanation was that the N330 reflected the greater effort required to process the violation of the expectancy that gaze would be shifted to an object. In addition, the N330 was believed to reflect activity in the posterior Superior Temporal Sulcus. Tipples et al. (2013) also found an enhanced negative occipito-temporal ERP for incongruent gaze shifts (occurring slightly earlier at N300). Furthermore, Tipples et al. (2013) found an enhanced N300 when arrows provided the directional shifts of attention, and so suggested there is a domain general mechanism for detecting and processing unexpected events.

N2pc

The role of the N2 posterior contralateral (N2pc) is not clearly established within social gaze, only being shown to my knowledge in one study to date involving eye gaze specifically (Galfano et al., 2011), as it is more commonly investigated within visual search paradigms (e.g. Grubert & Eimer, 2015), and thought to reflect attentional processing (Luck, 2014). The N2pc component is calculated by subtracting ERPs at the sites ipsilateral to the target stimulus, from contralateral ERPs (Grubert & Eimer, 2015). This ERP component comprises greater negative activity at the posterior sites contralateral to the side on which the stimuli are presented, implicated in spatial attentional shifting (Galfano et al., 2011). Galfano et al. (2011) used the N2pc as an index of spatial attention reorientation to the target required when incongruent gaze cueing occurred. Galfano et al. (2011) predicted, and found, a greater N2pc elicited from incongruent gaze cueing than congruent.

To summarise, ERPs P3, N300, N170 and N2pc have all been implicated in gaze processing, but there have been some mixed findings and most of the work has investigated observing averted or direct gaze, rather than gaze leading. What little evidence there is for gaze leading is that the N170 is not modulated by gaze response, but that the P350 can be, but not if participants are told a computer, rather than a human, is generating the gaze response (Caruana et al., 2015; Caruana et al., 2017).

Autism-like Traits

Initiating joint attention has been identified as a specific deficit in those who have a diagnosis of autism with greater deficits usually found for initiating joint attention, over responding to joint attention (see Mundy, 2017; Mundy & Newell, 2007, for reviews). This has generated a growing field of research into interventions to improve this social skill with a recent meta-analysis concluding that these interventions can be effective (Murza et al., 2016). Because of the sub-optimality of joint attention processes in the clinical population, participants' level of autism-like traits were measured to explore any correlations with the

ERP effects in the broader phenotype. This was measured using participant's self-reported autism-like traits (AQ) using the Autism Spectrum Quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001).

It is possible that ERPs vary according to level of autism-like traits, for example, the N170 has been reported as occurring later (e.g. Hileman, Henderson, Mundy, Newell, & Jaime, 2011) or being smaller (Churches, Damiano, Baron-Cohen, & Ring, 2012) in those with a diagnosis of autism, compared to controls. However, there have been mixed findings and a recent systematic review found no consistent differential effects on the N170 for those with autism (Feuerriegel, Churches, Hofmann, & Keage, 2015).

Experiment 8

Current Study Aims

This study aimed to find out whether the N170 is modulated according to gaze response following a simple gaze leading saccade from a face to an object. This would provide evidence of the detection of the response elicited by gaze leading and the time course of this. No information was offered to participants about whether a human or the computer controlled the responses, simply presenting the faces and explaining their gaze would either be followed or not followed by the onscreen faces. In addition, participants performed a gender identification task, so their task was orthogonal to the gaze interaction. This means that any effects which emerge are unlikely to be driven by higher order processing judgments about the gaze interaction.

Experiment 8 Method

Participants. After ethical approval was obtained from a local ethics committee, 36 Psychology undergraduate student participants (mean age = 19.9 years, *SD* = 1.33; 7 males), gave written, informed consent and were granted course credits in return for participation. All participants reported normal or corrected-to-normal vision and no history of neurological disorder. Participants were positioned comfortably in a chin rest and had rest breaks between four blocks of trials. One participant was excluded from analysis because the EEG signal was poor in the regions of interest. Henceforth, *n*=35. An a priori power analysis anticipating a medium effect size (based on mean P3 differences reported by Caruana et al., 2017) with a power of 0.80 (Cohen, 1988) calculated using G*Power3, (Faul et al., 2007) produced a required sample of 34. However, testing was stopped at the end of a scheduled block of testing for convenience with 36 participants tested.

Stimuli. Images of six smiling faces, three male and three female, (280 x 374 pixels) were taken from the NimStim face set (see Figure 20; Tottenham et al., 2009). Smiling faces were used to make the task a little more interesting for participants whose sole job was to report gender. Each face comprised of three versions, with eyes looking right, looking left or straight ahead. There were two further NimStim faces (one female and one male) used for the practice block. There were eight images of everyday kitchen objects (220 x 78 pixels) taken from Bayliss, Frischen, Fenske, and Tipper (2007). All stimuli were presented on a black background using E Prime 2.0.

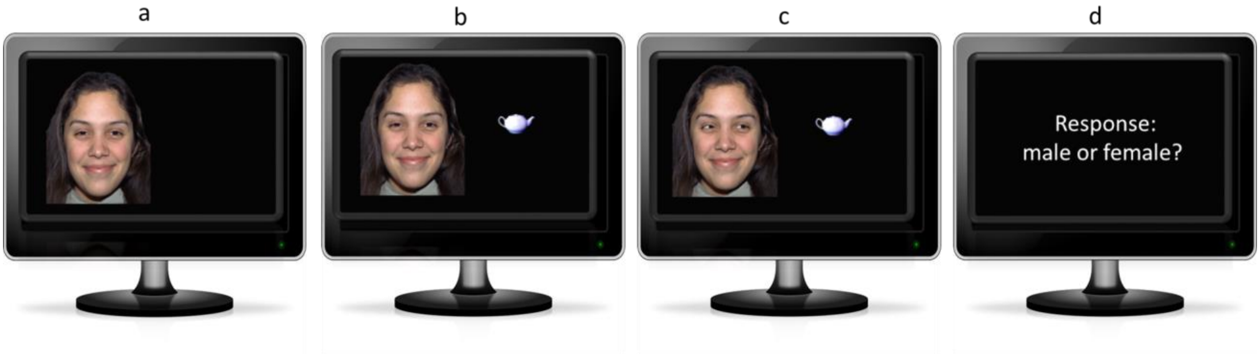

Figure 20. Trial sequences and examples of stimuli. a) The participant fixates an onscreen face, displayed for 1000 ms. b) An object appears and the participant immediately saccades to the object. c) After 800 ms, the onscreen face responds with either a congruent or incongruent gaze shift, displayed for 2500 ms. d) The participant is prompted to report the gender of the face they just saw (displayed until response).

Apparatus and materials. A 64-channel active electrode system (Brain Products GMbH) with a cap (BrainCap-64 channels) and an amplifier (BrainAmp MR 64 PLUS) was used for EEG acquisition. Viewing distance was approximately 70 cm from eyes to a 24″ computer monitor (resolution 1920 × 1080 pixels). A standard keyboard was used for participants' manual responses. The Autism Spectrum Quotient Questionnaire (see Appendix A) was used as a measure of the level of autism-like traits (Baron-Cohen et al. 2001), presented using E Prime. Participants rated how strongly they agreed or disagreed with each item (e.g. I prefer to do things the same way over and over again) on a four point Likert scale ranging from definitely agree to definitely disagree.

Procedure. Participants were positioned in a comfortable chair in front of a computer screen 70 cm from their face. Participants completed a practice "gaze leading" block of six trials followed by four gaze leading blocks of 60 trials each where six faces, presented randomly, followed gaze leading 50 percent of the time and did not follow gaze leading (averted gaze) the other 50 per cent of the time. Therefore, there were a total of 120 trials per condition. In two of the gaze leading blocks, the faces were presented 2.5 cm to the left of the centre of the screen with the object appearing 13.5 cm to the right of the faces. In the other two blocks, faces were presented 2.5 cm to the right of centre, with the object appearing 13.5 cm to the left of the faces. Block order was counterbalanced. Finally, participants completed the Autism Quotient Questionnaire on the computer. Participants were given rest breaks for as long as they needed between each block.

In all the gaze leading blocks the face was presented looking straight ahead (so with direct gaze towards the participants) and this was displayed for 1000 ms whilst the participants were instructed to fixate on the face. Next, the object appeared to the right or left of the face positioned in line with the line of gaze of the onscreen face and 13.5 cm from the bridge of the nose of the face to the centre of the object. Participants were instructed to saccade to the object as soon as it appeared and keep fixating on the object until they noticed the gaze shift had occurred in their peripheral vision. After 800 ms, gaze onset occurred either to follow (joint attention) or not follow (averted gaze) the participant’s gaze towards the object. The 800 ms time frame between the object appearing and the gaze onset was informed by previous work on how long a time interval feels naturalistic between gaze leading and response (Pfeiffer et al., 2012) and a small pilot study in which participants were asked to rate which of four time durations felt most naturalistic from a choice of 400, 800, 1000 and 1400 ms, using the same stimuli set up as the subsequent experiment. Once the gaze response had been displayed for 2500 ms, the face and object cleared to reveal a prompt to report whether the face was male or female. Participants used the index finger of their right hand to press letter key “m” for male and the index finger of their left hand to report “f” for female. There was an inter-trial interval of a blank screen jittered with a random interval of 750-1250 ms following participant response and before the next face appeared to start the next trial. Participants performed a gender identification task in order to ensure the task was orthogonal to any judgments about whether their gaze was followed or not (see Figure 20).

Participants were shown examples of their typical EEG artifacts on a monitor in the testing room, including their horizontal saccades. They were informed that, therefore, experimenters would be monitoring their eye movements using the EEG signal throughout the experiment.

Data acquisition. Accuracy and reaction time for the gender identification task was recorded for every trial. EEG was recorded using a Brain Vision actiCAP system with 63 active electrodes. Participants wore an elastic nylon cap (10/10 system extended). A further electrode was placed under the left eye to monitor horizontal eye movements (EOG). The continuous EEG signal was recorded at a 500 Hz sampling rate using FCz as a reference electrode. All electrodes had connection impedance below 50 kΩ before recording commenced.

Continuous EEG data were pre-processed and analysed offline using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014). High and low pass half-amplitude cutoffs were set at 0.1 and 40 Hz, respectively. Any bad channels were interpolated with the spherical interpolation function from EEGLAB. Artifacts were removed in two stages. Firstly, trials containing excessive artifacts were rejected by manual inspection (2.4 per cent of trials). Secondly, horizontal eye movements and blinks were identified using the “runica” ERPLAB function for independent component analysis (ICA). Two procedures helped inform the components selected for removal for each participant. Firstly, the scalp maps for all components were examined to identify those which were eye artifacts (both blinks and saccades). Secondly, the maximal contribution to ERPs was assessed during the timeframe within which the saccades could occur (from -200 ms before gaze onset to 4300 ms after), and those eye artifact components which contributed the most were selected. Continuous data were segmented into epochs of 1000 ms (from -200 ms to 800 ms relative to gaze onset). EEG data was then re-referenced to an average reference and averaged. The total mean number of trials per condition, out of 120, following artifact removal, were 117 for congruent gaze shifts (range 109-120 trials) and 117 for averted gaze (range 106-120 trials).

Two regions of interest (ROIs; see Figure 21) were selected based on previous research and visual inspections of ERPs. A left occipital-parietal ROI was comprised of the four electrode sites P5/P7/PO3/PO7 and a right occipital-parietal ROI comprised of the four electrode sites P6/P8/PO4/PO8. These sites are commonly associated with face processing, gaze processing and attentional processes (see, for example, Eimer, 2011; Hietanen, Leppänen, Nummenmaa, & Astikainen, 2008; Schmitz, Scheel, Rigon, Gross, & Blechert, 2012; Tipples et al., 2013).

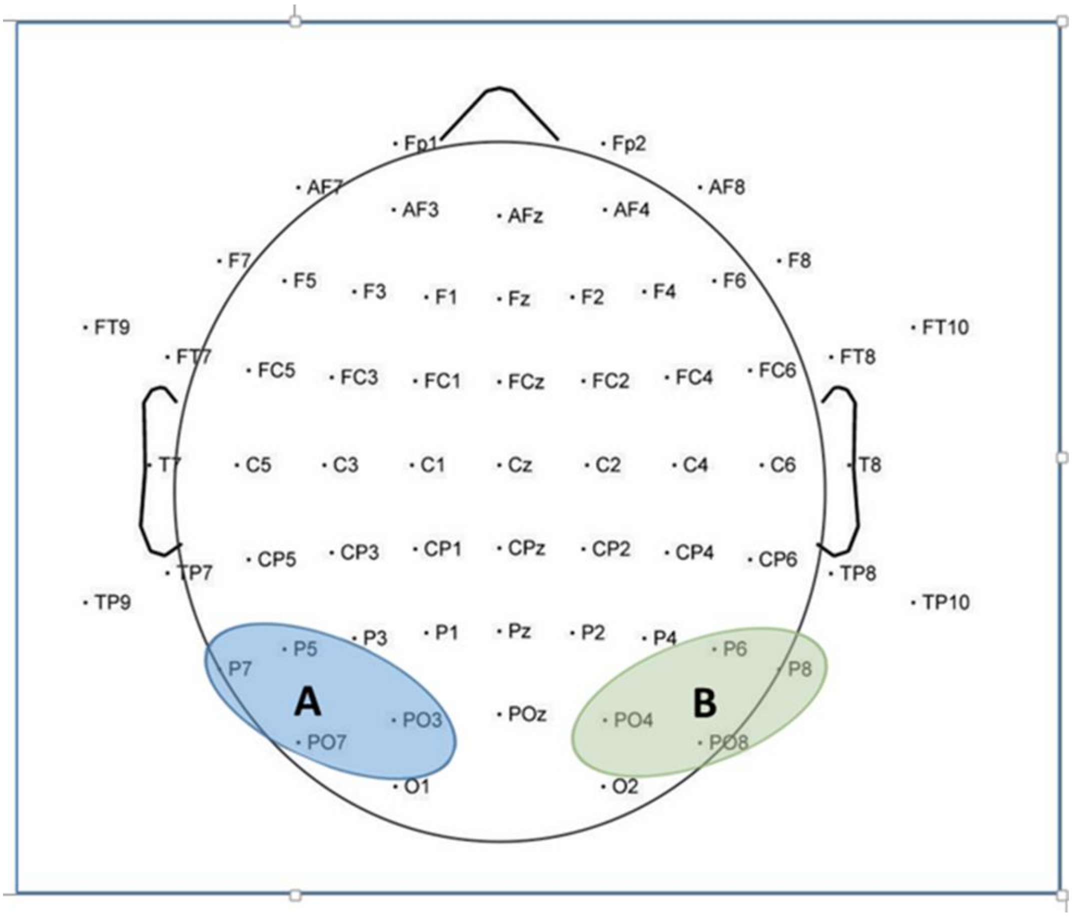

Figure 21. Regions of interest for the ERP analysis. **A**: Blue, Left parietal-occipital ROI **B**: Green, Right parietal-occipital ROI.

ERP trials were time locked to the onset of the gaze stimulus. The amplitudes for ERPs were measured as the mean of all data points between 170-230 ms relative to the mean of all data points in the 200 ms pre-stimulus baseline. This time window was chosen based on a combination of previous research and visual inspection of grand averaged and individual participant's average ERPs (averaged across conditions, as recommended by Luck, 2014).

Experiment 8 Results

Gender categorisation data. Accuracy for identifying gender of faces was very high and did not vary across gaze response condition, as mean accuracy for both conditions was 99 per cent. The mean reaction time (RT) for reporting gender was 558 ms, 95% CI [484, 632] (*SD*=214) for congruent gaze shifts, and 552 ms, 95% CI [479, 624] (*SD*=210) for averted gaze. For faces presented on the left, mean RT was 557 ms, 95% CI [484, 631] (*SD*=214) and, for faces presented on the right, mean RT was 549 ms, 95% CI [474, 624] (*SD*=218).

Reaction times to report the gender of the faces were subjected to an ANOVA with gaze response (congruent or averted) and face position (faces presented on the left or right) as within subject factors. There was no main effect of gaze response, $F(1,34)=0.72$, MSE=1618.65, p=.401, $\eta_p^2 = .021$, and no main effect of face position, $F(1,34)=0.23$, MSE=2219.69, $p=.463$, $\eta_p^2 = .007$.

AQ data. The mean AQ score was 16.63 (SD= 1.33) and the range was 3 to 32. There were no significant correlations between the mean AQ and RT in any of the conditions, rs<0.08, ps>.63. Similarly, there were no correlations between the magnitude of ERP and AQ score, other than one medium, negative correlation between amplitude in the right ROI for averted gaze and AQ, $r(35)=0.348$, $p = .040$. The greater the AQ score, the lower the mean negative amplitude for averted gaze in the right ROI.

ERP Analysis. Grand-averaged ERPs for the mean amplitudes for the time window 170-230 ms for the left and right ROIs can be found in Figure 22**.**

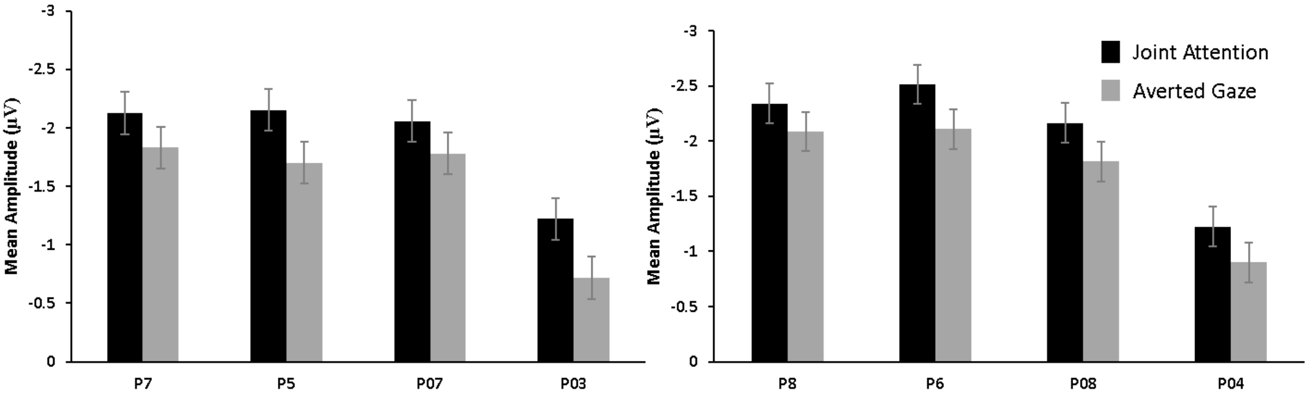

Figure 22. Mean amplitudes for the time window 170-230 ms for each electrode in the regions of interest. The graph on the left shows the left ROI electrodes and graph on the right shows the right ROI electrodes. Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus & Masson (1994).

The mean amplitudes in the time window 170-230 ms were subjected to a two way ANOVA with gaze response (joint attention or averted) and hemisphere (left or right) as within subject factors. There was a main effect of gaze response, $F(1,34) = 13.00$, $MSE= 4.481$, $p = .001$, $\eta_p^2 = 0.28$, and no main effect of hemisphere, $F(1,34) = 1.59$, $MSE= 1.32$, $p = .216$, $\eta_p^2 = 0.05$ and no interaction between gaze response and hemisphere, $F(1,34)= 0.14$, $MSE= 0.02$, $p=.711$, $\eta_p^2 = .004$. Follow up contrasts showed greater negativity (mean difference = -0.36 µV, *SD* = 0.59) for joint attention over averted, *t*(34)= -3.61, *p* = .001, d_z= 0.30, 95% CI [0.12, 0.48] representing a small to medium effect, and no difference (mean difference = 0.19 µV, *SD* = 0.91) between right hemisphere and left hemisphere, *t*(34)= 1.26, *p* = .216, d_z= -0.21, 95% CI [-0.40, 0.09]. At Figure 23 are the grand-averaged ERPs for each electrode in both ROIs and at Figure 24 is a scalp map showing the gaze response difference effects.

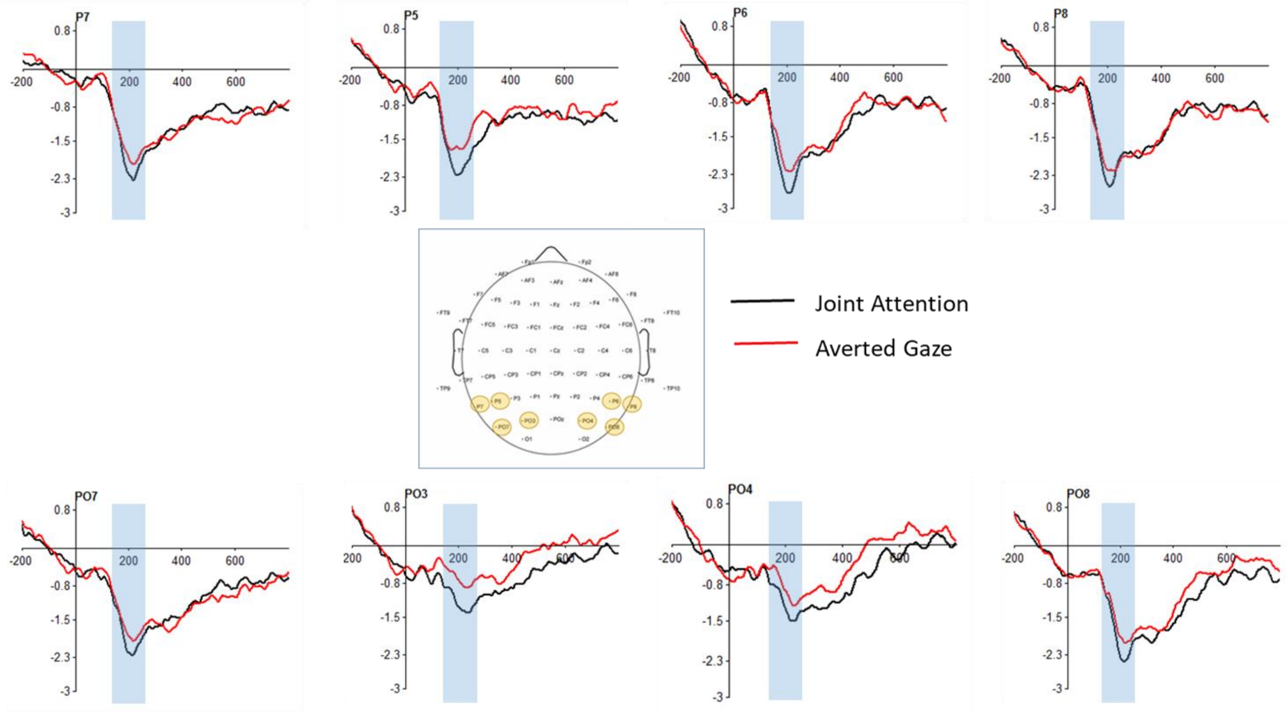

Figure 23. Grand averaged ERPs (n = 35) in response to gaze onset for each electrode in the left (P5, P7, PO3, PO7) and right (P6, P8, PO4, PO8) parietal-occipital Regions of Interest (ROI), showing greater mean negativity at 170-230 ms after gaze shift for joint attention gaze response than for the averted gaze response.

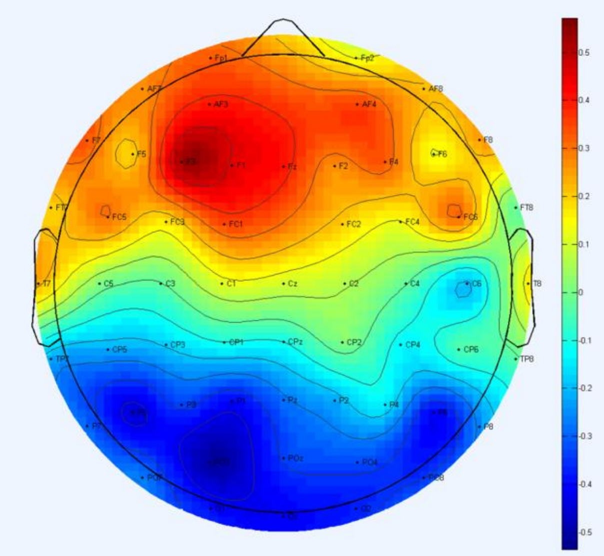

Figure 24. A scalp map showing the gaze response difference effects, calculated as the mean amplitude (in µV) for the joint attention gaze response (where negativity was greater) subtracted from the mean amplitude for the averted gaze response between 170 and 230 ms after gaze shift.

The grand-averaged ERP in response to gaze onset at midline electrode Cz is shown in Figure 25, along with electrode P8, as an example of the N170-like waveform. This is shown to evidence a vertex positive potential (VPP) at Cz which is known to accompany the N170 component. As can be seen from Figure 29, the VPP “mirrors” the N170. The close association shown here between the N170 and VPP is typical of N170 studies (see Eimer, 2011; Joyce & Rossion, 2005). Therefore, this supports the assertion that it is likely to be the N170 component elicited by the gaze shift that can be observed in these data (see later discussion of this).

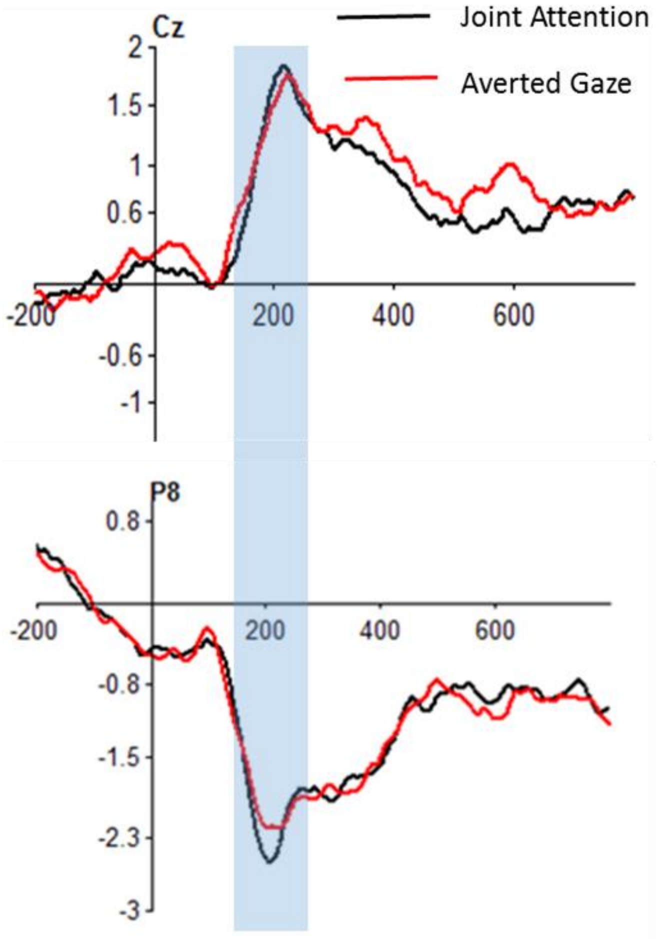

Figure 25. Grand averaged ERPs (n = 35) in response to gaze onset for electrodes P8 and Cz, showing the N170 component at P8 and the vertex positive potential (VPP) component at Cz.

To examine any effects of the onscreen face location, a further 2 x 2 x 2 ANOVA was performed on the mean amplitudes between 170 and 230 ms with face location (left and right), hemisphere (left and right ROI) and gaze response (joint attention and averted gaze) as within subjects factors. There were no main effects of face location, $F(1,34) = 0.09$, $MSE = 0.24$, $p = .762$ $\eta_p^2 = 0.003$, or hemisphere, $F(1,34) = 0.63$, $MSE = 0.99$, $p = .433$, $\eta_p^2 = 0.018$. There was a main effect of gaze response, $F(1,34) = 15.58$, $MSE = 10.21$, $p < .001$, $\eta_p^2 = 0.31$. There was an interaction between face location and hemisphere, $F(1,34) = 24.59$, $MSE = 55.33$, $p < .001$ $\eta_p^2 = 0.42$ (representing a large effect), and no interactions between face location and gaze response, $F(1,34) = 0.34$, $MSE = 0.02$, $p = .856$ $\eta_p^2 = 0.001$, nor between face location, hemisphere and gaze response, $F(1,34) = 0.70$, $MSE = 0.15$, $p = .409$ $\eta_p^2 = 0.02$. Follow-up comparisons showed significant differences between face location and hemisphere, with greater mean amplitude for faces presented on the left in the right ROI than the left ROI, $t(34) = 4.05$, $p < .001$, $d_z = -0.66$, 95% CI [-1.01, -0.30] (mean difference = 1.00

µV, *SD* = 1.47), and greater mean amplitude for faces presented on the right in the left ROI than the right ROI , *t*(34)= 3.54 , *p* = .001, d_z= 0.55, 95% CI [0.22, 0.88] (mean difference = 0.77 µV, *SD* = 1.29), representing medium to large effects.

Experiment 8 Discussion

These data demonstrate an electrophysiological difference between observing a congruent (joint attention) gaze shift and an averted gaze shift, in response to a horizontal gaze leading saccade to an object. This means the neuro-cognitive system differentiates between when a joint attention bid is reciprocated and when it is not. A specific inference from the data is that by around 200 ms after gaze response, the brain detects whether the response is congruent or incongruent. This is remarkable given the small stimulus change in this experiment (the gaze shift) and the fact that this very small change is presented in the periphery. This is enough to elicit an N170-like waveform and for that waveform to be modulated by gaze congruency. This is consistent with some of the previous work showing N170 modulation during gaze processing (Itier et al., 2007; Latinus et al., 2014; Watanabe et al., 2002). This is also consistent with previous behavioural findings that faces who follow our gaze leading capture our attention more than those who do not (Edwards et al., 2015), because a distinction between gaze response must be made by the cognitive system in order to drive this attentional effect. Furthermore, the results might be consistent with previous findings of an implicit sense of agency over a gaze shift response (see Experiments 1-5 in Chapters 2 and 3; Stephenson, Edwards, Howard, & Bayliss, 2018). This is because a reciprocated gaze shift must be distinguished from averted gaze in order to experience a sense of agency over the response. This would need to be investigated much further, though, for example, by investigating whether gaze shifts participants passively observe (without gaze leading first) elicit the same type of effect. If they do, then this would suggest the effect is more one of congruency than agency for gaze leading.

These findings are the first, to my knowledge, to follow the findings of Caruana et al. (2015, 2017) exploring the electrophysiological correlates of gaze leading. There are important distinctions to be made between the current paradigm and the previous work, which mean a direct comparison cannot be made but can explain why Experiment 8 found an N170-like modulation, whilst Caruana (2015, 2017) did not. Caruana et al. (2015, 2017) used a computer-simulated onscreen virtual partner whom participants were told was controlled by a real person in another room who was playing a find and seek game with them. The paradigm used was a virtual game where participants gaze cued their partner (in fact, a computer simulation) to locations where prisoners were escaping. Therefore, the paradigm was more interactive for Caruana et al. (2015,2017) and had more context than that used in Experiment 8 because participants observed whether their attempts to lead the gaze of the avatar prison guard were successful or not. Experiment 8 employed photographs of real faces (as opposed to an onscreen avatar), with no suggestion that those faces were being controlled by a real person, and no game element, just a horizontal gaze leading saccade from a face to an object. Therefore, Experiment 8's paradigm is simpler and likely to employ lower level perceptual processes than the higher-order interactional processes likely to be elicited by the game employed by Caruana et al. (2015, 2017). A final, important difference is that Caruana et al. (2015, 2017) time-locked the ERP to the avatar's gaze shift response which occurred once participants had saccaded back from a cued location to the face (the avatar gaze shift occurred whilst the participant fixated the face), whereas, in this study, ERPs were time-locked to the onscreen face's gaze shift whilst participants fixated the referent object. Therefore, the gaze shift occurred peripherally to fixation in this study, whilst the gaze shift occurred at fixation in Caruana et al. (2015, 2017).

What Caruana et al. (2015) and Caruana et al. (2017) together offered evidence for was a P350 modulation according to gaze response, but only when participants were led to believe they were playing the game with a real person (Caruana et al., 2015) and not when told a computer generated the gaze shift responses (Caruana et al., 2017). This fits with a suggestion made by Carrick et al. (2007), that activity after 300 ms may be affected by social context and meaning, whilst activity at N170 may index gaze shifts, unaffected by social context or mental state judgments. This could be the case because the finding that mean amplitudes between 170 and 230 ms are greater to congruent gaze, could index a detection of gaze response which is elicited by observing a change in attentional state towards one's own gaze location, rather than being associated with any attributions of mental state, given this paradigm had no context within which to make such attributions. This paradigm also employed a task (face gender identification) which was orthogonal to eye gaze response. This is also consistent with the theoretical position of Caruana et al. (2017) that some mental attribution processes may be at work in the P350 their study elicited, given no P350 modulations occurred in the context of an interaction with a computer program, rather than when told a real person was controlling the gaze responses.

The data in Experiment 8 support the notion that we need to process gaze shift responses to joint attention in order to facilitate on-going social interaction. After all, if our gaze shift has not been detected, we need to try again to achieve shared attention. In addition, we need to monitor gaze shift responses which we did not intend to cause, when our object of gaze has been detected and followed. Indeed, it may even be that we are looking at something that may cause us social embarrassment; in those situations, too, it is important to know if our gaze has been followed. The findings show an ERP neural correlate which distinguishes between these responses to our gaze leading, a critical part of social cognition.

It is notable that the averaged ERP waveform peak is later than that which is typically observed for an N170, peaking around 200 ms, rather than 170 ms (see Eimer, 2011, for a review). This may be because participants were fixating on the object when the gaze shift on the face occurred in their periphery. Indeed, this accords with another study who found later N170 onset when faces stimuli are presented in periphery. For example, when presenting fearful faces in periphery, a similar latency was found at 30 degrees eccentricity with peaks of around 200 ms for the N170 component, which occurred earlier when faces were presented closer to centre, at 15 degrees eccentricity (Rigoulot et al., 2011). Most N170 studies have stimuli presented at fovea, so this study has that important distinction.

The waveform observed here is elicited by a peripheral change and time-locked to gaze onset, rather than face onset. Therefore, it may appear a little different in form to the typical N170 components associated with face or gaze processing, although it still does resemble an N170. It is worth considering whether the waveform is being driven by a typical N2pc component, often associated with attentional shifts. Interestingly, both ERPs (N170 and N2pc) elicit greatest effects at the same posterior sites (commonly P7/P8 and PO7/PO8) as each other (see e.g Eimer, 2011; Luck, 2012; Woodman, Arita, & Luck, 2009) so it is possible there is an overlap in the cognitive processes which underlie them. In Experiment 8, multiple cognitive processes are probably contributing to the N170-like waveform. The N2pc is enhanced negativity peaking approximately between 200 and 300 ms after stimulus onset in electrodes contralateral to the location being attended (Woodman et al., 2009). The waveform observed here is a little early for the characteristic N2pc, but is a little delayed to be truly characteristic of the face-sensitive N170. The analysis shows greater negativity at sites contralateral to the face location, like a typical N2pc, but there is no interaction between hemisphere, face location and gaze response. There was a midline VPP at electrode Cz which is typically associated with N170, occurring in the same time window as the N170 (See Figure 29). Taking these considerations together, I believe the waveform this paradigm has elicited is likely to reflect cognitive processes following gaze shift response comprising both attentional (N2pc-like) and face/gaze (N170-like) cognition, although the waveform is more consistent with the N170 component.

Finally, there was a medium correlation between autism-like traits and magnitude of the N170, meaning that the higher the AQ score, the lower the mean 170-230 ms negative amplitude magnitude. However, this was only found for averted gaze in the right ROI and the sample size lacks sufficient power for a reliable correlational analysis. I, therefore, do not make any strong claims based on this finding. However, it is worth mentioning that that there may be individual differences associated with ERP components elicited by responses to gaze leading. This may accord with previous findings of smaller ERP effects for those with an autism diagnosis (e.g. Churches et al., 2012). Future studies would need to explore this possibility.

Limitations and Future Directions

In this study, faces were equally likely to follow gaze as not follow gaze, meaning their gaze responses were not predictable for any trial. Future work could build on these findings by investigating whether the amplitude enhancement for congruent gaze shifts would be eliminated when participants can predict the gaze response. It is possible the ERP may be differently modulated for faces who participants learn will always follow gaze, compared to those who never do. Similarly, an oddball paradigm could be deployed in which faces who usually follow gaze occasionally stop following; this may modulate an early ERP similar to that which emerged here, or it may modulate the P3 component, commonly associated with processing unexpected events (Polich, 2007).

Future studies could also investigate whether the gaze leading saccade from the face to the object before the congruent gaze shift occurs is critical to elicit the neural differentiation for congruent and incongruent gaze. This could be done by adding a fixation cross before the object appears so participants fixate the object all the time, with no prior saccade from face to object. This would mean gaze shift would occur (but still in the periphery like Experiment 8) when participants are already fixating the object. If the same results occurred, that would suggest what drives the effects is the gaze shift towards the object of attention, whether elicited following deliberate gaze leading or following incidental joint gaze. However, if the modulation was not found, then that would suggest gaze leading is necessary for the neuro-cognitive system to distinguish the congruency of gaze shift.

Conclusions

In sum, these data show that when processing gaze shifts in response to gaze leading, there is a modulation of an N170-like ERP component, with greater negativity for congruent gaze shifts over averted gaze responses. This electrophysiological evidence complements previous work which shows dedicated neural networks which can both differentiate between initiating and responding to joint attention, and between the congruency of gaze shift responses (Schilbach et al., 2010). This work also complements behavioural data showing an implicit sense of agency over reciprocated gaze alignment with our own direction of gaze (Stephenson et al., 2018) and an attentional capture effect for faces who follow joint attention bids (Edwards et al., 2015).

SECTION 3: General Discussion

Chapter 6: General Discussion and Conclusion

The aim of this thesis was to add to what is known about the neural and cognitive processes which result from initiating joint attention towards a referent object, causing responding gaze shifts in others. Specifically, there were three empirical aims: to explore whether a sense of agency is felt over gaze responses the gaze leader causes in others, any effects upon the gaze leader's recollection of unfamiliar faces according to gaze response and to explore electrophysiological correlates. This chapter will offer an overview of results across all empirical chapters, together with limitations and future directions. Then, I will offer a new, neuro-cognitive model of joint and shared attention informed both by previous research and by the data in this thesis. How this model fits with existing theories will be described, followed by a thesis summary.

Results Overview

Agency. In Chapters 2 and 3 (Experiments 1-5), it was demonstrated that the temporal binding effect, often used as an index of an implicit sense of agency, occurs when gaze shift responses result from gaze leading during joint attention. These effects evidence an implicit sense of agency over congruent gaze shifts elicited by us during a simulated social interaction. Experiments 1, 2 and 5 each demonstrated no sense of agency when gaze shifts are observed, without any gaze leading saccades towards the object of the gaze shift. Experiments 1-3 each evidenced a contrasting effect of an implicit sense of agency over congruent gaze shifts when participants had executed a horizontal saccade to the referent object when they were told they caused the gaze shift. Experiments 2 and 3 demonstrated a weaker, but reliable (of about half the magnitude of binding when gaze leading preceded the gaze shift), implicit sense of agency over congruent gaze shifts towards an object participants were fixating upon, even without a horizontal saccade first from the face to that object. This was suggestive of a hierarchical system in which the greatest binding effects result from deliberate gaze leading from a face to an object and agency is attributed, but less so, when joint attention results incidentally. However, this did not replicate in Experiment 5 (Chapter 3), where no binding was found for congruent gaze shifts in an experiment where participants never experienced gaze leading from the face to an object and were never told their gaze shifts elicited onscreen responses, suggesting context and participant belief will modulate sense of agency without a deliberate joint attention bid.

Experiments 1, 3 and 4 included a non-social response for comparison with gaze shift responses. A change to a fixation cross, following gaze leading of the same saccade magnitude as the face stimulus conditions between fixation crosses (Experiment 4), showed no binding effects. This suggested any temporal compression effects associated purely with executing saccades cannot account for the binding effects elicited within the context of an interaction with an onscreen face. I infer that there is something special about a congruent gaze shift in response to gaze leading which is most likely to induce the greatest sense of agency. Experiment 1 demonstrated, in a control condition, that a low level spatial shift towards an object with a phase scrambled image, with no prior gaze leading to cause it, will not induce temporal binding. Experiment 3 explored the same spatial shift from a phase scrambled display, but when the shift occurred, participants were fixating the object, so it was a congruent spatial shift towards their object of gaze. This did elicit temporal binding, surprisingly. However, I attribute this as most likely caused by the ambiguity of the situation when it is adaptive for us to over-attribute agency if there is any doubt about what is happening as the cost of under-attribution for causing events in the world is worse than over-attributing agency. If in doubt, assume we caused something, so we can plan what we need to do about it next, if anything. This is perhaps one of the most challenging findings to explain in this thesis and so may prove the most fruitful for future studies to explore further.

Face recognition memory. Experiments 6 and 7 in Chapter 4 showed that there is no effect on recognition of faces according to their previously observed gaze responses to gaze leading. This is perhaps because it is adaptive to be able to recognise faces previously encountered regardless of gaze response, as cooperative followers and norm violators are equally important to remember. Across both experiments, there were, albeit with a weak effect, faster reaction times to report old/new faces who had not followed gaze over the faces who did, suggesting there might be circumstances during gaze behaviours in which norm violators are easier to remember than cooperative partners.

Neural mechanisms. Experiment 8 in Chapter 5 shows that the neuro-cognitive system by around 200 ms distinguishes between a congruent gaze shift response to gaze leading, and an incongruent one. This is evidenced by greater mean negative amplitude between 170 and 230 ms following gaze response, likely to be an

N170 ERP component modulation and which evidences the detection of a successful joint attention bid.

Autism-like traits. There were no meaningful, reliable correlations between levels of autism-like traits and sense of agency (Experiments 1-3), or face recollection (Experiments 6-7), nor modulation of the N170 component (Experiment 8). This might be surprising given the previous findings in the joint attention literature which have often pointed to differences in gaze behaviours and also differences in the broader phenotype (reviewed in Chapter 1). However, this is also in line with studies which show how often there can be similarities to typically developing populations, not just differences, in gaze processing and behaviours driven by gaze for those with autism or nonclinical populations with high autism-like traits (see e.g. Pell et al., 2016). It should be noted, though, the lack of any correlations might be due to the fact the experiments in this thesis all examined individual differences in the broader phenotype in nonclinical and small samples where the modal AQ score was low.

Further Discussion of Results

Now that the results have been overviewed, some further thoughts and ideas will be offered about what these findings could mean for each of the three empirical strands of the thesis, followed by a discussion of how all three strands may relate to each other in a integrated discussion of the results.

Agency Findings. In Chapter 2, the idea was introduced that sense of agency may not be what the phenomenon of temporal binding actually taps into (Dewey & Knoblich, 2014; Hughes et al., 2013). Hughes et al. (2013) argue that the effects might be caused mainly by differences in temporal prediction, citing temporal attention evidence that sensory processing can be weakened by temporal prediction. Hughes et al. (2013) accept, however, that it is probably a combination of processes such as attention, prediction and causality which alter sensory outcome of actions. I agree that a variety of cognitive processes may lead to the temporal binding effect but this does not negate the most likely explanation that temporal binding does reflect a sense of agency, given the typical stark contrast in binding between active and passive conditions found in the literature and found in the experiments in Chapters 2 and 3. Attentional processes are likely to play a part in the perception of the passage of time. After all, we commonly talk about how time “flies” or “drags” usually experiencing a sense of time passing quickly when we are busy and, therefore, our attention is absorbed in what we are doing. This is when we are in control of our actions, that is, experiencing the greatest sense of agency over them.

Taking this idea of attentional processes at work in sense of agency, this may help explain some of the more unexpected findings in Chapter 3. In Experiment 4, although there was no significant binding, there was a significant difference in temporal reproduction errors between the saccade and no saccade task with greater time interval reproductions when not making a saccade than when making a saccade between crosses. Therefore, time was perceived to pass more slowly when no action was performed. The fact a saccade had to be performed was likely to engage more attention than just fixating a cross with no action required. If attentional engagement with a motor task (a saccade here) helps compress time, then making a saccade in the context of interacting with an onscreen face is likely to compress time more, as more attention is engaged. This could help explain why the gaze leading task *with a face* in Experiments 1-3 resulted in significant temporal compression effects, whilst that *with a fixation cross* in Experiment 4 did not. This could speak to an additive effect of social context to sense of agency over and above effects caused by a motor action alone.

In Experiment 5, there were no significant binding effects when fixating the face throughout or when fixating the object throughout. In both conditions, there was no motor action (saccade) to contribute to a sense of agency. However, there was a perception of the time interval being longer when fixating the face alone, compared with fixating the object. Attentional processes are different when awaiting a gaze shift which must be detected in the periphery (when fixating the object) compared to fixating the face where gaze shifts can be detected directly. This may account for the differences in the perception of time.

In addition to temporal attention, Hughes et al. (2013) emphasised the involvement of prediction in attenuating sensory outcomes. Social outcomes we cause in others can vary a great deal in their predictability. It is usually easier to predict outcomes we cause in inanimate objects who have no free will than the outcomes from humans who may respond to our actions less predictably or, at least, have a greater range of possible responses. Faces always responded to gaze leading in Experiments 1-3 and 5 with a congruent gaze response as the outcome, which meant the paradigm was 100% predictable. It is possible different results would be found if the predictability of gaze response was manipulated. Future studies could explore whether the effects would be attenuated by less likelihood of a congruent response. I speculate that agency effects might be robust to such manipulation because it is equally important to detect an incongruent gaze shift, which may be self-attributed just as much as a congruent one, and because of the findings of binding in the ambiguous phase-scrambled condition demonstrated in Experiment 3. If in doubt, perhaps the system over-attributes agency, and this is an interesting avenue for future studies to reveal.

In Chapter 2, Synofzik et al’s (2013) model of optimal cue integration for sense of agency was introduced. This model included both retrospective and predictive processes at work with feelings of agency (implicit agency) and judgements of agency (explicit agency) conceptualised as related, yet dissociable. The findings from Experiments 1-5 fit well with that model. There were no reliable correlations between implicit agency (binding effects) and explicit agency ratings, consistent with the two processes being dissociable. At the same time, agency ratings were lower for the conditions in which there was no binding, supporting the idea that explicit and implicit agentic processes are related, nevertheless. The optimal cue integration account does not mention how social context may feed into the feelings of agency so is something the findings in Experiments 1-5 could add. There is an affective level in the optimal cue integration account which includes emotional appraisal and reward anticipation feeding into both feelings and judgments of agency. Perhaps a further level could be added of social context, related to the affective level.

Face Recognition Findings. The face recognition results found similar levels of recognition ability for both previously encountered followers and non-followers after gaze leading. It could be, as already discussed, that the equal importance of both responses in social interactions can explain the lack of effects found. The binary old/new response may be too blunt a tool in itself to fully test this and future work could apply, instead, a more continuous measure to be more sensitive to any underlying encoding differences. However, there was one weak finding that non-followers are better encoded in memory demonstrated by faster reaction times (which is a continuous measure) to identify whether a face had been seen before for non-followers over followers. A change in design may reveal a stronger effect of this, for example, by participants encountering the faces more often and/or introducing a much longer delay before the recall task. Another future avenue would be to analyse the eye metrics further. Participants were instructed to look back at the face following their peripheral detection of the gaze shift response. Therefore, the eye metrics could be analysed to ascertain if saccade latency here or duration of fixation on the face once they saccaded back to the face correlated with recognition accuracy and whether there were any differences between conditions for this. Such differences in fixation duration or saccade latency, were there to be any, between followers and non-followers may have influenced subsequent recognition.

Neural Findings. The N170 modulation shown in Experiment 8 is the first of its kind to examine any N170 effects on responses to gaze leading. It is notable that Experiment 8 examined the effects of the gaze response when participants were fixating the object and detecting the gaze shift in their peripheral vision This can be contrasted with previous ERP gaze studies. Previous studies have looked at ERP modulation either for direct versus averted gaze towards the participant or observing gaze shifts not elicited by the participant (e.g. Latinus et al., 2014; Senju et al., 2006) and all have done so when participants fixated the face. None have found modulation of the N170 for peripherally presented gaze shifts. Therefore, the findings in Experiment 8 are the first to find N170 modulation for such gaze shifts. It remains for future studies to build on this finding to ascertain if the effect is solely due to a congruency effect generally or due to a congruency effect only found following gaze leading. To test for this, a future work needs to employ a gaze leading paradigm coupled with a control condition in which gaze congruency is observed without gaze leading first (e.g. when already fixating on the object).

Integrated Discussion of Results. Taking all eight experiments together, we now know that a sense of implicit agency can be felt over causing gaze shifts to align with our own, that we can encode new faces well in memory after brief gaze leading interactions, whatever the outcome, and that there is electrophysiological evidence of the detection of congruent or incongruent gaze responses. It may be that these three findings are linked by the overarching theme of sense of agency. Perhaps the high accuracy in recalling the unfamiliar faces in Experiments 6 and 7 was helped by the fact participants gaze led the faces and the fact they sensed agency over the outcomes. Controlling another's gaze may boost memory for that person as the interaction is more meaningful than when passively seeing a new face. This needs to be tested, though, in future research by introducing a control recall task, for example, where there is passive observation of new faces to see whether causing the gaze shift drives higher recollection accuracy. The weak effect in Experiments 6 and 7 of faster reaction times to non-followers could also be due to the violation of expected agentic control over those faces. The N170 modulation according to gaze outcome might be a signature of agentic control over gaze shifts but it remains to be tested whether this is simply a congruency effect that would be elicited in the absence of a causal gaze leading action.

Limitations and Future Directions

Agency. Although the majority of the literature supports the idea that temporal binding is a measure of implicit sense of agency (e.g. Haggard, 2017, for a review), there are some dissenters (e.g. Hughes et al., 2013), and so it cannot be ruled out completely that the temporal binding effects demonstrated here are not a proxy for an implicit sense of agency. However, I believe the best explanation for the binding effects in the active tasks in Experiments 1-3 is an implicit sense of agency, given no such effects are found in passive tasks, coupled with the ever growing literature which also supports this position and the points made about attentional processes at work made in the further discussion of agency findings above.

Future research could employ a gaze contingent paradigm to see if this leads to an increase in the magnitude of the effects found. The findings of attenuated binding for incidental joint attention found in the Passive Object task in Experiments 2 and 3 were not replicated in Experiment 5, meaning more research is needed to resolve this. Relatedly, fruitful areas for future research would be to explore the finding of large binding effects in the Scrambled Passive Object task of Experiment 3, to explore further whether introducing ambiguity and uncertainty into a paradigm can still elicit binding when there is any observed shift towards our object of gaze. It could be that it is so fundamentally important to our social interactions that any congruent shift, whether it be from eyes or other stimuli, will always be processed as potentially self-caused.

Face recognition. Experiments 6 and 7 with the particular paradigm employed was just a starting point to exploring any effects on face recognition following gaze leading. There are other ways to measure recollection beyond the old-new task which was employed. Using an n-back task to explore working memory effects (Kane, Conway, Miura, & Colflesh, 2007) could form the basis for a future study. Greater exposure to the faces than were employed here might facilitate enhanced encoding in memory. Finally, measuring recall from longer-term memory could be explored.

Neural mechanisms. Future work could build on these findings by investigating whether the amplitude enhancement for congruent gaze shifts would be eliminated when participants can predict the gaze response. It is possible the ERP may be differently modulated for faces who participants learn will always follow gaze, compared to those who never do. Similarly, an oddball paradigm could be deployed in which faces who usually follow gaze occasionally stop following; this may modulate an early ERP similar to that which emerged here, or it may modulate the P3 component, commonly associated with processing unexpected events (Polich, 2007). Future studies could also investigate whether the gaze leading saccade from the face to the object before the congruent gaze shift occurs is critical to elicit the neural differentiation for congruent and incongruent gaze.

Finally, there is a general limitation of all of the experiments in this thesis which is that effects were explored within a computer-simulated joint attention paradigm with a virtual other, rather than a real world interaction between dyads. A review of the literature has identified that the field has progressed from static to interactive paradigms over the years, and the importance of such developments (Pfeiffer et al., 2013). The experiments in this thesis employ an interactive methodology as the virtual partner responds to participant gaze behaviours. However, future studies can build on these findings to explore whether the effects demonstrated here extend to real-world interactions. This would, of course, introduce new challenges, making it harder to ensure tightly controlled, measurable variables, of the type demonstrated here using an interactive, simulated approach. The field recognises how grappling with these issues has resulted in compromises to ecological validity in social gaze-based research, but the field is, nevertheless, increasingly working towards attempts to overcome this (Schilbach, 2015).

Model of Shared Attention and Associated Neural Mechanisms

The findings in this thesis, together with the previously reviewed literature on joint and shared attention, enable the formulation of a novel model to capture the processes at work during a joint attention interaction and the neural regions involved. The model aims to capture how both people in a joint attention interaction have to coordinate their behaviour, how this leads to a state of shared attention which, in turn, facilitates a number of subsequent social cognitive processes. This model of joint and shared attention is found at Figure 26.

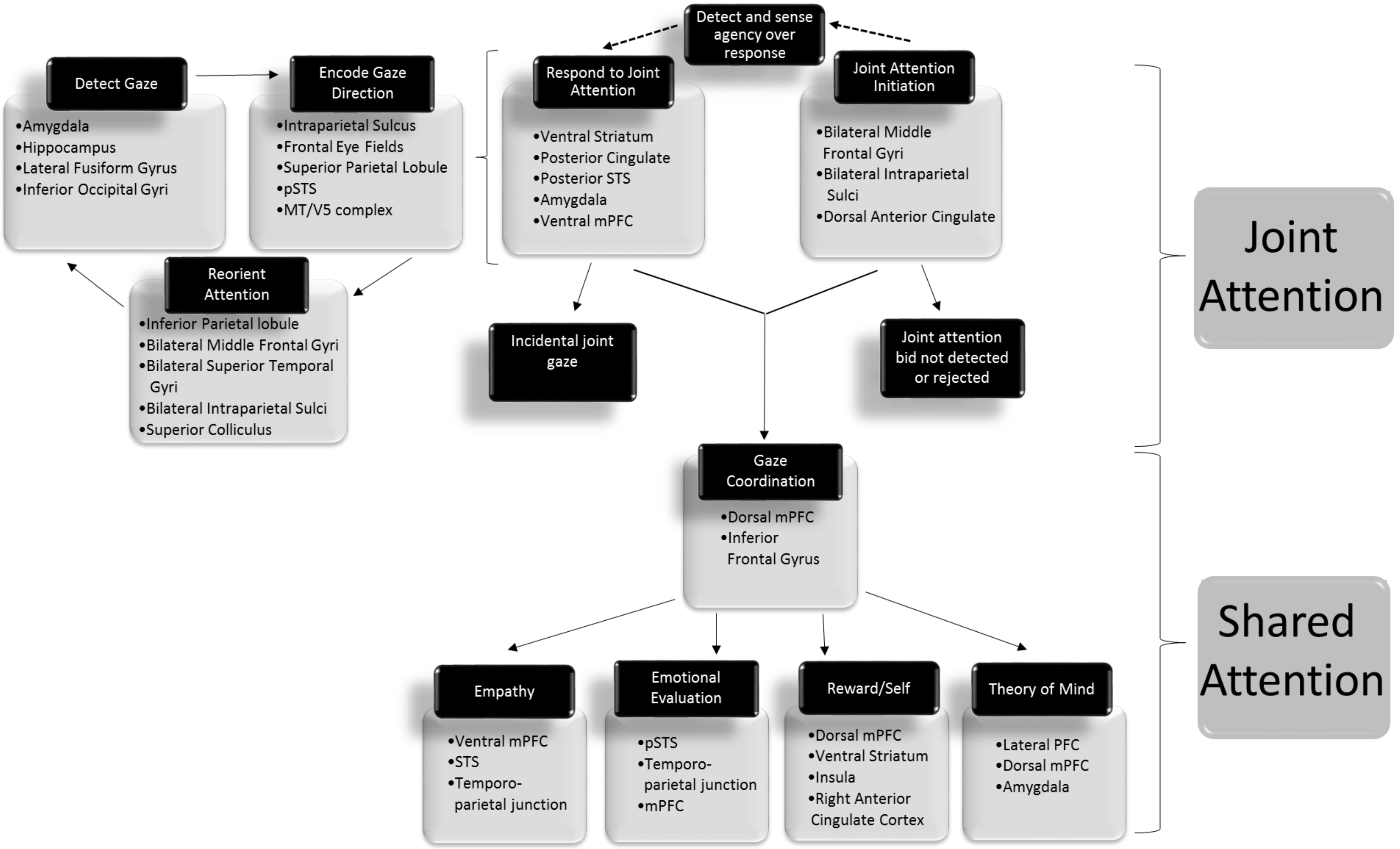

Figure 26. Neurocognitive Model of Joint and Shared Attention. The two agents need to coordinate gaze in order to engage in shared attention. The initiator needs to detect and sense agency over the response received. Coordinating gaze requires both agents to be aware of their joint gaze, elevating ‘joint’ to ‘shared’ attention. This then facilitates the social-cognitive processes of empathy, emotional evaluation, sense of reward and self and theory of mind. The sub-serving brain regions are identified in this model.

Two agents during joint attention need to coordinate their attention and so the model shows how this leads to the higher level state of sharing attention, in turn, facilitating key components of social cognition: empathy, emotional evaluations, sense of self and reward and theory of mind. In addition, the model demonstrates that *both* the initiator and responder during the interaction need to detect the presence of gaze, encode each other's gaze direction and then reorient their attention. In addition, the initiator needs to detect and sense agency over the response received, which is one part of the model the data in this thesis specifically informs.

How experiments 1-8 contribute to the model. The model is based on previous research into all the processes at work during joint and shared attention and the associated brain regions involved. The experiments in this thesis specifically contribute the processes of the detection of response and sense of agency experienced over the follower's responses to their joint attention initiation bid (as shown by the dashed lines at the top of the model in Figure 26). The gaze leading paradigms used throughout these experiments all provide some evidence of these two critical processes (detection and agency) at work for the gaze initiator. Gaze leading can facilitate the gaze coordination found in shared attention. In turn, ultimately this can lead to the consequent cognitive outcomes identified in the model. What the model suggests is that gaze leading is a critical part of the social interaction which requires agency over gaze outcomes, in turn, eventually facilitating other social cognitive outcomes of empathy, emotional evaluation, sense of reward and self and theory of mind.

It is debateable how 'automatic' the processes of, firstly, detecting the outcome of gaze leading and, secondly, the resultant sense of agency are. The rapid 200 ms peak of congruency detection evidenced by Experiment 8, suggests the detection may be reflexive. The implicit sense of agency evidenced in Experiments 1-4 suggest, too, a reflexivity and perhaps the sense of agency does always reach conscious awareness. This is because the self-reported explicit agency did not always tally with the implicit binding measure. For example, there were low agency ratings in Experiments 2 and 3 for passive object conditions, but significant temporal binding effects and in Experiment 3 there were low agency ratings for scrambled object condition, yet large binding effects. This speaks again to a dissociation between explicit and implicit agency which fits with the optimal cue intgration account of agency (Synofzik et al., 2013).

Application of the model to theory. The extensive network sub-serving gaze processing, initiating joint attention, responding to joint attention and shared attention consequences are summarised in the model. The theoretical framework offered is that initiating shared attention has fundamental benefits for the initiator; we are motivated to share attention as part of the human capacity for social cognition and intergenerational transmission of culture, including language. Just as gaze following allows access to mentalising about other's intentions, beliefs and expected behaviour, so initiating shared attention allows us to share our thoughts and experiences with others. This motivation to share our thoughts and experiences with others, supported by joint attention, has been identified by Tomasello et al. (2005) and argued to be what sets us apart as a species and facilitates shared intentionality and, critically, allows our culture to evolve.

A recent theory, termed sociomotor action control (Kunde, Weller, & Pfister, 2017) is consistent with the model. In the model the joint attention initiator needs to detect the response to their gaze leading action in another person, in order to coordinate mutual gaze and lead to on-going cognitive outcomes. It is this detection of our action outcomes on *other people's behaviour* that is captured by the idea of sociomotor action control which is that the responses elicited in another's behaviour feedback to inform further action control (Kunde et al., 2017). I add to this idea that social responses from other people are much less predictable than typical action-outcome in inanimate objects manipulated by us. Arguably, the variance in possible outcomes from another person whose behaviour can change on a whim, is far greater than the variance in expected outcomes from inanimate objects. Therefore, we need to be particularly flexible in our assessment of feedback from social outcomes; the system must be capable of processing a huge range of responses. In the context of shared attention, possible outcomes include being ignored in our gaze leading bid and having to re-establish eye contact and repeat the gaze leading saccade. I also note that, as shown in the model, incidental joint gaze may occur because our gaze can be followed without any deliberate intent on our part to establish shared attention.

The model captures the neural mechanisms of the gaze detection process, the coordination needed between both initiator and responder and the potential resulting cognitive processes (empathy, emotional evaluations, sense of self/reward and theory of mind) which are integral to the way we interact as human beings. Our motivation to engage with others is facilitated through shared attention which is adaptive to our functioning in social groups and the shared intentionality we can engage in which makes us so successful as a species (Tomasello & Herrmann, 2010).

The model of shared attention contributes to wider theories about social attention and social cognition. For example, the model can lend support to Frith's "we-mode" theory that, when agents are interacting, they engage in a collective mode of cognition and tend to co-represent actions of social partners (Gallotti & Frith, 2013). This is supported by studies showing activation of the inferior frontal gyrus when engaging in mutual gaze, specifically coordinating gaze with a social partner (Cavallo et al., 2015; Koike et al., 2016) which is the same region where evidence for human mirror neurons has been offered, using a repetition suppression paradigm (Kilner, Neal, Weiskopf, Friston, & Frith, 2009). Future research could investigate whether repetition suppression is found both when executing a repeated joint attention initiation to an object and when observing another person repeating a joint attention initiation. This would enable exploration of whether there is evidence of mirror neurons within the inferior frontal gyrus specifically for joint attention bids, which would support the idea both of a human mirror neuron system and the overlapping theory of co-representing a social partner's actions.

Koike et al. (2016) used hyperscanning functional magnetic resonance imaging where two participants shared attention and found synchronisation of neural activity of the inferior frontal gyrus (IFG) during mutual gaze and also IFG activation during both initiating and responding to joint attention. Furthermore, eye blinks were coordinated during a joint attention task. This is further evidence of a shared representational state during shared attention which facilitates theory of mind and other key elements of social cognition.

The neural network in the model. The model reflects past neuroimaging findings which broadly support the Parallel and Distributed-Processing Model hypothesised by Mundy & Newell (2007) and Mundy et al. (2009) in which joint attention initiation is mainly served by the anterior attention network and joint attention response by the posterior attention network, but with an emphasis on the connections between these networks. A further intriguing neural correlate is that shown by Schilbach et al. (2010) who demonstrated enhanced ventral striatum activity for initiating joint attention, suggesting this is a rewarding experience. This activity also correlated with self-reported subjective feelings of pleasantness. The greater the activity change in the ventral striatum, the greater the sense of pleasantness reported for looking at objects with another person. In this case the other person was an onscreen face but participants were told that the onscreen face was controlled by a real person. This type of research, which seeks to examine online social interactions rather than offline, has become popular in recent years (see Pfeiffer et al., 2013, for a review).

A further study from the same lab showed that gaze based behaviours with another person activated the ventral striatum, and it did not matter whether the participants believed their partner had a shared goal or not (Pfeiffer et al., 2014). Another study found increased striatum activity when initiating joint attention was reciprocated with gaze following, compared to an averted gaze response (Gordon et al., 2013). Finally, the ventral striatum was activated more even when participants simply passively observed actors in a video clip engaging in a shared purpose than when the actors were simply acting in parallel (Eskenazi et al., 2015).

Cognitive outcomes in the model. The four key cognitive outcomes from shared attention are identified in the model: empathy, emotional evaluation, reward/sense of self and theory of mind. I will consider each of these in turn. Firstly, the brain regions implicated both by joint attention and the processing of empathy are the pSTS and the temporo-parietal junction (TPJ), whilst another overlapping region involved is the ventral MPC (see Bernhardt & Singer, 2012, for a review, and Bzdok et al., 2012, for an activation likelihood activation meta-analysis). Secondly and relatedly, making emotional evaluations can result from shared attention which can be split into evaluations about oneself and evaluations of others' emotions. Those regions involved in our own emotional evaluations are the insula and the right anterior cingulate cortex (ACC), whilst those used when evaluating other's emotions are the STS and the TPJ, and those used for both types of evaluation are the amygdala, the lateral PFC and the dorsal mPFC (see Lee & Siegle, 2009, for a review). Thirdly, initiating shared attention has been associated with the dorsal mPFC (Schilbach et al., 2010) which has been implicated in processing self-referential information (Bergström, Vogelsang, Benoit, & Simons, 2014; Schmitz & Johnson, 2007), and with regions associated with processing reward like the ventral striatum, insula and right ACC (Redcay et al., 2010; Schilbach et al., 2010). Because these two processes of reward and self-referential information implicate the same brain regions, I have combined those processes together in the model. Fourthly, sharing attention with another person facilitates the human attribute of theory of mind and its accompanying potential for cooperation, teaching, control and communication. The neural mechanisms of theory of mind have been identified as the pSTS, the TPJ and the mPFC (Saxe, 2006; and see Schurz, Radua, Aichhorn, Richlan, & Perner, 2014, for a meta analysis).

To summarise, the model captures the neural mechanisms of the gaze detection process, the coordination needed between both initiator and responder and the potential resulting cognitive processes (empathy, emotional evaluations, sense of self/reward and theory of mind) which are integral to the way we interact as human beings. Our motivation to engage with others is facilitated through shared attention which is adaptive to our functioning in social groups and the shared intentionality we can engage in which makes us so successful as a species (Tomasello & Herrmann, 2010).

Thesis Summary

The empirical chapters of this thesis (Experiments 1 to 8) together offer new evidence of a sense of agency over gaze shifts we elicit in others, of rapid neural differentiation between congruent and incongruent gaze shift responses, and evidence an equally efficient ability to recollect unfamiliar faces following both congruent and averted gaze shifts. Together, these insights contribute to our understanding of cognitive processes at work during joint and shared attention, specifically, adding to what we know about the initiator of the interaction. These data have helped inform a new neuro-cognitive model of these processes, which I now summarise.

The new neurocognitive model of joint and shared attention in this thesis seeks to capture three main strands. Firstly, the model captures the many findings and advances made over the past twenty years about the brain regions involved in the different processes which make up sharing attention using eye gaze signals. Secondly, the model shows how these processes relate to one another in an iterative fashion, and, thirdly, identifies the key outcomes which can result from sharing attention; empathy, emotional evaluation, sense of self/reward and theory of mind. An implicit sense of agency for causing eye gaze shifts in others, together with the neural detection and differentiation of gaze responses demonstrated in this thesis, are captured within the model. Specifically, experiencing agency over gaze shifts and differentiating a congruent from an incongruent response, is necessary to detect the response and coordinate gaze during the on-going social interaction which, in turn, can lead to empathy, emotional evaluations, sense of self, reward and theory of mind. These elements of social cognition are the foundations upon which human cooperation and communication are based.

References

Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology, 18*(2), 166–172.

Adolphs, R., & Spezio, M. (2006). Role of the amygdala in processing visual social stimuli. *Progress in Brain Research, 156,* 363–378.

American Psychiatric Association. (2013). DSM 5. American Psychiatric Association.

Anderson, J. R., & Gallup, G. G. (2015). Mirror self-recognition: a review and critique of attempts to promote and engineer self-recognition in primates. *Primates*, *56*(4), 317–326.

Bahill, A. T., & Stark, L. (1979). The trajectories of saccadic eye movements. *Scientific American*, *240*(1), 108–117. https://doi.org/10.1038/scientificamerican0179-108

Bard, K. A. (2009). Social cognition: Evolutionary history of emotional engagements with infants. *Current Biology, 19*(20), R941–R943.

Barlas, Z., & Obhi, S. (2013). Freedom, choice, and the sense of agency. *Frontiers in Human Neuroscience, 7,* 514.

Baron-Cohen, S. (1997). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *The Maladapted Mind: Classic Readings in Evolutionary Psychopathology,* 207–239.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism Spectrum Quotient : Evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. https://doi.org/10.1023/A:1005653411471

Bayliss, A. P., Bartlett, J., Naughtin, C. K., & Kritikos, A. (2011). A direct link between gaze perception and social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 634.

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104*(3), 644–653. https://doi.org/10.1016/j.cognition.2006.07.012

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, *13*(6), 1061–1066.

Bayliss, A. P., Pellegrino, G. di, & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *The Quarterly Journal of Experimental Psychology*, *58*(4), 631–650.

Bell, R., Buchner, A., & Musch, J. (2010). Enhanced old–new recognition and source memory for faces of cooperators and defectors in a social-dilemma game. *Cognition*, *117*(3), 261–275.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*(6), 551-565.

Bergström, Z. M., Vogelsang, D. A., Benoit, R. G., & Simons, J. S. (2014). Reflections of oneself: neurocognitive evidence for dissociable forms of self-referential recollection. *Cerebral Cortex*, *25*(9), 2648–2657.

Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review* *of Neuroscience*, *35*.

Beyer, F., Sidarus, N., Fleming, S., & Haggard, P. (2018). Losing Control in Social Situations: How the Presence of Others Affects Neural Processes Related to Sense of Agency. *ENeuro*, *5*(1), ENEURO-0336.

Billeci, L., Narzisi, A., Campatelli, G., Crifaci, G., Calderoni, S., Gagliano, A., … Muratori, F. (2016). Disentangling the initiation from the response in joint attention: an eye-tracking study in toddlers with autism spectrum disorders. *Translational Psychiatry*, *6*(5), e808.

Binetti, N., Harrison, C., Coutrot, A., Mareschal, I., & Johnston, A. (2015). Reciprocating the gaze of others: how we look and how long we like to be looked at. *Journal of Vision*, *15*(12), 172.

Birmingham, E., Stanley, D., Nair, R., & Adolphs, R. (2015). Implicit social biases in people with autism. *Psychological Science*, *26*(11), 1693–1705.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237–242.

Bottini, S. (2018). Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis. *Research in Autism Spectrum Disorders*. https://doi.org/10.1016/j.rasd.2017.10.001

Boucenna, S., Narzisi, A., Tilmont, E., Muratori, F., Pioggia, G., Cohen, D., & Chetouani, M. (2014). Interactive technologies for autistic children: A review. *Cognitive Computation*, *6*(4), 722–740.

Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology*, *119*(2), 145.

Brignani, D., Guzzon, D., Marzi, C. A., & Miniussi, C. (2009). Attentional orienting induced by arrows and eye-gaze compared with an endogenous cue. *Neuropsychologia*, *47*(2), 370–381.

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*, 67–78.

Buehner, M. J. (2012). Understanding the past, predicting the future: causation, not intentional action, is the root of temporal binding. *Psychological Science*, *23*(12), 1490–1497.

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796.

Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., & Henson, R. N. A. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology*, *17*(1), 20–25.

Call, J., & Tomasello, M. (2005). What chimpanzees know about seeing revisited: An explanation of the third kind. *Joint Attention: Communication and Other Minds*, 45–64.

Cameron-Faulkner, T., Theakston, A., Lieven, E., & Tomasello, M. (2015). The Relationship Between Infant Holdout and Gives, and Pointing. *Infancy*, *20*(5), 576–586. https://doi.org/10.1111/infa.12085

Capozzi, F., Becchio, C., Willemse, C., & Bayliss, A. P. (2016). Followers are not followed: Observed group interactions modulate subsequent social attention. *Journal of Experimental Psychology: General*, *145*(5), 531.

Carlin, J. D., & Calder, A. J. (2013). The neural basis of eye gaze processing. *Current Opinion in Neurobiology*, *23*(3), 450–455.

Carlson, T. A., Hogendoorn, H., & Verstraten, F. A. J. (2006). The speed of visual attention: What time is it? *Journal of Vision*. https://doi.org/10.1167/6.12.6

Carpenter, M., & Call, J. (2013). How joint is the joint attention of apes and human infants? In J.M Metcalfe & H.S Terrace (Eds.) Agency and Joint Attention, (49-61). New York: Oxford University Press.

Carrick, O. K., Thompson, J. C., Epling, J. A., & Puce, A. (2007). It??s all in the eyes: neural responses to socially significant gaze shifts. *NeuroReport*. https://doi.org/10.1097/WNR.0b013e3280ebb44b

Caruana, N., Brock, J., & Woolgar, A. (2015). A frontotemporoparietal network common to initiating and responding to joint attention bids. *Neuroimage*, *108*, 34–46.

Caruana, N., de Lissa, P., & McArthur, G. (2015). The neural time course of evaluating self-initiated joint attention bids. *Brain and Cognition*. https://doi.org/10.1016/j.bandc.2015.06.001

Caruana, N., de Lissa, P., & McArthur, G. (2017). Beliefs about human agency influence the neural processing of gaze during joint attention. *Social Neuroscience*. https://doi.org/10.1080/17470919.2016.1160953

Cavallo, A., Lungu, O., Becchio, C., Ansuini, C., Rustichini, A., & Fadiga, L. (2015). When gaze opens the channel for communication: Integrative role of IFG and MPFC. *Neuroimage*, *119*, 63–69.

Chambon, V., & Haggard, P. (2013). 14 Premotor or Ideomotor: How Does the Experience of Action Come About? *Action Science: Foundations of an Emerging Discipline*, 359.

Chambon, V., Sidarus, N., Haggard, P., Blakemore, S. J., Berberian, B., Sarrazin, J. C., … Vogeley, K. (2013). Priming determinist beliefs diminishes implicit (but not explicit) components of self-agency. *Consciousness and Cognition*, *17*(4), 133–151. https://doi.org/10.1016/j.concog.2008.03.004

Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2012). An online neural substrate for sense of agency. *Cerebral Cortex*, *23*(5), 1031–1037.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239.

Churches, O., Damiano, C., Baron-Cohen, S., & Ring, H. (2012). Getting to know you: The acquisition of new face representations in autism spectrum conditions. *NeuroReport*, *23*(11), 668–672. https://doi.org/10.1097/WNR.0b013e3283556658

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. NJ: Hillsdale.

Constantino, J. N., Kennon-McGill, S., Weichselbaum, C., Marrus, N., Haider, A., Glowinski, A. L., … Jones, W. (2017). Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature*. https://doi.org/10.1038/nature22999

Conty, L., N’Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, *45*(13), 3024–3037. https://doi.org/10.1016/j.neuropsychologia.2007.05.017

Cravo, A. M., Claessens, P. M. E., & Baldo, M. V. C. (2009). Voluntary action and causality in temporal binding. *Experimental Brain Research*, *199*(1), 95–99.

Cravo, A. M., Claessens, P. M. E., & Baldo, M. V. C. (2011). The relation between action, predictability and temporal contiguity in temporal binding. *Acta Psychologica*, *136*(1), 157–166.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25*(2), 141–168.

Cumming, G., & Calin-Jageman, R. (2016). *Introduction to the new statistics: Estimation, open science, and beyond*. Routledge.

Dalmaso, M., Edwards, S. G., & Bayliss, A. P. (2016). Re-encountering individuals who previously engaged in joint gaze modulates subsequent gaze cueing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(2), 271.

Danquah, A. N., Farrell, M. J., & O'Boyle, D. J. (2008). Biases in the subjective timing of perceptual events: Libet et al.(1983) revisited. *Consciousness and Cognition*, *17*(3), 616–627.

David, N., Gawronski, A., Santos, N. S., Huff, W., Lehnhardt, F. G., Newen, A., & Vogeley, K. (2008). Dissociation between key processes of social cognition in autism: Impaired mentalizing but intact sense of agency. *Journal of Autism and Developmental Disorders*, *38*(4), 593–605. https://doi.org/10.1007/s10803-007-0425-x

David, N., Newen, A., & Vogeley, K. (2008). The “sense of agency” and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, *17*(2), 523–534.

David, N., Obhi, S., & Moore, J. W. (2015). Sense of agency: examining awareness of the acting self. *Frontiers in Human Neuroscience, 9*, 310.

Debruille, J. B. (2007). The N400 potential could index a semantic inhibition. *Brain Research Reviews*. https://doi.org/10.1016/j.brainresrev.2007.10.001

Decety, J., & Porges, E. C. (2011). Imagining being the agent of actions that carry different moral consequences: an fMRI study. *Neuropsychologia*, *49*(11), 2994–3001.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desantis, A., Hughes, G., & Waszak, F. (2012). Intentional binding is driven by the mere presence of an action and not by motor prediction. *PLoS One*, *7*(1), e29557.

Desantis, A., Waszak, F., & Gorea, A. (2016). Agency alters perceptual decisions about action-outcomes. *Experimental Brain Research*. https://doi.org/10.1007/s00221-016-4684-7

Dewey, J. A., & Knoblich, G. (2014). Do implicit and explicit measures of the sense of agency measure the same thing? *PloS One*, *9*(10), e110118.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576–585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001

Ebert, J. P., & Wegner, D. M. (2010). Time warp: Authorship shapes the perceived timing of actions and events. *Consciousness and Cognition*, *19*(1), 481–489.

Edwards, S. G., Stephenson, L. J., Dalmaso, M., & Bayliss, A. P. (2015). Social orienting in gaze leading: a mechanism for shared attention. *Proceedings of the Royal Society B: Biological Sciences, 282*(1812), 20151141.

https://doi.org/10.1098/rspb.2015.1141

Eggebrecht, A. T., Elison, J. T., Feczko, E., Todorov, A., Wolff, J. J., Kandala, S., … Estes, A. M. (2017). Joint Attention and Brain Functional Connectivity in Infants and Toddlers. *Cerebral Cortex*, *27*(3), 1709–1720.

Eimer, M. (2011). The face sensitive N170 component of the event-related potential. In J. V. Calder, Andrew J., Rhodes, Gillian, Johnson, Mark H., Haxby (Ed.), *The Oxford Handbook of face perception* (pp. 329–344). Oxford: Oxford University Press.

Ellis, H. D. (1975). Recognizing faces. *British Journal of Psychology*, *66*(4), 409–426.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, *24*(6), 581–604.

Eskenazi, T., Rueschemeyer, S.-A., de Lange, F. P., Knoblich, G., & Sebanz, N. (2015). Neural correlates of observing joint actions with shared intentions. *Cortex*, *70*, 90–100.

Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, *18*(2), 254–261.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14), 9602–9605.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, *5*(1), 39–60.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feuerriegel, D., Churches, O., Hofmann, J., & Keage, H. A. D. (2015). The N170 and face perception in psychiatric and neurological disorders: A systematic review. *Clinical Neurophysiology*, *126*(6), 1141–1158. https://doi.org/10.1016/j.clinph.2014.09.015

Freeman, S. F. N., Gulsrud, A., & Kasari, C. (2015). Brief report: linking early joint attention and play abilities to later reports of friendships for children with ASD. *Journal of Autism and Developmental Disorders*, *45*(7), 2259–2266.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694.

Frith, C. (2012). Explaining delusions of control: The comparator model 20 years on. *Consciousness and Cognition*, *21*(1), 52–54.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*(16), R724–R732.

Furlan, M., Smith, A. T., & Walker, R. (2015). Activity in the human superior colliculus relating to endogenous saccade preparation and execution. *Journal of Neurophysiology*, *114*(2), 1048–1058.

Galfano, G., Sarlo, M., Sassi, F., Munafò, M., Fuentes, L. J., & Umiltà, C. (2011). Reorienting of spatial attention in gaze cuing is reflected in N2pc. *Social Neuroscience*, *6*(3), 257–269. https://doi.org/10.1080/17470919.2010.515722

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14–21.

Gallagher, S. (2012). Multiple aspects in the sense of agency1. *New Ideas in*

Psychology, *30*(1), 15–31.

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, *17*(4), 160–165.

Gamer, M., Schmitz, A. K., Tittgemeyer, M., & Schilbach, L. (2013). The human amygdala drives reflexive orienting towards facial features. *Current Biology*. https://doi.org/10.1016/j.cub.2013.09.008

Gillespie-Lynch, K. (2013). Response to and initiation of joint attention: overlapping but distinct roots of development in autism. *OA Autism*, *1*(2), 13.

Goods, K. S., Ishijima, E., Chang, Y.-C., & Kasari, C. (2013). Preschool based JASPER intervention in minimally verbal children with autism: Pilot RCT. *Journal of Autism and Developmental Disorders*, *43*(5), 1050–1056.

Goot, M. H., Tomasello, M., & Liszkowski, U. (2014). Differences in the nonverbal requests of great apes and human infants. *Child Development*, *85*(2), 444–455.

Gordon, I., Eilbott, J. A., Feldman, R., Pelphrey, K. A., & Vander Wyk, B. C. (2013). Social, reward, and attention brain networks are involved when online bids for joint attention are met with congruent versus incongruent responses. *Social Neuroscience*, *8*(6), 544–554. https://doi.org/10.1080/17470919.2013.832374

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–549.

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: a longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*(6), 839–848.

Gregory, S. E. A., & Jackson, M. C. (2017). Joint attention enhances visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 237.

Gregory, S. E. A., & Jackson, M. C. (2018). Barriers block the effect of joint attention on working memory: Perspective taking matters. *Journal of Experimental Psychology: Learning, Memory and Cognition*. doi: 10.1037/xlm000622.

Grgič, R. G., Crespi, S. A., & De’Sperati, C. (2016). Assessing self-awareness through gaze agency. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0164682

Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-color and multiple-color visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 86.

Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. *Autism*, *18*(4), 346–361.

Guastella, A. J., & Hickie, I. B. (2016). Oxytocin treatment, circuitry, and autism: a critical review of the literature placing oxytocin into the autism context. *Biological Psychiatry*, *79*(3), 234–242.

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

Haggard, P., & Clark, S. (2003). Intentional action: conscious experience and neural prediction. *Consciousness and Cognition*, *12*(4), 695–707.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385.

Hamilton, A. F. de C. (2016). Gazing at me: the importance of social meaning in understanding direct-gaze cues. *Philosophical Transactions of the Royal Society*

B: Biological Sciences. https://doi.org/10.1098/rstb.2015.0080

Happé, F., & Frith, U. (2014). Annual research review: towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry*, *55*(6), 553–577.

Haxby, J. V, Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233.

Haxby, J. V, Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*(1), 59–67.

Hedley, D., Brewer, N., & Young, R. (2011). Face recognition performance of individuals with Asperger syndrome on the Cambridge face memory test. *Autism Research*, *4*(6), 449–455. https://doi.org/10.1002/aur.214

Hietanen, J. K., Leppänen, J. M., Nummenmaa, L., & Astikainen, P. (2008). Visuospatial attention shifts by gaze and arrow cues: An ERP study. *Brain Research*, *1215*, 123–136. https://doi.org/10.1016/j.brainres.2008.03.091

Hileman, C. M., Henderson, H., Mundy, P., Newell, L., & Jaime, M. (2011). Developmental and individual differences on the P1 and N170 ERP components in children with and without autism. *Developmental Neuropsychology*, *36*(2), 214–236. https://doi.org/10.1080/87565641.2010.549870

Hoehl, S., Wiese, L., & Striano, T. (2008). Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLoS One*, *3*(6), e2389.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*(1), 80–84.

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131–134.

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, *139*(1), 133–151. https://doi.org/10.1037/a0028566

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude estimation reveals temporal binding at super-second intervals. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1542.

Humphreys, G. R., & Buehner, M. J. (2010). Temporal binding of action and effect in interval reproduction. *Experimental Brain Research*, *203*(2), 465–470.

Itier, R. J., Alain, C., Sedore, K., & McIntosh, A. R. (2007). Early Face Processing Specificity: It's in the Eyes! *Journal of Cognitive Neuroscience*, *19*(11), 1815–1826. https://doi.org/10.1162/jocn.2007.19.11.1815

Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience and Biobehavioral Reviews*, *33*(6), 843–863. https://doi.org/10.1016/j.neubiorev.2009.02.004

Jaswal, V. K., & Akhtar, N. (2018). Being vs. Appearing Socially Uninterested: Challenging Assumptions about Social Motivation in Autism. *The Behavioral and Brain Sciences*, 1-84.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clinical Neurophysiology*, *116*(11), 2613–2631.

Kaisler, R. E., & Leder, H. (2016). Trusting the Looks of Others: Gaze Effects of Faces in Social Settings. *Perception*, 45(8), 875-892. https://doi.org/10.1177/0301006616643678

Kaminski, J., Bräuer, J., Call, J., & Tomasello, M. (2009). Domestic dogs are sensitive to a human’s perspective. *Behaviour*, *146*(7), 979–998.

Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2005). Domestic goats, Capra hircus, follow gaze direction and use social cues in an object choice task. *Animal Behaviour*, *69*(1), 11–18.

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–2128. https://doi.org/10.1098/rstb.2006.1934

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., … Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring: A PET study. *Brain*, *122*(4), 779–783.

Khalighinejad, N., Di Costa, S., & Haggard, P. (2016). Endogenous Action Selection Processes in Dorsolateral Prefrontal Cortex Contribute to Sense of Agency: A Meta-Analysis of tDCS Studies of “Intentional Binding.” *Brain Stimulation*. https://doi.org/10.1016/j.brs.2016.01.005

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(32), 10153–10159. https://doi.org/10.1523/JNEUROSCI.2668-09.2009

Kim, K., & Mundy, P. (2012). Joint Attention, Social-Cognition, and Recognition Memory in Adults. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2012.00172

Kirchner, J. C., Hatri, A., Heekeren, H. R., & Dziobek, I. (2011). Autistic symptomatology, face processing abilities, and eye fixation patterns. *Journal of Autism and Developmental Disorders*, *41*(2), 158–167.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*. https://doi.org/10.1037/h0043688

Knöll, J., Morrone, M. C., & Bremmer, F. (2013). Spatio-temporal topography of saccadic overestimation of time. *Vision Research*. 83, 56-65. https://doi.org/10.1016/j.visres.2013.02.013

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution*, *40*(5), 419–435. https://doi.org/10.1006/jhev.2001.0468

Koike, T., Tanabe, H. C., Okazaki, S., Nakagawa, E., Sasaki, A. T., Shimada, K., … Sadato, N. (2016). Neural substrates of shared attention as social memory: A hyperscanning functional magnetic resonance imaging study. *NeuroImage*, *125*, 401–412. https://doi.org/10.1016/j.neuroimage.2015.09.076

Kok, A. (2001). On the utility of P300 amplitude as a measure of processing capacity. *Psychophysiology, 38*, 557–577.

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition*, *172*, 46–58.

Kuhn, G., Pagano, A., Maani, S., & Bunce, D. (2015). Age-related decline in the reflexive component of overt gaze following. *The Quarterly Journal of Experimental Psychology*, *68*(6), 1073–1081.

Kühn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. *Cortex*, *49*(7), 1935–1942.

Kunde, W., Weller, L., & Pfister, R. (2017). Sociomotor action control. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1316-6

Landes, T. L., Kashima, Y., & Howe, P. D. L. (2016). Investigating the effect of gaze cues and emotional expressions on the affective evaluations of unfamiliar faces. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0162695

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50–59.

Latinus, M., Love, S. A., Rossi, A., Parada, F. J., Huang, L., Conty, L., … Puce, A. (2014). Social decisions affect neural activity to perceived dynamic gaze. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsv049

Lawton, K., & Kasari, C. (2012). Teacher-implemented joint attention intervention: Pilot randomized controlled study for preschoolers with autism. *Journal of Consulting and Clinical Psychology*, *80*(4), 687.

Leavens, D., & Racine, T. P. (2009). Joint attention in apes and humans: are humans unique? *Journal of Consciousness Studies, 16*(6–1), 240–267.

Lee, K. H., & Siegle, G. J. (2009). Common and distinct brain networks underlying explicit emotional evaluation: a meta-analytic study. *Social Cognitive and Affective Neuroscience*, *7*(5), 521–534.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subjects designs. *Psychon B Rev*, *1*(4), 476–490.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*(April), 1–14. https://doi.org/10.3389/fnhum.2014.00213

Luck S. J. (2012). Electrophysiological Correlates of the Focusing of Attention within Complex Visual Scenes: N2pc and Related ERP Components. In S.J Luck & E.S Kappenman (Eds.). *The Oxford Handbook of Event-Related Potential Components* (3-30). New York: Oxford University Press. doi.org/10.1093/oxfordhb/9780195374148.013.0161

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.

Lundqvist, D., Flykt, A., & Öhman, A. (n.d.). *The Karolinska Directed Emotional Faces (KDEF). CD ROM from Department of Clinical Neuroscience. Psychology section, Karolinska Institutet; 1998*. ISBN 91-630-7164-9.

McNeil, J. E., & Warrington, E. K. (1993). Prosopagnosia: A Face-specific Disorder. *The Quarterly Journal of Experimental Psychology Section A*, *46*(1), 1–10. https://doi.org/10.1080/14640749308401064

Merchant, H., & Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences*, *8,* 22–27.

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, *17*(1), 136–144.

Moore, J. W., Lagnado, D., Deal, D. C., & Haggard, P. (2009). Feelings of control: contingency determines experience of action. *Cognition*, *110*(2), 279–283.

Moore, J. W., Middleton, D., Haggard, P., & Fletcher, P. C. (2012). Exploring implicit and explicit aspects of sense of agency. *Consciousness and Cognition*, *21*(4), 1748–1753.

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and Cognition*, *21*(1), 546–561.

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, *18*(4), 1056–1064.

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6-through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, *21*(3), 283–298.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience*. https://doi.org/10.1038/nn1488

Mundy, P. (2017). A Review of Joint Attention and Social-Cognitive Brain Systems in Typical Development and Autism Spectrum Disorder. *European Journal of Neuroscience*, (December 2016), 1–18. https://doi.org/10.1111/ejn.13720

Mundy, P., Kim, K., Mcintyre, N., & Lerro, L. (n.d.). Brief Report: Joint Attention and Information Processing in Children with Higher Functioning Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-016-2785-6

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

Mundy, P., Sullivan, L., & Mastergeorge, A. M. (2009). A parallel and distributed-processing model of joint attention, social cognition and autism. *Autism Research*, *2*(1), 2–21.

Murza, K. A., Schwartz, J. B., Hahs-Vaughn, D. L., & Nye, C. (2016). Joint attention interventions for children with autism spectrum disorder: A systematic review and meta-analysis. *International Journal of Language and Communication Disorders*, *51*(3), 236–251. https://doi.org/10.1111/1460-6984.12212

Myllyneva, A., & Hietanen, J. K. (2016). The dual nature of eye contact: To see and to be seen. *Social Cognitive and Affective Neuroscience*, *11*(7), 1089–1095. https://doi.org/10.1093/scan/nsv075

Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology*, *20*(01), 79–97. https://doi.org/10.1017/S0954579408000047

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*(3), 135–143. https://doi.org/10.1016/j.tics.2008.12.006

O'Hearn, K., Schroer, E., Minshew, N., & Luna, B. (2010). Lack of developmental improvement on a face memory task during adolescence in autism. *Neuropsychologia*, *48*(13), 3955–3960.

O'shea, J., Muggleton, N. G., Cowey, A., & Walsh, V. (2004). Timing of target discrimination in human frontal eye fields. *Journal of Cognitive Neuroscience*, *16*(6), 1060–1067.

Oberwelland, E., Schilbach, L., Barisic, I., Krall, S. C., Vogeley, K., Fink, G. R., … Schulte-Rüther, M. (2016). Look into my eyes: Investigating joint attention using interactive eye-tracking and fMRI in a developmental sample. 130, 248-

260. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2016.02.026

Obhi, S. S. (2012). The troublesome distinction between self-generated and externally triggered action: a commentary on Schüür and Haggard. *Consciousness and Cognition*, *21*(1), 587–588.

Obhi, S. S., & Hall, P. (2011). Sense of agency and intentional binding in joint action. *Experimental Brain Research*, *211*(3–4), 655–662. https://doi.org/10.1007/s00221-011-2675-2

Parker, K. J., Oztan, O., Libove, R. A., Sumiyoshi, R. D., Jackson, L. P., Karhson, D. S., … Phillips, J. M. (2017). Intranasal oxytocin treatment for social deficits and biomarkers of response in children with autism. *Proceedings of the National Academy of Sciences*, 201705521.

Pell, P. J., Mareschal, I., Calder, A. J., von dem Hagen, E. A., Clifford, C. W., Baron-Cohen, S., & Ewbank, M. P. (2016). Intact priors for gaze direction in adults with high-functioning autism spectrum conditions. *Molecular Autism*, *7*(1), 25.

Perrett, D. I., & Emery, N. J. (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 13(5), 683-694.

Perrett, D. I., Hietanen, J. K., Oram, M. W., Benson, P. J., & Rolls, E. T. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *335*(1273), 23–30.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London B: Biological Sciences*, *223*(1232), 293–317.

Pfeiffer, U. J., Schilbach, L., Jording, M., Timmermans, B., Bente, G., & Vogeley, K. (2012). Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in Psychology*, *3*.

Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., & Vogeley, K. (2014). Why we interact: on the functional role of the striatum in the subjective experience of social interaction. *NeuroImage*, *101*, 124–137.

Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience & Biobehavioral Reviews*, *37*(10), 2516–2528.

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, *209*(4), 481–493. https://doi.org/10.1007/s00221-011-2579-1

Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition*, *16*(2), 241–254.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., & Hietanen, J. K. (2011). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Social Cognitive and Affective Neuroscience*, *6*(4), 486–494. https://doi.org/10.1093/scan/nsq068

Poonian, S. K., & Cunnington, R. (2013). Intentional binding in self-made and observed actions. *Experimental Brain Research, 229*(3), 419–427.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Psouni, E., Falck, A., Boström, L., Persson, M., Sidén, L., & Wallin, M. (2018). Together I Can! Joint Attention Boosts 3-to 4-Year-Olds' Performance in a Verbal False-Belief Test. *Child Development*. 1-16. https :// doi.org /10.1111/ cdev.13075

Redcay, E., Kleiner, M., & Saxe, R. (2012). Look at this: the neural correlates of initiating and responding to bids for joint attention. *Frontiers in Human Neuroscience*, *6*(June), 1–14. https://doi.org/10.3389/fnhum.2012.00169

Redcay, E. & Saxe., R. (2013). Do you see what I see? The neural bases of joint attention. In J.M Metcalfe & H.S Terrace (Eds.) *Agency and Joint Attention*, (206-215). New York: Oxford University Press.

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, *40*(2), 149–166.

Rhodes, G., Jeffery, L., Taylor, L., & Ewing, L. (2013). Autistic traits are linked to reduced adaptive coding of face identity and selectively poorer face recognition in men but not women. *Neuropsychologia*, *51*(13), 2702–2708.

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*(5), 1281–1302. https://doi.org/10.1037/a0037004

Rigoulot, S., D'Hondt, F., Defoort-Dhellemmes, S., Despretz, P., Honoré, J., & Sequeira, H. (2011). Fearful faces impact in peripheral vision: Behavioral and neural evidence. *Neuropsychologia*, *49*(7), 2013–2021. https://doi.org/10.1016/j.neuropsychologia.2011.03.031

Rochat, P., & Striano, T. (2000). Perceived self in infancy. *Infant Behavior and Development*, *23*(3–4), 513–530.

Rosati, A. G., Arre, A. M., Platt, M. L., & Santos, L. R. (2016). Rhesus monkeys show human-like changes in gaze following across the lifespan. In *Proc. R. Soc. B* (Vol. 283, p. 20160376). The Royal Society.

Rossion, B. (2014). Understanding face perception by means of human electrophysiology. *Trends in Cognitive Sciences*, *18*(6), 310–318. https://doi.org/10.1016/j.tics.2014.02.013

Rossion, B. (2018). Humans are visual experts at unfamiliar face recognition. *Trends in Cognitive Sciences*, *22*(6), 471–472.

Ruess, M., Thomaschke, R., & Kiesel, A. (2018). Intentional binding of visual effects. *Attention, Perception, & Psychophysics*, 1–10.

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin and Review*, *16*(2), 252–257. https://doi.org/10.3758/PBR.16.2.252

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism*, *6*(1), 2. https://doi.org/10.1186/2040-2392-6-2

Sai, F., & Bushnell, I. W. R. (1988). The perception of faces in different poses by 1-month-olds. *British Journal of Developmental Psychology*, (6), 35–41.

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.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*(2), 235–239.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266.

Schilbach, L. (2015). Eye to eye, face to face and brain to brain: Novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Current Opinion in Behavioral Sciences*. 3, 130-135. https://doi.org/10.1016/j.cobeha.2015.03.006

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., … Vogeley, K. (2010). Minds Made for Sharing: Initiating Joint Attention Recruits Reward-related Neurocircuitry. *Journal of Cognitive Neuroscience*, *22*(12), 2702–2715. https://doi.org/10.1162/jocn.2009.21401

Schmitz, J., Scheel, C. N., Rigon, A., Gross, J. J., & Blechert, J. (2012). You don't like me, do you? Enhanced ERP responses to averted eye gaze in social anxiety. *Biological Psychology*, *91*(2), 263–269. https://doi.org/10.1016/j.biopsycho.2012.07.004

Schmitz, T. W., & Johnson, S. C. (2007). Relevance to self: A brief review and framework of neural systems underlying appraisal. *Neuroscience & Biobehavioral Reviews*, *31*(4), 585–596.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *42*, 9–34.

Schweinberger, S. R., Kloth, N., & Jenkins, R. (2007). Are you looking at me? Neural correlates of gaze adaptation. *NeuroReport*, *18*(7), 693–696. https://doi.org/10.1097/WNR.0b013e3280c1e2d2

Sebanz, N., & Lackner, U. (2007). Who's calling the shots? Intentional content and feelings of control. *Consciousness and Cognition*, *16*(4), 859–876.

Segal, N. L., Goetz, A. T., & Maldonado, A. C. (2016). Preferences for visible white sclera in adults, children and autism spectrum disorder children: implications of the cooperative eye hypothesis. *Evolution and Human Behavior*, *37*(1), 35–39.

Senju, A., Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience*, 1(3-4), 220-234. https://doi.org/10.1080/17470910600989797

Shteynberg, G. (2015). Shared attention. *Perspectives on Psychological Science*, *10*(5), 579–590.

Sperduti, M., Pieron, M., Leboyer, M., & Zalla, T. (2014). Altered pre-reflective sense of agency in autism spectrum disorders as revealed by reduced intentional binding. *Journal of Autism and Developmental Disorders*, *44*(2), 343–352. https://doi.org/10.1007/s10803-013-1891-y

Stavropoulos, K. K. M., & Carver, L. J. (2013). Research review: social motivation and oxytocin in autism–implications for joint attention development and intervention. *Journal of Child Psychology and Psychiatry*, *54*(6), 603–618.

Stephenson, L. J., Edwards, S. G., Howard, E. E., & Bayliss, A. P. (2018). Eyes that bind us: Gaze leading induces an implicit sense of agency. *Cognition*, *172*(December 2017), 124–133. https://doi.org/10.1016/j.cognition.2017.12.011

Stetson, C., Cui, X., Montague, P. R., & Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron*, *51*(5), 651–659.

Stewart, M. E., Griffiths, T. D., & Grube, M. (2015). Autistic traits and enhanced perceptual representation of pitch and time. *Journal of Autism and Developmental Disorders*, 1–9.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*. 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*. https://doi.org/10.3389/fpsyg.2013.00127

Takahashi, K., & Watanabe, K. (2013). Gaze cueing by pareidolia faces. *I-Perception*, *4*(8), 490–492.

Thiel, C. M., Zilles, K., & Fink, G. R. (2004). Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: an event-related fMRI study. *Neuroimage*, *21*(1), 318–328.

Thiel, C. M., Zilles, K., & Fink, G. R. (2005). Nicotine modulates reorienting of visuospatial attention and neural activity in human parietal cortex. *Neuropsychopharmacology*, *30*(4), 810–820.

Tipples, J., Johnston, P., & Mayes, A. (2013). Electrophysiological responses to violations of expectation from eye gaze and arrow cues. *Social Cognitive and Affective Neuroscience*, *8*(5), 509–514. https://doi.org/10.1093/scan/nss024

Tomasello, M., & Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development*, 70(1), i-152.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). In search of the uniquely human. *Behavioral and Brain Sciences*, *28*(5), 721–727.

Tomasello, M., & Farrar, M. J. (1986). Object permanence and relational words: A lexical training study. *Journal of Child Language*, *13*(3), 495–505.

Tomasello, M., Hare, B., & Fogleman, T. (2001). The ontogeny of gaze following in chimpanzees, Pan troglodytes, and rhesus macaques, Macaca mulatta. *Animal Behaviour*, *61*(2), 335–343.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320. https://doi.org/10.1016/j.jhevol.2006.10.001

Tomasello, M., & Herrmann, E. (2010). Ape and human cognition: What's the difference? *Current Directions in Psychological Science*, *19*(1), 3–8.

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, *64*, 231–255.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(1), 3–48.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*(1), 1–25.

Turati, C. (2004). Why Faces Are Not Special to Newborns: An Alternative Account of the Face Preference. *Current Directions in Psychological Science*, *13*(1), 5–8. https://doi.org/10.1111/j.0963-7214.2004.01301002.x

Velzen, J. van, & Eimer, M. (2003). Early posterior ERP components do not reflect the control of attentional shifts toward expected peripheral events. *Psychophysiology*, *40*(5), 827–831.

Wallace, G. L., & Happé, F. (2008). Time perception in autism spectrum disorders. *Research in Autism Spectrum Disorders*, *2*(3), 447–455. https://doi.org/10.1016/j.rasd.2007.09.005

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*(3), 640–663.

Watanabe, S., Kakigi, R., Miki, K., & Puce, A. (2006). Human MT/V5 activity on viewing eye gaze changes in others: a magnetoencephalographic study. *Brain Research*, *1092*(1), 152–160.

Watanabe, S., Miki, K., & Kakigi, R. (2002). Gaze direction affects face perception in humans. *Neuroscience Letters*, *325*(3), 163–166. https://doi.org/10.1016/S0304-3940(02)00257-4

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: experiencing control over the movements of others. *Journal of Personality and Social Psychology*, *86*(6), 838.

Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: A review of behavioral studies. *Neuroscience and Biobehavioral Reviews*, *36*(3), 1060–1084. https://doi.org/10.1016/j.neubiorev.2011.12.008

Wen, W., Yamashita, A., & Asama, H. (2015). The influence of action-outcome delay and arousal on sense of agency and the intentional binding effect. *Consciousness and Cognition*, *36*, 87–95.

Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry*, *44*(3), 456–468.

Wolf, R. C., Philippi, C. L., Motzkin, J. C., Baskaya, M. K., & Koenigs, M. (2014). Ventromedial prefrontal cortex mediates visual attention during facial emotion recognition. *Brain*, *137*(6), 1772–1780.

Wolf, W., Launay, J., & Dunbar, R. I. M. (2016). Joint attention, shared goals, and social bonding. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12144

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

Woodman, G. F., Arita, J. T., & Luck, S. J. (2009). A cuing study of the N2pc component: An index of attentional deployment to objects rather than spatial locations. *Brain Research*, *1297*, 101–111. https://doi.org/10.1016/j.brainres.2009.08.011

Yabe, Y., Dave, H., & Goodale, M. A. (2017). Temporal distortion in the perception of actions and events. *Cognition*. https://doi.org/10.1016/j.cognition.2016.10.009

Yabe, Y., & Goodale, M. A. (2015). Time flies when we intend to act: temporal distortion in a go/no-go task. *Journal of Neuroscience*, *35*(12), 5023–5029.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature*, *414*(6861), 302.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141.

Young, A. W., & Burton, A. M. (2017a). Are We Face Experts? *Trends in Cognitive*

Sciences, 22(2), 100-110. https://doi.org/10.1016/j.tics.2017.11.007

Young, A. W., & Burton, A. M. (2017b). Recognizing Faces. *Current Directions in Psychological Science*, *26*(3), 212–217. https://doi.org/10.1177/0963721416688114

Appendices

Appendix A: Stephenson et al. (2018)

Eyes that Bind Us: Gaze Leading Induces an Implicit Sense of Agency

Lisa J. Stephenson, S. Gareth Edwards, Emma E. Howard, & Andrew P. Bayliss

University of East Anglia, United Kingdom

Corresponding author: Andrew P. Bayliss, School of Psychology, University of East Anglia, Norwich, NR4 7TJ, United Kingdom. Email: Andrew.P.Bayliss@uea.ac.uk.

Abstract

Humans feel a sense of agency over the effects their motor system causes. This is the case for manual actions such as pushing buttons, kicking footballs, and all acts that affect the physical environment. We ask whether initiating joint attention – causing another person to follow our eye movement – can elicit an implicit sense of agency over this congruent gaze response. Eye movements themselves cannot directly affect the physical environment, but joint attention is an example of how eye movements can indirectly cause social outcomes. Here we show that leading the gaze of an on-screen face induces an underestimation of the temporal gap between action and consequence (Experiments 1 and 2). This underestimation effect, named 'temporal binding,' is thought to be a measure of an implicit sense of agency. Experiment 3 asked whether merely making an eye movement in a non-agentic, non-social context might also affect temporal estimation, and no reliable effects were detected, implying that inconsequential oculomotor acts do not reliably affect temporal estimations under these conditions. Together, these findings suggest that an implicit sense of agency is generated when initiating joint attention interactions. This is important for understanding how humans can efficiently detect and understand the social consequences of their actions.

Keywords

Gaze leading, Joint attention, Sense of agency, Social cognition, Temporal binding.

Eyes that Bind Us: Gaze Leading Induces an Implicit Sense of Agency

1. Introduction

The effects our motor system have on the environment need to be accurately detected. Action monitoring in humans gives rise to a *sense of agency* whereby we become conscious of our own actions (Gallagher, 2000). Such actions might be grasping objects or pushing buttons. However, some of the most important actions we execute do not directly affect the non-social, physical world, but do affect the *social* world. That is, some actions lead to changes in other people's actions (e.g. Caspar, Christensen, Cleeremans, & Haggard, 2016). One such ubiquitous social action is that when we look somewhere, other humans may spontaneously reorient their own gaze in the same direction, thus establishing joint attention (Frischen, Bayliss & Tipper, 2007). Joint attention is an everyday but important example that shows that, although eye movements cannot directly affect inanimate objects (aside from modern emerging gaze-controlled technologies, Slobodenyuk, 2016), changes in our gaze direction can influence other *people*. Moreover, saccades are the most common action we perform; we foveate a new area of the visual field 3-5 times each second (Schiller, 1998). However, there is little evidence that saccades evoke a sense of agency in a similar way to manual actions. We, therefore, tested whether an implicit sense of oculomotor agency over a conspecific's gaze shift response emerges in joint attention.

Because eye movements are a special form of action, they may not necessarily engage the same mechanisms underpinning agency as those engaged by other effectors. Nevertheless, there is a clear advantage in having robust agency detection systems for social outcomes elicited by our own actions, so a common mechanism that generalises between all effectors and outcome types could also be posited. Efficiently detecting the social effects we have caused may be critical to understanding others' actions and support mental state ascription (Happé, Cook, & Bird, 2016). Thus, the importance of understanding the role for agency in social action is critical for the understanding of social cognition.

There is one recent paper that suggests that people can learn to understand the contingencies between their saccades and a bouncing ball stimulus on a screen (Grgič, Crespi, & de'Sperati, 2016), which is an initial piece of evidence that the effects of saccades can be explicitly self-attributed. However, explicitly measuring sense of agency does not provide a full picture and can be problematic. This is because explicit measures are somewhat limited as self-reported feelings of control over an action depend on the actor's own ability for introspection (Barlas & Obhi, 2013; David et al., 2008; Sebanz & Lackner, 2007). Moreover, as Gallagher (2012) points out, self-agency is not normally something of which we are typically aware. Explicit measures are further criticised for their susceptibility to response bias and impression management (Obhi, 2012). Because of this, an alternative is to measure sense of agency implicitly with a measure that does not ask the participant to introspect about their explicit experience of control. Inferring sense of agency from implicit measures of correlated, potentially underlying mechanisms, has been a revealing approach (Barlas & Obhi, 2013). This can be achieved by exploiting an effect known as temporal binding (Haggard, Clark, & Kalogeras, 2002), whereby perception of the temporal distance between act and outcome is compressed for self-generated acts, and relatively accurate when judging the gap between two non-self-related stimuli (Moore & Obhi, 2012, for a review). This is why the temporal binding effect is theorised to measure an implicit sense of agency (see Haggard, 2017, for review).

Here, we adopt a twofold approach of measuring the sense of agency: temporal binding (which we offer as an implicit measure of agency) and self-reported ratings of felt control (an explicit measure of agency). We considered this necessary because explicit measures and binding effects do not always correlate, suggesting they may not reflect the exact same processes (e.g. Dewey & Knoblich, 2014, but see Ebert & Wegner, 2010, where changes in temporal binding were found to be related to explicit self-reports of agency). This possible dissociation between explicit and implicit agency are incorporated into an optimal cue integration account where implicit agency operates at a sensorimotor level, whilst explicit agency emerges following higher level processing (see Synovik et al., 2013).

Relatedly, sense of agency may arise both from predictive model-based mechanisms and postdictive mechanisms (Blakemore, Wolpert, & Frith, 2002; Haggard, 2017; Synofzik, Vosgerau, & Voss, 2013). According to the predictive model, the sense of agency is produced when there is a match between the predicted and the actual sensory outcome from an action (Blakemore et al., 2002). The retrospective or postdictive model, however, conceptualises a comparison between the action's idea and action's effect and a sense of agency arises if they are similar (Chambon & Haggard, 2013). Moore, Wegner, and Haggard (2009) argued that different, and varied, agency cues are integrated to result in a sense of agency (e.g. consequences of actions and sensorimotor prediction). Moore, Middleton. Haggard, and Fletcher (2012) tested this by exploring whether explicit and implicit agency were modulated differently by sequential patterns of action and outcome. Their results supported a model in which explicit and implicit agency can be thought of as dissociable, but, they argued, the two are not completely independent systems. This is consistent with Synovik et al's (2013) optimal integration cue account in which explicit and implicit agency can both be included. Given this reviewed evidence, we aimed to measure the temporal binding effect associated with an implicit sense of agency and collect self-report explicit ratings of agency as a manipulation check.

The temporal binding phenomenon has been associated with implicit sense of agency over physical actions that cause auditory (e.g. Barlas & Obhi, 2014), and visual outcomes (Cravo, Claessens, & Baldo, 2011). Investigations of interpersonal agency have been more limited, though agency is recognised as a critical aspect of joint action (Sebanz, Bekkering, & Knoblich 2006). Some studies have demonstrated a sense of agency over others' actions during joint tasks (Obhi & Hall, 2011; Pfister, Obhi, Rieger, & Wenke, 2014), and by illusory agent misidentification (e.g. Wegner, Sparrow, & Winerman, 2004). Interpersonal dynamics can modulate agency (e.g. under social coercion, Caspar et al., 2016). Social outcomes of physical acts have been studied by Yoshie and Haggard (2013), who showed that the valence of human vocalisations that served as a consequence of their participants' actions modulated temporal binding (but see Moreton, Callan, & Hughes, 2016). These studies offer some evidence that a *social outcome* from a button press can elicit binding. In one version of this paradigm, participants are asked to replicate the time interval they have just experienced (e.g. Humphreys & Buehner, 2010). We apply this notion of social sense of agency, measured using a time interval reproduction paradigm, to a crucial component of social cognition – joint attention - a key way in which humans communicate.

The above-reviewed binding evidence suggests that the socio-affective consequences of actions are coded in a generally similar way to non-social outcomes. Previous studies have shown saccade control can be guided by action-outcome effects, albeit in a non- social context (e.g Huestegge & Kreutzfeldt, 2012; Riechelmann, Pieczykolan, Horstmann, Herwig, & Huestegge, 2017). Relatedly, one eye-tracking study demonstrated that action-effect associations are made by the oculomotor system within a social context (Herwig & Hortsmann, 2011). Participants learned that their saccades triggered changes to onscreen facial expressions and adjusted their saccade accordingly. When they anticipated their saccade would trigger a smiling face, saccades landed near the mouth region and when they anticipated triggering a frown, saccades landed near the eyebrow region. This revealing finding illustrates how oculomotor actions can be influenced by perceived outcomes within a social context.

The actions studied thus far in the temporal binding literature are mostly restricted to button presses (see Moore & Obhi, 2012, for a review). In joint attention, the initiating act is an eye movement, whereby the gaze leader looks at an object, and a follower orients their attention to the same object (Frischen et al., 2007). Recent work has shown that people more efficiently detect instances when their gaze has been followed (Edwards, Stephenson, Dalmaso, & Bayliss, 2015), and that leading others' gaze has consequences for subsequent interactions with those individuals (Bayliss et al., 2013; Dalmaso, Edwards & Bayliss, 2016). Having one's eyes followed may necessarily involve the generation of a sense of agency over another's congruent gaze response. Indeed, people do explicitly express a feeling of control (Pfeiffer et al., 2012) and naturalness (Bayliss et al., 2013) in such scenarios. Establishing with temporal binding that similar processes underpin implicit agency in social gaze orienting as with physical acts, would be an important advance in our understanding of how social attention operates. Specifically, such a finding could help to explain why noticing that someone else has followed your gaze to establish joint attention is such a powerful experience, despite it being a common occurrence (e.g. Edwards et al., 2015; Bayliss et al., 2013). That is, rather than merely detecting that one's gaze has been followed, we interpret the social response as a causal outcome of our initial action.

Alternatively, it may not be this straightforward. There are also reasons to think that social agency might operate very differently to non-social agency. We have an enormous amount of experience of our physical manipulations of objects in the environment producing temporally contiguous outcomes. For example, when we kick a ball, it immediately moves. Therefore, the temporal window within which we become aware that our actions have produced an outcome are easily predictable. However, when we produce an action in order to elicit an outcome in another person, the temporal contiguity of the outcome has much more variance, making it harder to predict (Kunde, Weller, & Pfister, 2017). For example, a person may not immediately respond to our request to pass us an object nor may they immediately respond to our gaze signals, if their attention was elsewhere. The variance inherent in social interactions is one reason why implicit agency might work differently in social compared with non-social contexts. On the one hand, the variance might mean that temporal binding effects associated with implicit sense of agency might not emerge at all because social agency detection relies on higher-level mechanisms such as Theory of Mind (Premack & Woodruff, 1978) to make sense of social cause-and-effect. On the other hand, the instability of social interactions might actually elicit very reliable effects because of the critical importance of social agency detection, which could be underpinned by a system flexible enough to tolerate the inherent variance. Therefore, whether saccades that cause a social outcome could elicit temporal binding associated with implicit agency is an interesting open question for work both on social cognition and action monitoring.

In two experiments, we tested the hypothesis that gaze leading elicits temporal binding, which is offered as a measure of an implicit sense of agency (see Haggard, 2017, for a review). Participants’ time interval reproductions between an object’s appearance and an onscreen face looking at that object were compared between two tasks: an active task when a gaze leading saccade was made to the object, and a passive task in which no such gaze leading was performed. Therefore, we predicted that we would find greater temporal binding when participants’ eyes were followed to an object (Active Gaze Leading conditions) than when no saccades to the object were made (Passive conditions). Our data are consistent with this hypothesis, providing evidence that an implicit sense of agency, inferred from temporal binding, is generated in the gaze leader when their gaze is followed, establishing joint attention. A third experiment examined whether making an eye movement alone could explain the temporal compression effects found in Experiments 1 and 2, but no reliable effects were detected.

2. Experiment 1

In Experiment 1, participants completed an interval reproduction task under three conditions manipulated within-subjects. In the active task, for which we predicted reliable temporal binding, participants replicated the time interval between an object's appearance, to which the participants were to immediately saccade, and the on-screen face's gaze shift towards the object. As typical for temporal binding paradigms, we compared performance in the 'active' condition with a 'passive' condition in which no action is made by the participant. In the "Passive Face Fixation" condition participants fixated the face throughout. To provide a further control against which to compare any binding effects in the active task, we added a "Passive Phase Scrambled Fixation" condition. Here, we replaced the face with a non-social stimulus. A strength of our design is that participants in all conditions estimated the temporal gap between the same two events – the object appearing and the main stimulus (a face in two of three conditions) changing. In the active condition, participants saccaded after the object's appearance, and were instructed that their saccade was the cause of the on-screen face moving its eyes. We also had participants complete the Autism Spectrum Quotient (AQ, Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), a self-reported measure of autism-like traits. In all experiments, we have reported how we determined our sample size, all data exclusions (if any), all manipulations and all measures.

2.1. Method

2.1.1. Participants

Thirty-two participants (mean age=20.6 years; 2 were men) completed the study in return for course credit. We determined our target sample size by considering our relevant observed effect sizes in a previous study using the interval reproduction task (d_z=.84-1.44; Howard, Edwards, & Bayliss, 2016) and from appraising the wider literature. Anticipating a large effect size $d_z = .8$, with $1-\beta = 0.95$ at $\alpha = .05$, would require $n = 23$. However, it seemed appropriate here to anticipate a potentially smaller effect size than typically observed in temporal binding experiments using non-social actions, given the inherent variance associated with social responses to our own actions. We therefore targeted a sample of n=32, as this is closer to those used by ourselves and others to address similar questions. Participants reported normal or corrected-to-normal vision. Ethical approval was granted by the School of Psychology Ethics Committee, University of East Anglia. All participants were drawn from the Psychology undergraduate programme, were naïve to the aims of the study and gave written, informed consent.

2.1.2. Stimuli

The female face stimulus was a grayscale photograph with a calm expression (280×374 pixels) taken from Bayliss, Bartlett, Naughtin and Kritikos (2011), and had three versions: eyes direct, eyes closed and looking right. The object stimuli set comprised eight objects commonly found in the kitchen (varying in size; see Bayliss et al., 2013). The centre of the face was located 5 cm left-of-centre onscreen. The objects were presented 11.5cm to the right of the face. For one of the three conditions, a phase-scrambled version of the face was produced, comprising a rectangle (280x374 pixels) with two smaller rectangles (37x26 pixels) placed where the eyes would be on the face. The smaller rectangles were phase scrambled versions of the face stimulus’ eye regions. Stimuli appeared on a black background and were presented using E-Prime 2.0 software (see Figure 1).

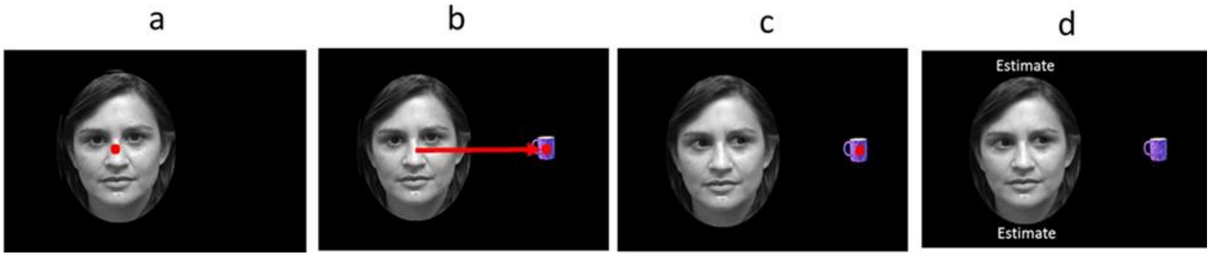

Fig. 1.

Trial sequence for the Active Gaze Leading task. Circles and the arrow were not displayed but represent where participants were instructed to fixate and the saccade from the face to the object, respectively. Participants looked at the face (a), displayed for 1000ms. Participants made a saccade (b) to the object as soon as it appeared. After a random inter-event interval of 400ms to 2300ms, gaze onset (c) occurred. After 1000ms, estimate instruction appeared (d) until response. Participants pressed and released the space bar to replicate the inter-event interval. The inter-event interval is the time between the object appearing and the gaze onset.

2.1.3. Apparatus and materials

Right eye position was tracked with an infrared eye tracker (Eyelink 1000, SR Research, Ontario, Canada; resolution 0.1^0, 500 Hz). A chin rest was used to maintain head stability. Viewing distance was 70cm from eyes to a 45 cm monitor (resolution 1024×768 pixels). A standard keyboard was used for manual responses. The Autism Spectrum Quotient Questionnaire was used as a measure of levels of autism-like traits (AQ; Baron-Cohen et al., 2001), presented using E Prime. A 1-8 scale was used for participants' self-reported feelings of agency in each condition, with 8 representing the highest feeling of agency.

2.1.4. Design

The within-subjects design had three blocked conditions of 56 trials per task. Block order was counterbalanced across participants. There were six possible orders with six participants experiencing one order, six participants undergoing another order, and the remaining four orders had five participants each. The conditions were Active Gaze Leading, Passive Face Fixation and Passive Phase Scrambled Fixation. The dependent measure was the proportional reproduction error (RE), calculated by dividing the reproduced time interval by the actual time interval to calculate mean proportional reproduction. Thus, 100% reproduction would be reproduction with no error at all. The inter-event interval was the time between an object's appearance and a subsequent on-screen gaze shift (Active Gaze Leading and Passive Face Fixation) or a spatial shift (Passive Phase Scrambled Fixation condition) towards the object. The temporal gap between the object's appearance (rather than the saccade) and the face's response was used to allow direct comparison between all conditions (as no saccades are made in passive conditions). We also had a correlational design to examine any associations between levels of AQ and degree of temporal binding.

2.1.5. Procedure

Each experimental block commenced with a standard nine point eye tracking calibration, then 8 practice trials, then 56 experimental trials (see Figure 1). In the Active Gaze Leading task, for which we predicted reliable temporal binding, each trial began with the presentation of the face on the left side of the screen, looking straight ahead. Participants were instructed to look at the face (presented for 1000 ms) until an object appeared on the right of the face. This sudden onset was the participant's cue to immediately saccade to it. Participants were told they must fixate on the object as soon as it appeared in the Active Gaze Leading task in order to cause the face to follow their gaze. Participants were instructed to fixate on the object after their gaze leading saccade, until the gaze shift occurred. After a randomly selected inter-event interval of 400-2300ms following the onset of the object, the face's gaze shifted to the right to look at the object. Participants were given no further instructions about where to look after their gaze leading saccade, apart from that they must maintain fixation on the object until the gaze shift occurred. After 1000ms, the word "Estimate" appeared (white font, Courier, 18pt) above and below the face. This prompted the participant to manually press and hold down the spacebar for a duration that to their best ability replicated the time interval between the object's appearance and the face's gaze shift towards it. Participants were given no feedback about their responses. Finally, after releasing the spacebar, the display cleared to black for 1000ms.

To be clear about the particulars of this 'Active' Gaze Leading condition, participants were told that their rapid saccade to the object was the causal event that made the face's eyes follow theirs. We were able to confirm that this was the impression that participants had with the explicit agency ratings task (details in Results section 2.2.2). We relied on the low variance of saccadic RT and spatial acuity in this very simple eye movement task to ensure that the minimum temporal gap of 400ms was greater than the vast majority of saccades. Moreover, timing the temporal gap from a single fixed onset that occurred in all conditions (the object onset) afforded us a straightforward and direct comparison across conditions.

The first control condition, in which we predict accurate temporal reproduction, was the ‘Passive Face Fixation’ task. This was identical to the Active Gaze Leading condition, except that 1) the participant maintained fixation throughout on the face, and 2) the face had closed eyes at the start of each trial before looking to the right following the appearance of the object. The final control condition, Passive Phase-scrambled task, used a rectangle comprised of the phase scrambled face, with two smaller, phase scrambled rectangular regions, which provided a spatial shift towards the object, instead of a gaze shift. The phase scrambled rectangles, positioned in the place the eyes would have been, shifted 2mm to the right after the inter-event interval. The size of the 2mm spatial shift was chosen as this was the same spatial shift as the eyes moved in the Active Gaze Leading condition. In both these passive control conditions, participants were instructed to fixate the face/phase-scrambled face throughout each trial, and replicate the interval between object onset and averted gaze onset. It was emphasised to them that they were not causing the gaze shift to occur. After each task (at the end of a 56 trial block) participants self-reported their degree of felt control over the face’s eye movements or the rectangles shifting. The instruction was “Please rate how much control you felt over the onscreen face’s eye movements/rectangles shifting from 1 to 8, 1 meaning no control at all to 8 meaning a lot of control.” Finally, participants completed the AQ on the computer.

2.2. Results

2.2.1. Proportional Reproduction

Trials in which participants' estimates were 3*SD*s above or below their individual means were removed (0.41% of trials). Mean proportional reproduction was calculated for each participant in each condition and submitted to statistical analysis (see Figure 2). We divided the reproduced time interval by the actual time interval to calculate mean proportional reproduction. Therefore, 100% reproduction represents perfect accuracy, anything greater than 100% is over-reproduction, and less than 100% is temporal compression (under-reproduction). We report Greenhaus-Geisser corrected degrees of freedom when applicable. Confidence intervals and standard errors around the means are based on 1000 bootstrap samples. We report confidence intervals around effect sizes and have used ESCI (Exploratory Software for Confidence Intervals) to calculate these (Cumming & Calin-Jageman, 2017).

First, in order to establish whether each condition produced temporal compression (reliable under-reproductions of the time between object and gaze onset), or relatively accurate reproductions, we performed single sample *t*-tests for each of the three conditions using proportional reproduction. This showed that temporal compression was only statistically significant in the Active Gaze Leading condition. Here, participants reproduced *M*=84% of the veridical time interval, 95% CI [73, 96] (*SD*=32%), *t*(31)=2.76, *p*=.01, d_z=0.69, 95% CI [0.18, 1.19]. In the two passive conditions, reproduction errors (REs) were low and did not differ statistically from 100% reproduction (Passive Face Fixation condition: *M*=100% reproduction, 95% CI [91, 112], *SD*=30%, *t*(31)=0.09, *p*=.926, d_z=0.02, 95% CI [-0.51,0.47]; Passive Phase-scrambled, *M*=94% reproduction, 95% CI [82, 100], *SD*=30%, *t*(31)=1.09, *p*=.286, d_z=0.27, 95% CI [-0.22; 0.76]. There was a main effect of task, *F*(1.53,47.42)=10.91, *MSE*=207, *p*<.001, η_p^2=0.260, and follow-up contrasts showed that the proportional temporal compression effect in the Active Gaze Leading condition was greater than in both the Passive Face Fixation, *t*(31)=3.73, *p*=.001, d_z=0.52, 95% CI [0.21,0.82] and

Passive Phase Scrambled Fixation conditions *t*(31)=3.17, *p*=.003, d_z=0.32, 95% CI [0.10,0.52]. Therefore, our hypothesis that having participants' deliberately-initiated saccade *followed* would result in greater temporal compression than passive conditions (where no saccades were made) was supported.

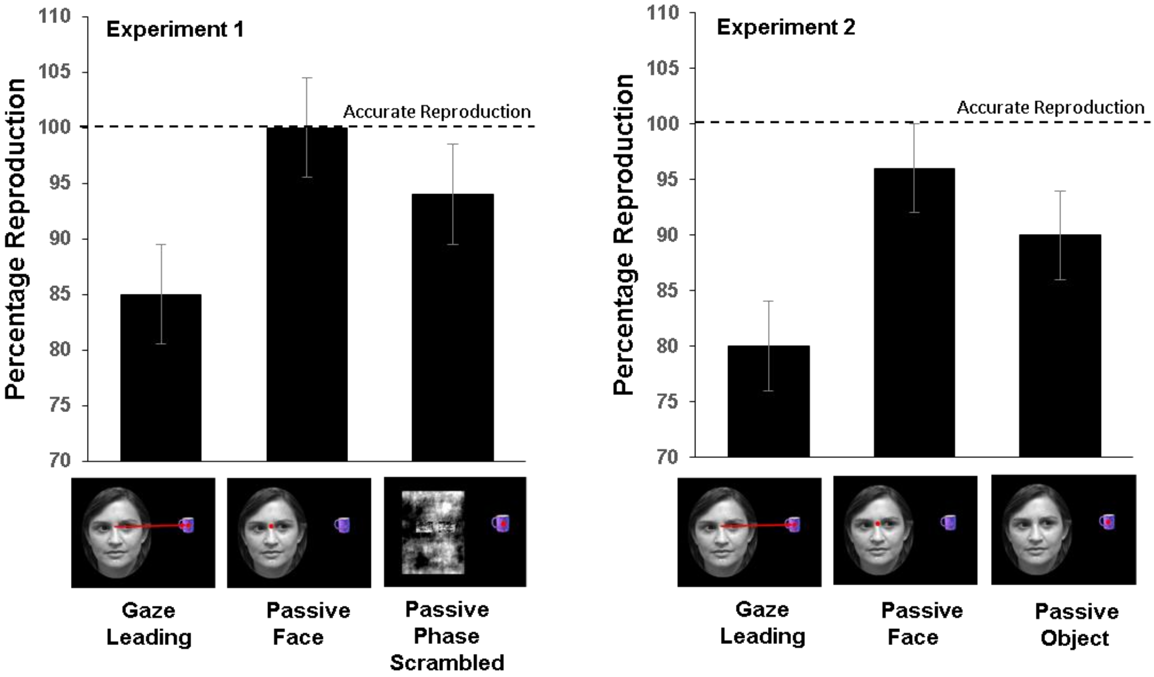

Fig. 2.
Mean percentage reproductions by condition for both experiments. In Gaze Leading tasks, participants looked first at the face, and then at an object as soon as it appeared. In the Passive Face or Passive Phase Scrambled tasks, participants looked at the face or scrambled face throughout. In the Passive Object task (Experiment 2), participants looked at the placeholder/object throughout. The images show how the face/scrambled stimulus was displayed when gaze onset occurred. Circles and the arrow were not displayed but represent where participants were instructed to fixate (and the saccade from the face to the object for the Active tasks). Error bars represent the standard error of the mean for within-subjects designs calculated using the procedure recommended by Loftus & Masson (1994).

2.2.2. Secondary measures, manipulation checks, and participant subset analyses

Mean self-reported explicit ratings of agency were greater for the Active Gaze Leading (*M*=4.44, *SD*=2.09), than both the Passive Face Fixation (*M*=2.25, *SD*=1.61) and Passive Phase Scrambled Fixation (2.03, *SD*=1.43) conditions; *t*'s>6, *p*'s<.001, d_z's>1**.** This shows that participants felt a degree of explicit agency in the Gaze Leading condition, supporting our inference that the temporal binding effect presented here reflects a sense of agency. The mean AQ score was 16.59 (*SD*=5.58), which is normative, and did not correlate significantly with reproduction error in any condition ($r< -.15$, $p>.4$).

We also considered potential concerns that something about performing a saccade per se might explain our data. Saccades can, indeed, affect time perception; a substantial amount of work has demonstrated an expansive effect (chronostasis; see review by Merchant & Yarrow, 2016), which if present in our data would of course increase our participants' estimates (i.e. this effect, if present, would work in opposition to our predicted and demonstrated effects). However, two studies have noted an opposing compressive effect (Morrone, Ross & Burr, 2005; Yabe & Goodale, 2015). These opposing effects are small and of similar magnitude so would cancel each other out were they to be present in our (rather different) task, so are unlikely to account for our data. In the critical Active Gaze Leading condition, mean saccadic reaction time was 220ms (*SD*=41ms) and mean saccade duration was 81ms (*SD*=44ms).

Further data exploration included checking for saccades executed after the onscreen face had moved its eyes, which was possible in our design. This could happen, for example, if the participant was rather slow on a trial with a short time interval. This could potentially affect the way that the participant perceived the agency of the social context. Such occurrences were present in nine participants, and on a maximum of three trials for a given participant (and a total of 0.7% of active trials). We reanalyzed the explicit and implicit data excluding all nine of these participants and found that the data pattern was very similar without these participants. Their mean explicit ratings are not different to those who never experienced this (*M*=4.5, *SD*=2.22 and M = 4.41, *SD*=2.15, respectively). Temporal compression was only statistically significant in the Active Gaze Leading condition. Here, participants reproduced *M*=84%, 95% CI [74,95] (*SD*=30%), of the veridical time interval $t(22)=2.49$, $p=.02$, $d_z=0.73$, 95% CI [0.13,1.3]. In the two passive conditions, reproduction errors were low and did not differ statistically from 100% reproduction, Passive Face Fixation condition: M=103%, 95% CI [93,113] SD=23%, $t(22)=0.597$, $p=.556$, $d_z=0.18$, 95% CI [-0.75,0.40]; Passive Phase-scrambled, M=98%, 95% CI [87,109], SD=25%, $t(22)=0.31$, $p=.763$, $d_z= 0.09$, 95% CI [-0.49,0.67].

To check whether passive tasks were compromised by saccades occurring contrary to the fixation instruction, we also examined erroneous saccades; on only 0.28% of trials were saccades made in error to the object during the Passive Face task and in 0.11% of trials in the Passive Scrambled condition. These few trials are unlikely to have had a critical impact on the data. Thus, overall, saccade metrics cannot parsimoniously explain the observed time underestimation in the Active task at the trial or participant levels.

As this is the first attempt to our knowledge using a temporal binding paradigm with saccades as the action, it is useful to examine whether our data share another commonality often observed in manual tasks in order to inform comparability across effectors. Previous temporal binding research using interval replication or estimation methodologies show stronger effects with longer intervals (Humphreys & Buehner, 2009; Wen, Yamashita, & Asama, 2015). In order to determine whether our data share this latter characteristic of the temporal binding phenomenon, we compared performance of each participant on the longer 50% of intervals they estimated with the shorter 50% of intervals they estimated. In order to establish whether this pattern is present in our data we instead used the *reproduction error* as the measure, calculated in milliseconds as the participants' reproduction of the temporal interval between two events minus the veridical temporal interval (rather than the proportion error used in the main analysis). The temporal compression effect was larger with the longer intervals, $t(31)=10.27$, $p<.001$, $d_z=1.75$. This corroborates the notion that the observed data reflects a temporal binding effect, rather than some form of previously unreported saccade-induced temporal discounting effect that would most likely be either proportional to saccade metrics, or in fact be stronger for short intervals, not weaker (given the timescale of saccades, and the timescale of previously observed interactions between saccades and time perception). We can, therefore, confidently assert this effect is temporal compression of a similar nature to that previously observed following manual actions that cause physical outcomes.

2.3. Discussion

Participants reliably under-reproduced the temporal gap between an object appearing in the periphery, and an on-screen face responding by looking towards the same object, only when participants moved their eyes to that object in the belief that they caused the face to follow their eyes. This is an indication that participants' eye movements resulted in an implicit sense of agency, the magnitude of which compares to temporal binding paradigms using manual actions that cause changes to the physical environment (Moore & Obhi, 2012). In both of our passive control conditions, our participants did not move their eyes to cause a social response, and they were rather accurate in their time reproductions. Therefore, we can be confident that the eye movement in the critical gaze leading condition caused the temporal compression associated with an implicit sense of agency.

3. Experiment 2

In Experiment 2 we sought to replicate the temporal binding effect in the Active Gaze Leading condition. It is notable that the Passive Face Fixation condition from Experiment 1 involved a face with closed eyes, whereas the Active Gaze Leading condition began the trials with direct gaze. This leaves open the possibility that this initial social contact of direct gaze is critical. To explore this, in Experiment 2, we instead had the active condition begin with closed eyes, and two passive control conditions begin with open eyes. One of the passive control conditions replicated that of Experiment 1, with face fixation throughout. The new passive control condition had participants gaze at the object throughout the trial, which allowed us to examine the importance of end-state gaze location. This was because we sometimes have our gaze followed after deliberate gaze leading, but we also have gaze followed incidentally when we happen to have been observed looking at an object. This is a scenario which is specifically found in a joint attention interaction, that is, gaze can be followed after deliberate gaze leading, but joint attention can result from a person following our passive attention to an object of interest, without any deliberate intention to engage in joint attention. It is, therefore, possible that agency may be experienced during joint attention when our gaze is followed incidentally, without a deliberate, gaze leading saccade. The new control condition enabled us to explore this possibility.

3.1. Method

A new sample of participants (*n*=32; mean age=19.7 years, four were men) was recruited from the same population as Experiment 1 and took part in return for course credits. The same stimuli were used as Experiment 1. The design involved changes to the three task conditions. The Active Gaze Leading condition was the same as Experiment 1 except that the onscreen face began each trial with closed eyes. The Passive Face Fixation task had the face commence with direct gaze. The new third condition, Passive Object Fixation, entailed the addition of a grey fixation dot (Courier, 18pt), which the participants were required to fixate at the start of each trial in this task and was where the object subsequently appeared. Therefore, in this Passive Object Fixation task, the onscreen gaze response occurred when participants were already looking at the object, not having first performed a gaze leading saccade to it. The procedure and task for participants was the same in all other respects for Experiment 2 as the previous experiment.

3.2. Results

3.2.1. Proportional Reproduction

Trials in which participants' estimates were 3*SD*s above or below their individual means were removed (0.28% of trials). The same processing and analysis was performed on the data as in Experiment 1. First, in order to establish whether each condition produced temporal compression (reliable under-reproductions of the time between object and gaze onset), or relatively accurate reproductions, we performed single sample *t*-tests for each of the three conditions on the proportional reproductions. This showed that temporal compression was statistically significant in the Active Gaze Leading condition. Here, participants reproduced the temporal gap by *M*=80%, 95% CI [73,86] (*SD*=19%), *t*(31)=6.18, *p*<.001, d_z=1.55, 95% CI [0.98, 2.10]. In the Passive Face condition, reproduction did not differ statistically from 100% reproduction (Passive Face Fixation condition: *M*=96%, 95% CI [88, 104], *SD*=23%, *t*(31)=1.00, *p*=.327, d_z=0.25, 95% CI [-0.24,0.74], but did in the Passive Object Fixation condition, *M*=90%, 95% CI [82,98], *SD*=22%, *t*(31)=2.70, *p*=.01, d_z=0.67, 95% CI [0.17;1.18]. There was a main effect of task, *F*(2,62) =21.45, *MSE* =.221, *p*<.001, η_p^2=0.409, and follow-up contrasts showed that the proportional temporal compression effect in the Active Gaze Leading condition was greater than in both the Passive Face Fixation, *t*(31)=6.02, *p*<.001, d_z=0.79, 95% CI [0.46, 1.11] and Passive Object conditions *t*(31)=4.17, *p*<001, d_z=0.51, 95% CI [0.23, 0.77].

3.2.2. Secondary measures, manipulation checks and participant subset analyses

As in Experiment 1, greater explicit agency was reported following the Active Gaze Leading (3.97, *SD*=1.79), than both the Passive Object Fixation (2.72, *SD*=1.57) and Passive Face Fixation (2.59, *SD*=1.50) conditions (*t*'s>3.6, *p*<.001, d_z's>0.7). The mean AQ score was 15.06 (*SD*=6.35), and did not correlate with reproduction error in any condition (r< -.15, *p*>.4). In the critical Active Gaze Leading condition, mean saccadic reaction time was 219ms (*SD*=57ms), and mean saccade duration for the gaze leading saccade was 79ms (*SD*=69). There were only 0.6% of trials where the onscreen face gaze shift occurred before the participant's saccade was completed. We performed the same check as Experiment 1, by re-analysing the data with the 9 participants excluded who experienced a gaze shift onscreen before their saccade was completed. This was for only an average of 1.22 trials. These nine participant's mean explicit ratings were not different to the rest of the sample (M = 3.66, SD=1.87 and M = 4.01, SD=1.75, respectively). The data showed a remarkably similar pattern. The Active Gaze Leading condition revealed temporal compression – participants reproduced 76%, 95% CI [68,84], *SD*=19% of the veridical time interval, $t(22)=6.12$, $p<.001$, $d_z=1.81$, 95% CI [1.11,2.48]. The Passive Face Fixation condition did not produce temporal compression (*M*=92% reproduction, 95% CI [82,101] *SD*=23%, $t(22)=1.77$, $p=.091$, $d_z=0.52$ 95% CI [-0.07,1.11]. However, the Passive Object Fixation task did reveal reliable under-reproductions, of about one third less than that in the active condition; *M*=84% reproduction, 95% CI [76,93] *SD*=19%, $t(22)=3.87$, $p=.001$, $d_z=1.14$, 95% CI [0.51,1.76].

Saccades to the object in error were made on only 0.33% of trials during the Passive Face task. In the Passive Object task of Experiment 2, saccades in error away from the object to the face were made on only 0.06% of trials. Therefore, passive tasks were not compromised by erroneous saccades, just like Experiment 1, as these were so small in number. We ran the same split half analysis of binding by temporal interval as Experiment 1, and again showed larger effects with the longer intervals, $t(31)=14.53$, $p<.001$, $d_z=2.57$, again supporting the notion that these are, indeed, temporal binding effects.

3.3. Discussion

We replicated both the binding effects for the Active Gaze Leading task and the null binding effects for the Passive Face Fixation task. Binding in the Passive Object Fixation task was significantly attenuated compared with the Active Gaze Leading task, but was nevertheless statistically reliable and is worthy of discussion so we address this further in the General Discussion below. For now, we note that there could perhaps be an implicit sense of agency (albeit reduced) which can be generated when there is a shift towards our object of gaze, even if we feel we have only incidentally caused the gaze shift, rather than intentionally.

4. Experiment 3

It is possible that saccades alone - devoid of social or agentic context - could produce binding. However, known saccade temporal disturbances have only previously been demonstrated at short intervals of around 100ms (e.g. Morrone et al., 2005), whilst ours are longer with an average of 1350ms. Nevertheless, it is worth checking if the mere oculomotor act of a saccade can produce similar effects. It is interesting to note that most temporal binding studies do not investigate whether a non-agentic *manual* action might produce distorted temporal judgements in and of themselves. However, because we know that saccades do produce some temporal distortion (Morrone et al., 2005; Yabe & Goodale, 2015), our approach affords an opportunity to explore this fundamental question. However, we also note here that, as our primary interest is in social cognition and agency, we look forward to further work being conducted on this question as it relates to core mechanisms of saccade control and temporal distortions because our single experiment may only provide indicative evidence one way or another. In Experiment 3, therefore, we tested two conditions with no social aspect or agentic expectation and predicted a null effect.

4.1 Method

A new sample of participants executed a saccade of the same amplitude as Experiments 1 and 2 between two fixation crosses in a Saccade task. They began fixation on a first cross and saccaded to a second cross, when it appeared. After the second cross appeared, the first cross enlarged. Participants then reproduced the interval between the second cross appearing and the first cross enlarging. In a No Saccade task, they maintained fixation on the first cross throughout, and reproduced the same time interval as the Saccade task. Thus, participants were exposed to a sequence of perceptual events, but none of these events were social, and they experienced both a saccade task with the same temporal and spatial characteristics of Experiments 1 and 2 and a no saccade task. Furthermore, they were given no information about whether their eye movements were causing anything to occur. This allowed us to test, for the first time to our knowledge, whether saccades alone – devoid of social context - can elicit temporal binding. A power analysis (GPower: Faul, Erdfelder, Lang, & Buchner, 2007) using the mean gaze leading effects from Experiments 1 and 2, found that n=29, would deliver 1-β power=0.95. Therefore, our final sample of n=31 (after removing one participant who did not follow instructions) was appropriate.

4.2 Results and Discussion

We found no significant under-reproduction in the Saccade Task, M=94%, 95% CI [79,109] (SD=40%), $t(30)$=0.81, p=.427, d_z=0.21, 95% CI [-0.29,0.70], nor in the No Saccade task, M=105%, 95% CI [95,115] (SD=27%) $t(30)$=0.983, p=.333, d_z=0.25, 95% CI [-0.75, 0.25]. As our prediction was for a null effect to emerge in the Saccade task, we aimed to assist the interpretability of this null by performing a Bayes one-sample t-test (Rouder, Speckman, Sun, Morey & Iverson, 2009), using the expected effect size parameter as the average effect size from the active conditions in Experiments 1 and 2 of 1.12. This produced a JZS BF=5.82 in favour of the null suggesting that, from these data, the null hypothesis is 5.82 times more likely than the alternative hypothesis. In addition, participants' ratings of explicit agency were low in both conditions; Saccade Task M=2.13 (SD=1.45) and the No Saccade Task M=2.10 (SD=1.64). In the Passive Fixation Cross task, saccades in error to the second fixation cross were made on only 0.95% of trials. Taken together, this suggests that the motor act of the eye movement itself is unlikely to account for the temporal compression effects we found in the social context of an interaction with an onscreen face.

5. General Discussion

We investigated the influence of gaze leading on the temporal compression effect known as temporal binding, which is associated with sense of agency. We showed, for the first time, that responses to our eye signals, like other motor actions, produce temporal binding within a simulated social interaction. This is offered as evidence for a form of oculomotor agency, which is informative for the understanding of social attention, and is more broadly of interest to the burgeoning field of technology with gaze-based interfaces (Slobodenyuk, 2016). Across four passive control conditions, we found no binding effects in three and an attenuated binding effect in the fourth. The explicit agency ratings supported our manipulation because greater ratings were made for active over passive tasks. We measured autism-like traits (AQ), but no relationship between binding and these were found. In a further control experiment, where fixation crosses replaced the face and object, we found no binding effects.

Given the importance of joint attention in human social interactions, and the fact that saccades do not - outside of the laboratory, or through certain assistive technologies - cause physical outcomes, it was sensible to first investigate joint attention. As it turned out, our data are typical for the temporal binding literature, so we would in fact predict that intentional saccades that cause a different type of social outcome, or even a non-social outcome, would also produce temporal binding. Our present data can therefore contribute to, and open up new questions for social cognition and for the role of agency in eye movements per se. Given the similarity of our data to that of studies investigating non-social agency, our data are consistent with a common mechanism which attributes agency for social and non-social outcomes. The confirmation that saccades can elicit binding is of general importance for a field in which most of the outcomes resulting in binding are a consequence of a button press (see Moore & Obhi, 2012, for a review). Relatedly, we note that in our active condition, the key saccade was voluntary, and it is therefore an interesting question as to whether or not reflexive exploratory saccades may drive similar agentic mechanisms.

Learned outcomes from saccades when exploring faces can feedback to elicit changes to subsequent interactions (Herwig & Hortsmann, 2011). Taking this together with our data, we can offer a conceptual framework in which agency is experienced for gaze responses, and this may be the mechanism needed for feedback to drive subsequent changes in saccadic behaviour. This would also help explain the changes in visual exploration people exhibit when inspecting faces with which they had previously engaged in joint attention (see Bayliss et al., 2013). This is also consistent with a theoretical framework of sociomotor action control offered by Kunde et al., (2017) whereby the social responses received from our actions feedback to plan subsequent social actions. Experiencing agency over the social responses to our actions is a prerequisite to that process. We need to detect agency over any gaze following we elicit in order to deduce whether we have successfully cued attention to the referent object, in order to then plan the on-going social engagement. Thus, detecting the influence that we have had over others' attentional states may be critical for everyday social interactions and even support theory of mind processes. Determining that mechanisms engaged via physical acts generalise to oculomotor agency adds to what we know about gaze leading in terms of attention (Edwards et al., 2015), and reward value (Schilbach et al., 2010; Gordon, Eilbott, Feldman, & Vander Wyk, 2013). Agency may be a key piece of the puzzle that supports joint action with co-ordination and cooperation (Sebanz & Knoblich, 2009).

The lack of binding in passive conditions shows that the mere presence of a social stimulus does not interfere greatly with accurate timing of intervals per se. However, the weaker but reliable binding effect in the Passive Object Fixation task of Experiment 2 is curious. This observation could merely reflect a carry-over effect from the active task blocks (given our repeated measures design). However, we examined those participants who completed the Passive Object task first, and found that the binding effect was present (*M*=87% reproduction) and of a similar magnitude to the binding effect for all participants (*M*=90%), so carry-over effects are an unlikely explanation for the effects we found. Therefore, a more interesting (but speculative) suggestion would be that object-oriented attention in the presence of a face gazing at the same object might affect time estimation, even in the absence of a recently preceding action. It could be the case that if we are looking at an object already, we may attribute some agency to an observed congruent eye shift; but the effect is stronger if we have *recently* saccaded to that object (as in the Active Gaze Leading condition). This chimes with work highlighting the critical importance of objects in joint attention (Bayliss & Tipper, 2006; Bayliss et al., 2013; Lobmaier, Fischer, & Schwaninger, 2006). It is perhaps this aspect of our data that might lead to future research into what might be ‘special’ about *social* agency – we can cause others to behave in a certain way due to our present state, or even because we have *not* acted. We need to detect these interactions as well. Therefore, there may be a hierarchical system which attributes the greatest sense of implicit agency for intentional gaze leading and then an attenuated sense of implicit agency if a gaze shift is detected when we are already directing our gaze towards an object incidentally. This notion implies the importance of causality, in addition to intentionality, in these effects (Buehner & Humphreys, 2009; Desantis, Hughes, & Waszak, 2012).

There are a host of boundary conditions that remain untested in order to establish the conditions necessary and sufficient to produce indices of implicit agency in social contexts. One important future condition to test is to establish whether the observed gaze response needs to be congruent with the participant’s saccadic action, or can be any response (e.g. to avert gaze, or to change emotional expression, for example). We speculate that possibly an incongruent gaze shift might elicit binding if we feel we have caused another to look away from our direction of interest. Whether this would be binding of the same magnitude as a congruent gaze shift (or no binding at all) would be interesting for future studies to explore. The current results identify just one instance in which temporal binding can occur following a causal eye movement. Although determining the specificity of this effect is of course important for understanding the nature of the mechanisms involved, if future work were to demonstrate that the effect does generalise widely, this would not necessarily reduce the direct importance of this mechanism for understanding how social cognition is supported by such basic sensorimotor mechanisms.

One potential complication for the interpretation of our findings is that in both active and passive conditions, participants must detect the onset of the object in their periphery (while they are looking at the face). However, in the active tasks, the onset of the responding gaze shift is to be detected in their periphery because the participant is now looking at the object having performed a saccade, while in the passive conditions, the participant detects the gaze shift at their point of fixation, having not moved their eyes. This difference could have affected the speed of detection of the gaze shift across conditions. However, were participants to be slower to detect the gaze shift in their peripheral vision in the active task, this would have extended their time estimations, which means that our binding effects may have, if anything, been artificially relatively reduced. Despite this difference potentially working against our predictions, medium (Experiment 1) and large (Experiment 2) binding effect sizes emerged.

Another notable aspect of our design is that we used closed eyes for the Passive Face task in Experiment 1 because we wanted to ensure participants could easily identify that the passive task was different to the active task (with open eyes), to ameliorate against potential carry-over effects. In Experiment 2, the face was depicted with closed eyes until averted gaze was displayed – no direct gaze towards the participant. The closed eyes at the outset could be interpreted as less agentic by participants, but this does not appear to be the case as explicit agency ratings were similar in both Experiments 1 and 2, as were the magnitude of binding effects (or even larger observed effect sizes in Experiment 2). We speculate that ambiguity may result in stronger attribution of agency when there is a spatial shift towards our direction of gaze. It may be adaptive to assume that we caused an outcome for which we believe – but are uncertain - that we were responsible for eliciting. The consequences of under-attribution of responsibility for a social outcome could be particularly costly, whilst a little over-self-attribution is unlikely to lead to adverse consequences. This explanation is consistent with recent findings reported by Desantis, Waszak, and Gorea (2016), who found that participants over-attribute self-agency when they are in an ambiguous situation. We suspect that this result may suggest that binding effects will emerge in instances where the end-point of joint gaze occurs (given that joint attention can be incidental, as well as deliberate – both of which are important to notice and interpret). This is another interesting line for future investigations with respect to social agency specifically.

Although the null effects on temporal estimation in Experiment 3 support the notion that the data from Experiment 1 and 2 do reflect a temporal binding effect in a social setting, it is worthwhile considering that one might have expected reliable temporal underestimation even in the context of a non-agentic, non-social saccade task of Experiment 3. Specifically, it is known that eye movements do lead to temporal underestimations (saccadic compression, e.g. Morrone et al., 2005), but this did not emerge clearly in Experiment 3 in our data. One explanation for this could be that the known saccadic-driven temporal effects may not be observable in the time intervals of the magnitude we employed here. Our temporal intervals varied around a mean of 1350ms, while the studies that have discovered saccade-triggered temporal distortions have typically employed much shorter intervals (~100ms, e.g. Morrone et al., 2005).

Another potential reason for the failure to observe this temporal compressive effect of saccades per se is possibly due to the action of an opposing temporally expansive process,‘chronostasis’, which could operate simultaneously under our experimental conditions leading to temporal equilibrium (see Merchant & Yarrow, 2016, for a review and see also Knöll, Morrone, & Bremmer, 2013; Yarrow et al., 2001). Achieving this equilibrium may be advantageous for spatio-temporal perceptual stability, and a naïve assumption would be that such equilibrium would emerge more readily after longer temporal intervals, hence we observed a null effect overall in Experiment 3. This is speculative, however, and it is clear that future explorations of the direct effects of saccades on timing estimates will assist with the contextualisation of our present data, and indeed with other work studying social cognition that involves interactive eye movements and other actions.

Future work could employ a gaze-contingent design to explore agency in social gaze interactions. The present work did not take this approach. If we had yoked more directly the action of the participant to the stimulus changes by using gaze-contingent stimuli, we could have expected our participants to report a greater explicit sense of agency than we found here, and the temporal binding effects might have also been more stable. We did not employ a gaze contingent design here because we wished to avoid the introduction of a confound. Specifically, in the Active Saccade task the to-be-estimated time interval would have included three periods of temporal lag that would not be present in the Passive conditions, making them not comparable without off-line adjustment. These lag periods are the saccade latency, the saccade duration and the eye-tracker uptake time to detect good fixation upon the object in order to cause the gaze shift. By not using gaze contingent stimuli, our chosen design afforded direct comparison of actual time intervals across conditions. Nevertheless, it is clear that future studies should employ gaze contingent designs that circumvent the issues we note above to overcome this limitation of the present research. This would allow for even more robust tests of hypotheses regarding the temporal dynamics of social gaze.

We found no reliable correlations between binding effects and autism quotient scores. It may nevertheless be important to test similar paradigms in clinical samples given previous findings of sub-optimality for joint attention initiation (Mundy & Newell, 2007), and decreased temporal binding effects in autism (Sperduti, Pieron, Leboyer, & Zalla, 2014). Relatedly, it is notable that some forms of psychosis, such as might be experienced by those with a diagnosis of schizophrenia, are associated with disrupted sense of agency (see Haggard, 2017, for a review). Therefore, this may generalise to problems with understanding other’s actions, which can be particularly problematic within the social setting of a joint attention interaction. These data are also of direct relevance for developers of gaze-controlled interfaces, a field that is currently grappling with issues of agency and control (Grgič et al., 2016; Slobodenyuk, 2016). For example, our findings can help inform research into making human-robot interactions more naturalistic when designing robots who can produce eye gaze responses to human gaze signals. Similarly, socially assistive robotics is a growing area where roboticists apply findings from cognitive science to inform the design of therapeutic interventions. Such interventions have been developed for a range of applications, including dementia, mental health, social communication for children with autism and stroke rehabilitation (see Matarić, 2017, for a review). Our research is also informative for developers of gaze-controlled interfaces more generally. Building on the boundary conditions for when eye movements can generate a similar sense of agency as other motor actions do, can inform how to make such technologies acceptable to users. Recent innovations of employing face/eye scanning in smartphones exemplify that using our eyes to control objects will soon be an everyday occurrence, so understanding oculomotor agency in social and non-social contexts is of direct relevance to medical and consumer product development.

To conclude, this study shows for the first time that temporal binding can occur when a social gaze response is perceived to result from intentional eye saccade bids for joint attention. We offer this as an implicit sense of agency effect that follows oculomotor actions that lead to a state of joint attention.

6. Author note

L.J. Stephenson, S. G. Edwards, E. E. Howard, and A.P. Bayliss, School of Psychology, University of East Anglia. Please address correspondence regarding this work to: Dr. Andrew Bayliss, School of Psychology, University of East Anglia, Norwich, NR4 7TJ, United Kingdom. Email: Andrew.P.Bayliss@uea.ac.uk.

Funding: This work was supported by University of East Anglia Studentships to L.J. Stephenson, S.G. Edwards and E.E. Howard, and by a Leverhulme Trust Project Grant RPG-2016-173 to A.P. Bayliss. Conflicts of interest: none.

7. Author Contributions

All authors developed the study concept and design. L.J. Stephenson collected and analysed the data. L.J. Stephenson and A.P. Bayliss interpreted the data and drafted the manuscript. All authors provided critical revisions. All authors approved the final version of the manuscript for submission.

References

Barlas, Z., & Obhi, S. S. (2013). Freedom, choice, and the sense of agency. *Frontiers in Human Neuroscience*, *7*. doi: 10.3389/fnhum.2013.00514.

Barlas, Z., & Obhi, S. S. (2014). Cultural background influences implicit but not explicit sense of agency for the production of musical tones. *Consciousness and Cognition*, 28, 94-103. doi:10.1016/j.concog.2014.06.013

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17. doi: 10.1023/A:1005653411471.

Bayliss, A. P., Bartlett, J., Naughtin, C. K., & Kritikos, A. (2011). A direct link between gaze perception and social attention. *Journal of Experimental Psychology: Human Perception and Performance*, 37(3), 634. doi:10.1037/a0020559.

Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. 2013). "Gaze leading": Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General*, 142(1), 76. doi: 10.1037/a0029286.

Bayliss, A. P., & Tipper, S. P. (2006). Gaze cues evoke both spatial and object-centered shifts of attention. *Attention, Perception, & Psychophysics, 68*(2), 310-318. doi: 10.3758 /BF03193678.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237-242. doi: 10.1016/S1364-6613 (02)01907-1.

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science, 20*(10), 1221-1228. doi: 10.1111/j.1467-9280.2009.02435.x.

Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion changes the sense of agency in the human brain. *Current Biology*, *26*(5), 585-592. doi: 10.10 16 /j.cub.2015.12.067.

Chambon, V., and Haggard, P. (2013). Premotor or ldeomotor: how does the experience of action come about? In W. Prinz, M. Beisert and A. Herwig (Eds.) *Action Science: Foundations of an Emerging Discipline* (359-380).Cambridge, MA: MIT Press.

Cravo, A. M., Claessens, P. M., & Baldo, M. V. (2011). The relation between action, predictability and temporal contiguity in temporal binding. *Acta Psychologica*, *136*(1), 157-166. doi:10.1016/j.actpsy.2010.11.005.

Cumming, G., & Calin-Jageman, R. (2017). *Introduction to the new statistics: Estimation, open science, and beyond*. New York and London: Routledge.

Dalmaso, M., Edwards, S. G., & Bayliss, A. P. (2016). Re-Encountering individuals who previously engaged in joint gaze modulates subsequent gaze cueing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(2), 271-284. doi: /10.1037/xlm0000159.

David, N., Newen, A., & Vogeley, K. (2008). The “sense of agency” and its underlying cognitive and neural mechanisms. *Consciousness and Cognition, 17*(2), 523-534. doi:10.1016/j.concog.2008.03.004.

Desantis, A., Hughes, G., & Waszak, F. (2012). Intentional binding is driven by the mere presence of an action and not by motor prediction. *PloS one, 7*(1). doi: 10.1371/journal.pone.0029557.

Desantis, A., Waszak, F., & Gorea, A. (2016). Agency alters perceptual decisions about action-outcomes. *Experimental Brain Research*, 1-9. doi: 10.1007/s00221-016-4684-7.

Dewey, J. A., & Knoblich, G. (2014). Do implicit and explicit measures of the sense of agency measure the same thing?. *PloS One*, *9*(10), e110118. 10.1371/ journal.pone.0110118.

Ebert, J. P., & Wegner, D. M. (2010). Time warp: Authorship shapes the perceived timing of actions and events. *Consciousness and Cognition*, *19*(1), 481-489. doi: 10.1016 /j.concog.2009.10.002.

Edwards, S. G., Stephenson, LJ.., Dalmaso, M., & Bayliss, A. P. (2015). Social orienting in gaze leading: A mechanism for shared attention. *Proceedings of the Royal Society B,* 82 (1812), 20151141. doi: 10.1098/rspb.2015.1141.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. doi: 10.3758 /BF03193146.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4), 694. doi: 10.1037/0033-2909.133.4.694.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences, 4*(1), 14-21. doi: 10.1016/S1364-6613(99)01417-5.

Gallagher, S. (2012). Multiple aspects in the sense of agency. *New Ideas in Psychology*, *30*(1), 15-31. doi: 10.1016/j.newideapsych.2010.03.003.

Gordon, I., Eilbott, J. A., Feldman, R., Pelphrey, K. A., & Vander Wyk, B. C. (2013). Social, reward, and attention brain networks are involved when online bids for joint attention are met with congruent versus incongruent responses. *Social Neuroscience, 8*(6), 544-554. doi: 10.1080/17470919.2013.832374.

Grgič, R. G., Crespi, S. A., & de'Sperati, C. (2016). Assessing Self-Awareness through Gaze Agency. *Plos One*, *11*(11), e0164682.doi:10.1371 /journal.pone.0164682.

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

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

Happé, F., Cook, J., & Bird, G. (2016). The Structure of Social Cognition: In (ter) dependence of Sociocognitive Processes. *Annual Review of Psychology*, *68*(1). doi: 10.1146/annurev-psych-010416-044046.

Herwig, A., & Horstmann, G. (2011). Action–effect associations revealed by eye movements. *Psychonomic Bulletin & Review, 18*(3), 531-537. doi: 10.3758 /s13423-011-0063-3.

Howard, E. E., Edwards, S. G., & Bayliss, A. P. (2016). Physical and mental effort disrupts the implicit sense of agency. *Cognition*, *157,* 114-125. doi: 10.1016/j. cognition.2016.08.018.

Huestegge, L., & Kreutzfeldt, M. (2012). Action effects in saccade control. *Psychonomic Bulletin & Review*, *19*(2), 198-203. doi: org/10.3758/s13423-011-0215-5.

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude estimation reveals temporal binding at super-second intervals. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1542. doi: 10.1037/a0014492.

Humphreys, G. R., & Buehner, M. J. (2010). Temporal binding of action and effect in interval reproduction. *Experimental Brain Research, 203*(2), 465-47. doi: 10.1 007/s00221-010-2199-1.

Knöll, J., Morrone, M. C., & Bremmer, F. (2013). Spatio-temporal topography of saccadic overestimation of time. *Vision Research, 83*, 56-65. doi: 10.1016/j.visres.2013.02.013.

Kunde, W., Weller, L., & Pfister, R. (2017). Sociomotor action control. *Psychonomic Bulletin & Review*, 1-15. doi: 10.3758/s13423-017-1316-6.

Lobmaier, J. S., Fischer, M. H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology, 53*(2), 117-122. doi: 10.1027/16183169. 53.2.117.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*(4), 476-490. doi: 10.3758/BF03210951.

Matarić, M. J. (2017). Socially assistive robotics: Human augmentation versus automation. *Science Robotics, 2*(4), eaam5410.

Merchant, H., & Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences, 8*, 22-27. doi: 10.1016/ j.cobeha. 2016.01.006.

Moreton, J., Callan, M. J., & Hughes, G. (2017). How much does emotional valence of action outcomes affect temporal binding?. *Consciousness and Cognition*, *49*, 25-34. doi:10.1016/j.concog.2016.12.008.

Moore, J. W., Middleton, D., Haggard, P., & Fletcher, P. C. (2012). Exploring implicit and explicit aspects of sense of agency. *Consciousness and Cognition*, *21*(4), 1748-1753. doi: 10.1016/j.concog.2012.10.005.

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

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, *18*(4), 1056-1064. doi: 10.1016/ j.concog.2009.05.004.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience*, *8*(7), 950-954. doi: 10.1038/nn1488.

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

Obhi, S. S. (2012). The troublesome distinction between self-generated and externally triggered action: A commentary on. *Consciousness and Cognition*, *21*(1), 587-588.

Obhi, S. S., & Hall, P. (2011). Sense of agency and intentional binding in joint action. *Experimental Brain Research*, *211*(3-4), 655-662. doi: 10.1007/s00221-011-2675-2.

Pfeiffer, U. J., Schilbach, L., Jording, M., Timmermans, B., Bente, G., & Vogeley, K. (2012). Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in Psychology, 3, 537.*

Pfister, R., Obhi, S. S., Rieger, M., & Wenke, D. (2014). Action and perception in social contexts: intentional binding for social action effects. *Frontiers in Human Neuroscience, 8.* doi: 10.3389/fnhum.2014.00667.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences, 1*(4), 515-526. doi: 10.1017/ S0140525X00076512.

Riechelmann, E., Pieczykolan, A., Horstmann, G., Herwig, A., & Huestegge, L. (2017). Spatio-temporal dynamics of action-effect associations in oculomotor control. *Acta Psychologica, 180,* 130-136. doi: 2017.09.003.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225-237. doi: org/10.3758/PBR.16.2.225.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., . . . Vogeley, K. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, 22(12), 2702-2715. doi: 10.1162/jocn.2009.21401.

Schiller, P. H. (1998). The neural control of visually guided eye movements. In: J. E. Richards (Ed.), *Cognitive Neuroscience of Attention* (pp. 3-50). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*(2), 353-367. doi: 10.1111/j.1756 8765.2009.01024.x.

Sebanz, N., & Lackner, U. (2007). Who's calling the shots? Intentional content and feelings of control. *Consciousness and Cognition*, *16*(4), 859-876. doi:10.1016/j.concog. 2006.08.002.

Slobodenyuk, N. (2016). Towards cognitively grounded gaze-controlled interfaces. *Personal and Ubiquitous Computing*, 1-13. doi: 10.1007/s00779-016-0970-4.

Sperduti, M., Pieron, M., Leboyer, M., & Zalla, T. (2014). Altered pre-reflective sense of agency in autism spectrum disorders as revealed by reduced intentional binding. *Journal of Autism and Developmental Disorders*, *44*(2), 343-352. doi: 10.1007/ s10803-013-1891-y.

Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: an interplay between prediction and postdiction. *Frontiers in Psychology, 4(127),* 1-8. doi: 10.3389/fpsyg.2013.00127.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: experiencing control over the movements of others. *Journal of Personality and Social Psychology, 86*(6), 838. doi: 10.1037/0022-3514.86.6.838.

Wen, W., Yamashita, A., & Asama, H. (2015). The influence of action-outcome delay and arousal on sense of agency and the intentional binding effect. *Consciousness and Cognition*, *36*, 87-95. doi: 10.1016/j.concog.2015.06.004.

Yabe, Y., & Goodale, M. A. (2015). Time flies when we intend to act: temporal distortion in a go/no-go task. *Journal of Neuroscience, 35*(12), 5023-5029. doi: org/10.1523 /JNEUROSCI.4386-14.2015.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. Nature, 414(6861), 302–305. http://dx.doi.org/10.1038/35104551.

Yoshie, M., & Haggard, P. (2013). Negative emotional outcomes attenuate sense of agency over voluntary actions. *Current Biology*, *23*(20), 2028-2032. doi: 10.1016/j.cub. 2013.08.03

Appendix B: AQ Questionnaire

The Adult Autism Spectrum Quotient (AQ)

Ages 16+

SPECIMEN, FOR RESEARCH USE ONLY.

For full details, please see:

S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin and E. Clubley, (2001)
The Autism Spectrum Quotient (AQ) : Evidence from Asperger Syndrome/High Functioning Autism, Males and Females, Scientists and Mathematicians
Journal of Autism and Developmental Disorders 31:5-17

Name:..

Sex:..

Date of birth:..................................

Today's Date................................

How to fill out the questionnaire

Below are a list of statements. Please read each statement very carefully and rate how strongly you agree or disagree with it by circling your answer.

DO NOT MISS ANY STATEMENT OUT.

Examples

E1. I am willing to take risks.	definitely agree	slightly agree	[✓ slightly disagree]	definitely disagree
E2. I like playing board games.	definitely agree	[✓ slightly agree]	slightly disagree	definitely disagree
E3. I find learning to play musical instruments easy.	definitely agree	slightly agree	slightly disagree	[✓ definitely disagree]
E4. I am fascinated by other cultures.	[✓ definitely agree]	slightly agree	slightly disagree	definitely disagree

1. I prefer to do things with others rather than on my own.	definitely agree	slightly agree	slightly disagree	definitely disagree
2. I prefer to do things the same way over and over again.	definitely agree	slightly agree	slightly disagree	definitely disagree

3. If I try to imagine something, I find it very easy to create a picture in my mind.	definitely agree slightly agree slightly disagree definitely disagree
4. I frequently get so strongly absorbed in one thing that I lose sight of other things.	definitely agree slightly agree slightly disagree definitely disagree
5. I often notice small sounds when others do not.	definitely agree slightly agree slightly disagree definitely disagree
6. I usually notice car number plates or similar strings of information.	definitely agree slightly agree slightly disagree definitely disagree
7. Other people frequently tell me that what I've said is impolite, even though I think it is polite.	definitely agree slightly agree slightly disagree definitely disagree
8. When I'm reading a story, I can easily imagine what the characters might look like.	definitely agree slightly agree slightly disagree definitely disagree
9. I am fascinated by dates.	definitely agree slightly agree slightly disagree definitely disagree
10. In a social group, I can easily keep track of several different people's conversations.	definitely agree slightly agree slightly disagree definitely disagree
11. I find social situations easy.	definitely agree slightly agree slightly disagree definitely disagree
12. I tend to notice details that others do not.	definitely agree slightly agree slightly disagree definitely disagree
13. I would rather go to a library than a party.	definitely agree slightly agree slightly disagree definitely disagree
14. I find making up stories easy.	definitely agree slightly agree slightly disagree definitely disagree
15. I find myself drawn more strongly to people than to things.	definitely agree slightly agree slightly disagree definitely disagree
16. I tend to have very strong interests which I get upset about if I can't pursue.	definitely agree slightly agree slightly disagree definitely disagree
17. I enjoy social chit-chat.	definitely agree slightly agree slightly disagree definitely disagree
18. When I talk, it isn't always easy for others to get a word in edgeways.	definitely agree slightly agree slightly disagree definitely disagree
19. I am fascinated by numbers.	definitely agree slightly agree slightly disagree definitely disagree
20. When I'm reading a story, I find it difficult to work out the characters' intentions.	definitely agree slightly agree slightly disagree definitely disagree
21. I don't particularly enjoy reading fiction.	definitely agree slightly agree slightly disagree definitely disagree

22. I find it hard to make new friends.	definitely agree	slightly agree	slightly disagree	definitely disagree
23. I notice patterns in things all the time.	definitely agree	slightly agree	slightly disagree	definitely disagree
24. I would rather go to the theatre than a museum.	definitely agree	slightly agree	slightly disagree	definitely disagree
25. It does not upset me if my daily routine is disturbed.	definitely agree	slightly agree	slightly disagree	definitely disagree
26. I frequently find that I don’t know how to keep a conversation going.	definitely agree	slightly agree	slightly disagree	definitely disagree
27. I find it easy to “read between the lines” when someone is talking to me.	definitely agree	slightly agree	slightly disagree	definitely disagree
28. I usually concentrate more on the whole picture, rather than the small details.	definitely agree	slightly agree	slightly disagree	definitely disagree
29. I am not very good at remembering phone numbers.	definitely agree	slightly agree	slightly disagree	definitely disagree
30. I don’t usually notice small changes in a situation, or a person’s appearance.	definitely agree	slightly agree	slightly disagree	definitely disagree
31. I know how to tell if someone listening to me is getting bored.	definitely agree	slightly agree	slightly disagree	definitely disagree
32. I find it easy to do more than one thing at once.	definitely agree	slightly agree	slightly disagree	definitely disagree
33. When I talk on the phone, I’m not sure when it’s my turn to speak.	definitely agree	slightly agree	slightly disagree	definitely disagree
34. I enjoy doing things spontaneously.	definitely agree	slightly agree	slightly disagree	definitely disagree
35. I am often the last to understand the point of a joke.	definitely agree	slightly agree	slightly disagree	definitely disagree
36. I find it easy to work out what someone is thinking or feeling just by looking at their face.	definitely agree	slightly agree	slightly disagree	definitely disagree
37. If there is an interruption, I can switch back to what I was doing very quickly.	definitely agree	slightly agree	slightly disagree	definitely disagree
38. I am good at social chit-chat.	definitely agree	slightly agree	slightly disagree	definitely disagree
39. People often tell me that I keep going on and on about the same thing.	definitely agree	slightly agree	slightly disagree	definitely disagree
40. When I was young, I used to enjoy playing games involving pretending with other children.	definitely agree	slightly agree	slightly disagree	definitely disagree

41. I like to collect information about categories of things (e.g. types of car, types of bird, types of train, types of plant, etc.).	definitely agree	slightly agree	slightly disagree	definitely disagree
42. I find it difficult to imagine what it would be like to be someone else.	definitely agree	slightly agree	slightly disagree	definitely disagree
43. I like to plan any activities I participate in carefully.	definitely agree	slightly agree	slightly disagree	definitely disagree
44. I enjoy social occasions.	definitely agree	slightly agree	slightly disagree	definitely disagree
45. I find it difficult to work out people's intentions.	definitely agree	slightly agree	slightly disagree	definitely disagree
46. New situations make me anxious.	definitely agree	slightly agree	slightly disagree	definitely disagree
47. I enjoy meeting new people.	definitely agree	slightly agree	slightly disagree	definitely disagree
48. I am a good diplomat.	definitely agree	slightly agree	slightly disagree	definitely disagree
49. I am not very good at remembering people's date of birth.	definitely agree	slightly agree	slightly disagree	definitely disagree
50. I find it very easy to play games with children that involve pretending.	definitely agree	slightly agree	slightly disagree	definitely disagree

Developed by:
The Autism Research Centre
University of Cambridge

Appendix C: AQ Sub-scales Correlational Analyses for Experiments 6 and 7

Experiment 6 AQ Sub-scale Analyses

The sub-scales are attention switching, attention to detail, communication, imagination and social skills. There was one significant, medium correlation between the AQ sub-scale of attention to detail and the number of high confidence hits for averted gaze faces, *r*(35)=0.342, *p*=.044. The higher the score on the AQ sub-scale for attention to detail, the greater the number of high confidence hits for recollection of non joint attention faces. This correlation is shown in a scatterplot in Figure A1.

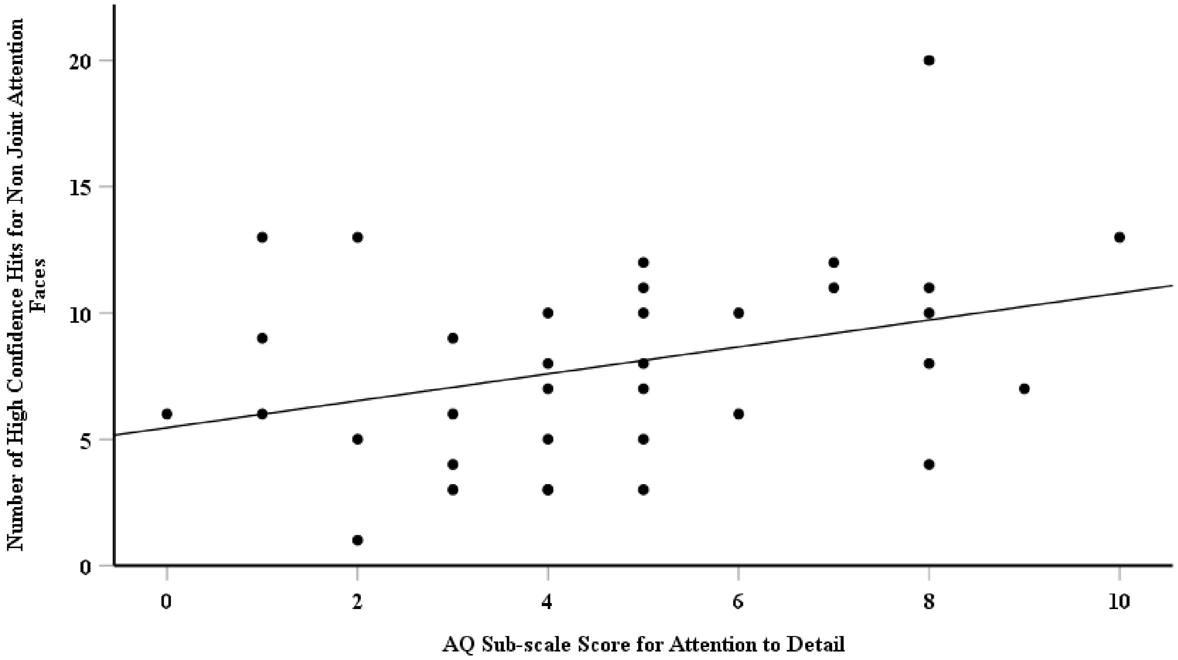

Figure A1. Scatterplot showing the positive correlation between scores on the AQ sub-scale for attention to detail and the number of high confidence hits for non joint attention faces.

There was no correlation between AQ and RT to identify gender collapsed across conditions, *r*(35)= 0.286, *p*= .096, nor between AQ and RT for joint attention faces, *r*(35)= 0.179, *p*= .303. However, there was a medium, positive correlation between AQ and RT for non-joint attention faces, *r*(35)= 0.363, *p*= .032. The greater the AQ score, the slower to identify gender after gaze leading when the response was averted gaze. This correlation is shown in Figure A2. Delving deeper into this finding, the correlation was driven by the AQ sub-scales of attention switching, *r*(35)=0.411, *p*=.014 and social skills, *r*(35)=0.373, *p*=.028.

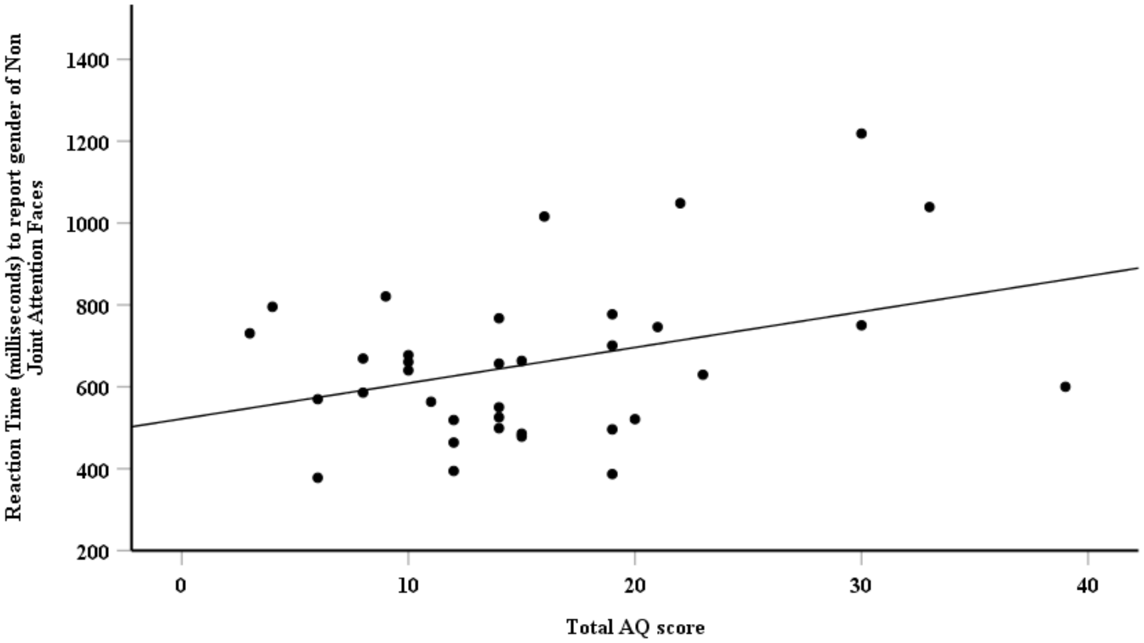

Figure A2. Scatterplot showing the positive correlation between AQ scores and the RT to report the gender of non joint attention faces.

There were no significant correlations between total AQ scores and RT to report old/new for joint or non joint attention faces or foils, or collapsed across conditions or the difference between RT to joint and non-joint attention faces, rs<0.201, ps>.075. However, there were eight significant, medium correlations between the AQ sub-scales of attention switching and attention to detail and RT performance on the old/new task. These correlations are summarised in Table A1.

Table A1.

Descriptive statistics RT (in milliseconds) for old/new task and correlation coefficients with AQ sub-scale scores, with 95% BCa confidence intervals reported in parenthesis. Confidence intervals are based on 1000 bootstrap samples.

Measure	*M*	*SD*	Attention Switching	Attention to Detail
1. RT for JA faces	657 [263, 1657]	294	0.358*	0.341*
2. RT for NJA faces	628 [335, 1174]	204	0.356*	0.448**
3. RT for foils	665 [371, 1423]	247	0.371*	0.430*
4. RT collapsed all conditions	650 [328, 1419]	239	0.376*	0.416*

Abbreviations: M = Mean. SD = Standard Deviation. AQ = Autism Quotient. JA = joint attention. NJA = non joint attention.
*Note: * p <.05, ** p<.01, two-tailed. N*=35.

Experiment 7

AQ Sub-scale Analyses

There were two significant correlations between number of hits in the face recognition task and some of the AQ sub-scales, although the correlation between the attention to detail sub-scale and number of high confidence hits for non joint attention faces in Experiment 6 did not replicate. There was a small to medium, negative correlation between the number of high confidence hits for joint attention faces and score on the imagination AQ sub-scale, *r*(58)= -0.24, *p*=.045. A scatterplot showing this correlation can be found at Figure A3. The higher the AQ score for imagination, the lower the number of high confidence hits for joint attention faces.

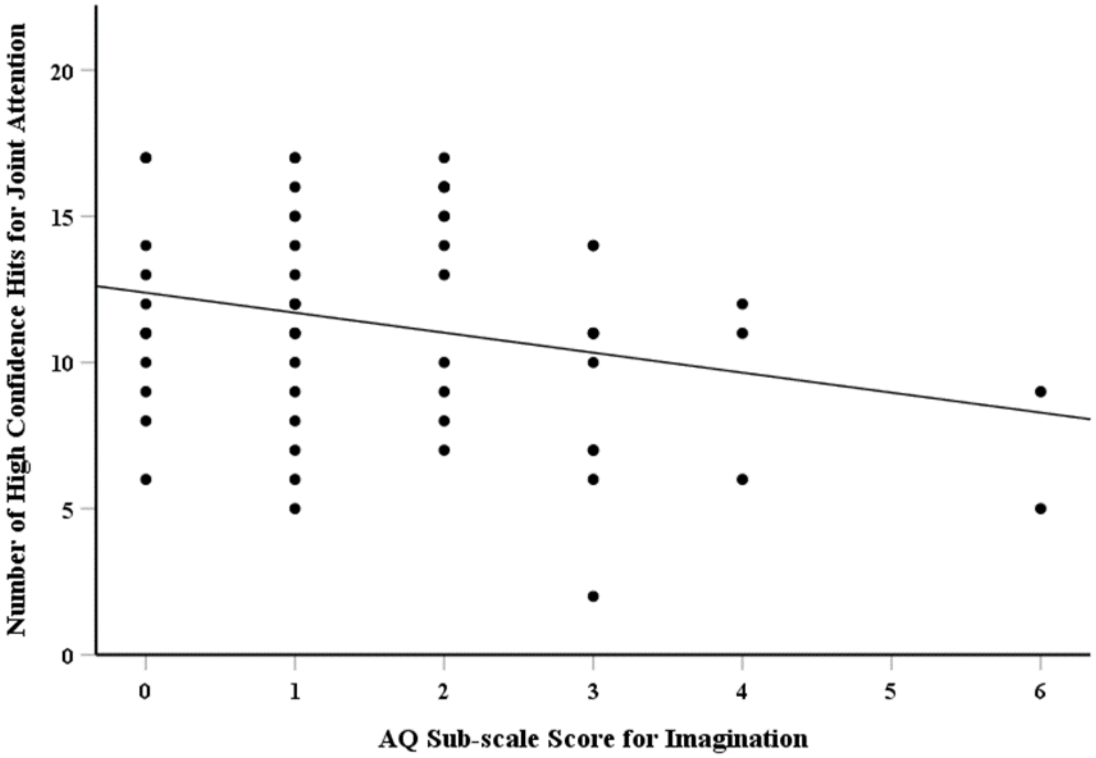

Figure A3. Scatterplot showing the negative correlation between scores on the AQ sub-scale for imagination and the number of high confidence hits for joint attention faces.

There was a medium, positive correlation between the difference between the high confidence hits for joint and non joint attention and score on the social skills AQ sub-scale, *r*(58)= 0.308, *p*=.019. A scatterplot showing this correlation can be found at Figure A4. The greater the difference between number of high confidence hits for joint and non joint attention, the higher the score on the social skills sub-scale.

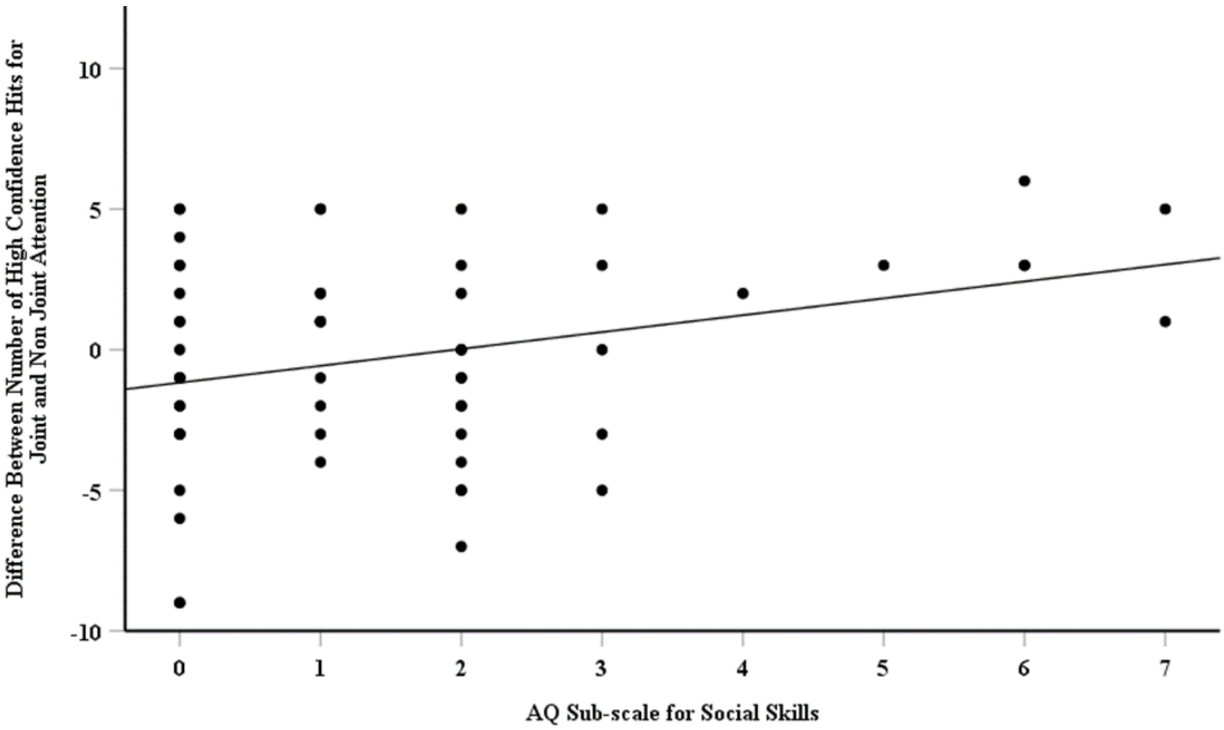

Figure A4. Scatterplot showing the positive correlation between scores on the AQ sub-scale for social skills and the difference between number of high confidence hits for joint and non joint attention faces. Positive differences mean more high confidence hits for joint attention over averted gaze faces. Negative differences mean more high confidence hits for non joint attention over joint attention faces.

There were no correlations between AQ and RT to identify gender collapsed across conditions, for joint attention faces, or non-joint attention faces, *r*s<0.041, *p*s>.761. The medium correlation between AQ and RT for non-joint attention faces found in Experiment 6, therefore, did not replicate. None of the AQ sub-scale correlations with RT for old/new found in Experiment 6 replicated.